Tail and Center Rounding of Probabilistic Expectations in the Health and Retirement Study[1],[2]

Pamela Giustinelli[3]
Department of Economics
Bocconi University

Charles F. Manski
Department of Economics
Northwestern University

Francesca Molinari
Department of Economics
Cornell University

Abstract

We study rounding of numerical expectations in the Health and Retirement Study between 2002 and 2014. We find that the vast majority of respondents provide weakly more refined responses in the tails of the 0-100 scale than in its center and that rounding varies across question domains and respondent characteristics. We exploit response tendencies across questions and waves to infer person-specific rounding in each question domain and scale segment. We replace each point-response with an interval representing the range of possible values of the true latent belief. We compare best-predictor estimates from face-value expectations with those implied by our intervals.

Keywords: Interval data; Subjective Probabilities; Survey data.

[1] **Acknowledgements:** We thank Maura Coughlin, Adam Karabatakis, and Miriam Larson-Koester for able research assistance. We received useful feedback from seminar participants at the HRS work-in-progress series, Bocconi University, University of Southampton, NYU CUSP, University of Michigan, Purdue University, Laval University, University of Oslo, Statistics Norway, University of Munich, and University of Padova, as well as from participants in the 2016 NYFed and ESRC RCMiSoC Workshop on Subjective Expectations and in the 2018 CESifo and ESRC RCMiSoC Workshop on Subjective Expectations and Probabilities in Economics.

[2] **Funding:** This work was supported by: the National Institute on Aging [grant numbers NIA P01 AG10179, P30 AG012846 for Giustinelli]; the National Science Foundation [grant numbers SES 1131500 through the University of Michigan node of the NSF-Census Research Network for Giustinelli, SES 1129475 for Manski, SES 1824575 for Molinari], and the Michigan Institute for Teaching and Research in Economics (MITRE)'s Undergraduate Student Support program (for Giustinelli).

[3] **Contact:** Pamela Giustinelli, Department of Economics, Bocconi University, Via Roentgen 1, 20136, Milan, Italy. Phone: +39 02 5836 3413. E-mail: pamela.giustinelli@unibocconi.it.

1. Introduction

Judgements about the likelihood of future events are an important input for predictions and decisions by citizens, policy makers, and researchers. From the early 1990s on, surveys designed by economists have increasingly measured respondents' subjective expectations for future events using a 0-100 scale of percent chance. This endeavor was prompted by earlier empirical evidence and theoretical arguments demonstrating the greater informativeness of elicited probabilities for binary events relative to verbal "yes/no" intention measures (Juster, 1966; Manski, 1990). This pattern has been repeatedly confirmed in more recent studies; for example, see Delevande and Manski (2010) and Gutsche et al. (2014) in the context of election polling.

The early research using numerical survey expectations devoted substantial effort to evaluating how individuals or groups of interest respond to the questions posed in specific domains and to assess the accuracy of elicited expectations. For example, Dominitz and Manski (1997) use subjective probabilities of absence of health insurance, victimization by burglary, and job loss from the Survey of Economic Expectations to study respondents' perceptions of economic insecurity. Fischhoff et al. (2000) analyze teens' expectations for 18 significant life events using subjective probabilities elicited in the 1997 National Longitudinal Study of Youth. More recent work has increasingly used expectations data to estimate microeconometric models of decision making under uncertainty and to analyze processes of expectations formation. Manski (2004, 2018), Attanasio (2009), Hurd (2009), van der Klaauw (2012), Armantier *et al.* (2013), Delavande (2014), Schotter and Trevino (2014), and Giustinelli and Manski (2018) review the literature from various perspectives.

Questions eliciting expectations on a 0-100 percent-chance scale in principle enable respondents to report beliefs to the nearest 1 percent, encouraging a common rounding convention with minimal data coarsening. But how do respondents use the scale in practice? The accumulated evidence reveals that respondents tend to round their responses. Responses that are not a multiple of 5 or 10 percent occur infrequently. When observed, they tend to occur near the endpoints of the scale to convey very small or large probabilities.

Rounding of expectations poses a series of challenges for statistical inference. First, rounding generates greater data coarsening than intended by the measurement scale. Second, the extent of rounding is not directly observable and may vary across respondents and/or questions. Third, the reasons why respondents round are incompletely understood.

Observed response patterns carry information about respondents' rounding practices, but they do not reveal why respondents round. Manski and Molinari (2010) hypothesize that respondents may round to simplify communication and/or to convey partial knowledge. If respondents round to simplify communication, rounding generates a form of measurement error. However, the structure of the errors produced by rounding is different from that occurring in the classical errors-in-variables model.

Manski and Molinari studied respondent-specific response patterns across all expectations questions asked in the 2006 wave of the Health and Retirement Study (HRS). They found strong evidence of rounding, with the extent differing across respondents. They proposed use of a person's response pattern across questions to infer the person's rounding practice, the result being interpretation of reported numerical values as interval data.

In this paper, we significantly expand study of respondent-specific rounding patterns by analyzing responses across all expectations questions asked in the core HRS questionnaire between 2002 and 2014. This enables us to learn important new features of rounding practices.[4]

Section 2 presents the main findings of our data analysis, with Supplementary Appendices reporting further details. We initially study each wave of the HRS separately and find that the respondent-specific rounding patterns reported by Manski and Molinari (2010) are stable across waves. We then pool data across waves. This yields rich respondent-specific data that enables us to probe more deeply than the earlier study.

We discover a tendency by about half of the respondents to provide more refined responses in the tails of the 0-100 scale than the center. In contrast, only about five percent of the respondents give more refined responses in the center than the tails. We find that respondents tend to report the values 25 and 75 more frequently than other values ending in 5. We also find that rounding practices vary somewhat across question domains, which range in the HRS from personal health to personal finances to macroeconomic events.

Based on our examination of rounding practices in Section 2, Section 3 develops a framework that interprets each numerical response given by a respondent as an interval. We propose a two-stage algorithm. The first stage classifies each respondent into one of a set of mutually exclusive and exhaustive rounding types and places an upper bound on the amount of rounding each respondent is inferred to apply when reporting their expectations. The second stage assigns an interval to each of the respondent's original point responses, which represents

[4] The HRS has measured probabilistic expectations biannually since 1992; Juster and Suzmann (1995) describe the initial design. Section P of the core questionnaire has been devoted to expectations measurement, each wave including about 25 to 35 questions spanning different domains of personal and macroeconomic uncertainty. From 2002 on, expectations have been consistently elicited on a 0-100 percent-chance scale, with many questions repeated across waves.

the range of values in which the respondent's underlying true belief is plausibly deemed to lie based on the respondent's inferred rounding type. These intervals can be interpreted as a measure of informativeness, or quality, of an increasingly used type of data (numerical probabilistic expectations) from an important data source (the HRS).

Our approach accommodates substantial heterogeneity in rounding practices. Within a specific question domain, a respondent's rounding type is a bivariate vector of the form (tail, center) rounding, partitioning the 0-100 scale into two symmetric tails (0-24 and 76-100) and a center (25-75). Thus, in addition to being person specific, the inferred degree of rounding may differ between tails and center and may vary across question domains. The assigned intervals vary across respondents and across values of the observed point responses.

We use our framework to study how rounding tendencies vary with observable characteristics of the respondents. We find that higher levels of educational attainment and of cognition are associated with a tendency to give more refined responses (less rounding) across all scale segments and question domains. On the other hand, the association of rounding with age appears to be non-linear, with youngest (50-59) and oldest (80+) respondents displaying a higher tendency to round than respondents in the intermediate age groups (60-69 and 70-79).

While this paper studies rounding as a subject of intrinsic interest, a reader may naturally ask how the interval data that our proposed approach generates from point responses may be used in statistical analyses. This matter has been addressed in the econometric literature studying conditional prediction with interval measurement of outcomes and/or covariates; see Manski and Tamer (2002) and Beresteanu and Molinari (2008). Manski and Molinari (2010) provide an illustrative application of best linear prediction of subjective expectations of HRS respondents for survival to age 75 with interval-measured expectations outcomes.

Section 4 demonstrates how interval data on subjective expectations can be employed as either an outcome variable or a covariate in prediction analyses of substantive interest. One application considers best linear prediction of the labor supply expectations of working HRS respondents, conditional on specified covariates. A second application uses longevity expectations and other covariates to predict hours worked.

As far as we are aware, only two previous papers systematically study rounding of responses to probabilistic expectations questions. One is Manski and Molinari (2010), on whose work we build. The other is Kleinjans and van Soest (2014), who develop and estimate a panel-data structural econometric model to analyze response patterns to each of six expectations questions in the HRS. Their analysis aims to investigate the extent to which probability reports are determined by genuine underlying probabilistic beliefs, rounding, a tendency to give so-called "focal" responses of (0, 50, 100), and selective item non-response. Despite the very different approaches taken, they find, as we do, that tendencies to round, give "focal" responses, and not respond tend to be persistent over time.[5]

Beyond readers who have interest in expectations data, we anticipate that general survey researchers will find this paper useful. Our study of tendencies to round responses to expectations questions should heighten concern that respondents may round responses to numerical questions in other contexts. Consider, for example, questions asking respondents to state their income or the number of hours they worked in the past week. Respondents may round

[5] Some authors have devoted special attention to responses of 0, 50, and 100 percent. Fischhoff and Bruine de Bruin (1999) and Bruine de Bruin *et al*. (2002) hypothesize that some respondents use 50 percent to signal epistemic uncertainty. Lillard and Willis (2001) and Hudomiet and Willis (2013) conjecture that respondents form full subjective distributions for the probability of an event and then report whichever of the values (0, 50, 100) is closest to the mode of their distribution. We analyze reports of (0, 50, 100) percent jointly with responses to the entire set of expectations questions asked.

their responses, with the extent of rounding differing across persons. Examination of a person's response pattern across different numerical questions, in the manner that we do here, may provide a credible way to infer that person's rounding practice. One may then interpret reported numerical values as intervals.

Some surveys elicit interval data directly. For example, the HRS uses *unfolding bracket* questions to enable respondents who are not willing to provide exact information about their income and assets to indicate whether the quantities of interest lie above or below a sequence of specified thresholds. Similarly, the Occupation Employment Statistics (Bureau of Labor Statistics, 2018) collects wage information in interval form. These intervals, too, can be analyzed using econometric methods for interval data referenced earlier.

Our interpretation of rounded responses as interval data provides an interesting counterpoint to previous statistical research on data coarsening (e.g., Heitjan and Rubin (1991), Heitjan (1994), and Gill et al. (1997)). In that literature, it is assumed that the researcher observes a random set $\mathcal{X}$ (an interval, group, or partial categorization) to which an unobservable random variable of interest x belongs with probability one. An assumption of "coarsening at random" is imposed, requiring that the probability of observing $\mathcal{X} = \mathrm{A}$ given $x = \mathrm{x}_0$ is constant for all x_0 in A, where A denotes a subset of the support of x. In contrast, the algorithm that we propose constructs sets $\mathcal{X}$ based on respondents' point responses and their tendencies for rounding across the entire set of questions eliciting subjective beliefs. Our approach does not assume ignorability of the coarsening mechanism and it allows for a coarsening mechanism that differs among respondents.

2. Exploratory Analysis of Response Patterns across Questions and Waves in the HRS

Since 2002 the HRS has devoted Section P of its core questionnaire to measurement of expectations in the domains of personal health, personal finances, and general economic conditions. Across seven biannual waves spanning 2002 to 2014, expectations have been elicited on a 0-100 percent chance scale. Several questions have been repeated across multiple waves. Table 1 shows the questions, organized by domain and the waves in which they were asked.

The number of questions per wave ranges between 22 in 2002 and 38 in 2006. Most questions are in the personal finances domain (between 11 and 23 per wave, 31 overall), followed by the personal health domain (between 3 and 9 per wave, 10 overall), and the domain of general economic conditions (between 2 and 7 per wave, 12 overall). A subset of 12 questions across the three domains were asked in all waves.

The number of responses varies across questions and waves, ranging from about 5,000 to 30,000 responses per question in each wave. The variation across questions stems from the fact that the HRS makes extensive use of skip sequencing. Thus, whether a respondent is asked a specific question depends on the previous answers given by the respondent and on whether the event specified by the question is relevant to the respondent.

The total number of responses generated by a question across the seven waves varies because questions have been added and removed over time. It also varies due to changes in sample composition across waves. The HRS sample has periodically been augmented with new cohorts of respondents who joined the study in specific waves. Respondents exit the study due to attrition or death.

Section 2.1 studies response patterns across questions in each wave, alternatively using all questions asked in the wave and the twelve questions asked in all waves. Focusing on the latter questions, we analyze the stability of response tendencies across pairs of waves. Supplementary Appendix A2 provides further detail, investigating patterns of response to specific questions. To ensure comparability with the analysis of the 2006 data by Manski and Molinari (2010), we condition also our analysis on respondents aged 50 or older.

Having established the temporal stability of rounding practices, Section 2.2 pools the HRS data across waves and analyzes response patterns separately by question domain. We pay particular attention to the location of responses inside the 0-100 scale and learn important features of respondents' response patterns in specific domains.

Throughout the paper, the notation $\mathbb{M}10$ and $\mathbb{M}5$ denotes responses that are multiples of 10 or 5 other than (0, 50, 100). When responses are non-rounded values, we distinguish those in the outer tails of the scale 1-4 and 96-99, from those between 6 and 94. NR denotes nonresponse.

2.1. Temporal Stability of Response Tendencies

2.1.1. Response Tendencies in Each Wave

Table 2 shows the fractions of respondents displaying each of seven mutually exclusive and exhaustive response patterns, progressing left to right from the most rounded to the least rounded. Column 3 gives the fraction of respondents who respond to no questions in the wave, coded in the HRS as "Don't know" or "Refuse." Column 4 gives the fraction of respondents who, when they respond, only use the values 0 and 100 in the corresponding wave. Column 5 gives the fraction who only use the values (0, 50, 100). Columns 6 and 7 give the fractions of respondents who answer at least one question with a value in $\mathbb{M}10$ and $\mathbb{M}5$ respectively. Column 8 gives the fraction of respondents who respond to at least one question with a non-round value in 1-4 or 96-99. Column 9, labelled “Some other,” gives the fraction who respond at least once with a non-round value in 6-94.

The set of expectations questions varies across waves. The top panel of Table 2 presents a version of the statistics where respondents are classified into one of the seven response patterns using only the twelve questions that were asked in all seven waves. The bottom panel uses the responses to all questions asked in a wave.

A very small fraction of respondents answer none of the questions posed to them. This fraction ranges between 0.009 and 0.027, depending on the set of questions used to classify respondents. Between 0.019 and 0.101 of respondents uses only the values (0, 100). Similar fractions of respondents use only the values (0, 50, 100). Most respondents give at least one answer in $\mathbb{M}10$ or in $\mathbb{M}5$. The fraction of $\mathbb{M}10$ respondents ranges between 0.263 and 0.337 across waves when all questions asked in a wave are used for classification and between 0.392 and 0.458 when only the questions common to all waves are used. Similarly, the fraction of $\mathbb{M}5$ respondents ranges between 0.427 and 0.513 when all questions are used for classification and between 0.295 and 0.353 when only the common set is used.

The fractions of respondents who give at least one response in the outer tails (1-4 or 96-99) or non-rounded values in 6-94 are sizeable but considerably smaller, especially the latter. The former fraction ranges between 0.101 and 0.144 when all questions are used for classification and between 0.054 and 0.092 when only the common set is used. The latter fraction ranges between 0.022 and 0.042 or between 0.011 and 0.020, depending on the set of questions used.

2.1.2. Transitions of Response Tendencies across Waves

The main message of Table 2 is that the response patterns found by Manski and Molinari (2010) in the 2006 wave of the HRS hold throughout the seven waves between 2002 and 2014. However, these are aggregate patterns that may partly be susceptible to variation across waves in sample composition. To address this issue, we compute transition matrices of response tendencies across waves. Specifically, for each pair of waves indicated by column, Table 3 reports the fractions of respondents classified as belonging to any rounding category in the first wave who transitioned to: the same rounding category in the second wave (1^{st} row), a finer or coarser adjacent category (2^{nd} row), and a more distant rounding category (3^{th} row). The reported calculations use the twelve questions in common to the seven waves.

We find that between 0.406 and 0.436 of the respondents remain in the same rounding category across any pair of adjacent waves. Between 0.373 and 0.386 transition to an adjacent category. Thus, between 0.788 and 0.813 of the respondents transition to the same or an adjacent category. Even transitions between the first and last waves, with fourteen years separating them, display high persistence, with over 0.78 of the respondents transitioning to the same or an adjacent category.

2.2. Pooling Data across Waves to Probe More Deeply into Response Tendencies

With temporal stability established, we henceforth pool the HRS data across waves. This greatly increases the number of expectations responses observed per respondent, multiplying it sevenfold for respondents interviewed in all waves between 2002 and 2014. Across all questions and waves, the average number of responses per respondent is 106.8. By domain, this figure ranges from 19.1 for personal health to 66 for personal finances. With such rich respondent-specific data, we can probe more deeply into rounding practices.

To obtain further insight, we scrutinize the rounding behavior of 100 respondents drawn at random and find two highly interesting patterns. First, a substantial fraction of respondents round more coarsely in the center of the 0-100 scale than in the tails. Second, respondents tend to use the percent-chance values 25 and 75 more than they do other values ending in 5. Supplementary Appendix SA2.2 describes the analysis underlying these findings.

Our study of 100 randomly drawn respondents does not reveal how prevalent the discovered features are across the whole sample of HRS respondents. To answer this question, we now refine our earlier categorization of rounding patterns. We define the center (C) of the percent-chance scale to be values in the range 26-74 and the tails (T) to be values in the ranges 0-24 and 76-100. The values 25 and 75 form the boundary between the tail and center. We group responses into nine categories, defined by their presence in T or C and by the degree to which they are multiples of smaller numbers. The categories are: $\mathbb{M}1$-T ≡ values in 1-24 or 76-99 that are not multiples of 5; $\mathbb{M}1$-C ≡ values in 26-74 that are not multiples of 5; $\mathbb{M}5$-T ≡ {5, 15, 85, 95}; $\mathbb{M}5$-C ≡ {35, 45, 55, 65}; $\mathbb{M}10$-T ≡ {10, 20, 80, 90}; $\mathbb{M}10$-C ≡ {30, 40, 60, 70}; $\mathbb{M}25$ ≡ {25, 75}; $\mathbb{M}100$ ≡ {0, 100}; $\mathbb{M}50$ ≡ {50}.

With this categorization, Table 4 shows the distribution of responses across respondents for each question asked in Section P between 2002 and 2014. The two main features detected by inspecting the random sample are decisively confirmed in the general HRS sample. Comparison of the frequencies of $\mathbb{M}25$ responses (in column 5) with the frequencies of the remaining $\mathbb{M}5$ responses ($\mathbb{M}5$-C in column 9 and $\mathbb{M}5$-T in column 8) reveals that the fraction of {25, 75} responses is always higher than the fraction of responses ending in 5 in the center of the scale

(35, 45, 55, 65). For most questions across the three domains, the fraction of {25, 75} responses is higher than the fraction of responses ending in 5 in the tails of the scale (5, 15, 85, 95).

Even more striking is comparison of the frequencies of responses in the tails versus those in the center. The fractions of $\mathbb{M}10$, $\mathbb{M}5$, and $\mathbb{M}1$ responses in the tails are higher than the corresponding fractions in the center for nearly all questions in Table 4 (but P47 and P190).

3. Transforming Expectations Responses into Interval Data

Generalizing the inferential approach proposed by Manski and Molinari (2010), this section develops a new algorithm that uses the response tendency of a respondent that we have documented in the previous sections to characterize rounding of responses to particular questions. The algorithm classifies each respondent into one of a set of mutually exclusive and exhaustive rounding types and transforms each original point response into an interval where the true latent belief is deemed to lie.

Our algorithm relies on considerably weaker and more credible assumptions than inference that uses expectations reports at face value. Nevertheless, we cannot be certain that the intervals we construct are accurate. The algorithm is subject to two potential forms of misclassification. First, a given survey response may be less rounded than the interval assigned by the algorithm; that is, the actual rounding interval may be a subset of the algorithm's interval. Then our use of the data is correct, but it yields inference that is less sharp than it would be if the true degree of rounding were known. Second, the actual rounding interval may not be completely contained in the algorithm's interval. Then the actual belief may lie outside our interval, making our use of the data incorrect. Still, use of the algorithm substantially lowers the risk and severity of the latter type of error relative to the standard approach that takes survey responses at face value.

The new algorithm embodies the data patterns documented in Section 2. Section 3.1 describes the determination of a respondent's rounding type. Section 3.2 presents the empirical distribution of the respondents' rounding types and studies how rounding tendencies vary with respondents' characteristics. Section 3.3 explains how a respondent's point response to a specific question and their rounding type are used to construct the interval associated with the observed point response.

3.1. Determination of Respondent Rounding Types

Based on the evidence in Section 2, we allow a respondent's rounding type to vary across question domains and between the tails and center of the measurement scale. Thus, within a specific domain of questions, a respondent's rounding type is a bivariate vector of the form (tail, center) rounding, partitioning the 0-100 scale into two symmetric tails (0-24 and 76-100) and a center (25-75). We believe that our specific choice of tails and center reasonably reflects the empirical patterns of HRS responses, but judgments need not be uniform. The algorithm can be easily adapted to different definitions of tails and center or extended to accommodate finer partitions of the 0-100 scale (e.g., outer tails, inner tails, center).

The new algorithm refines the earlier one posed by Manski and Molinari (2010) in multiple ways. One refinement is to separate tail from center rounding. Another is to classify persons who only use the response values (0, 25, 50, 75, 100) as rounding to the nearest 25 percent rather than to the nearest 5 percent. A further difference between the two algorithms is that here we use a tighter criterion for assignment of a person to a more refined rounding type.

To explain the tighter criterion, consider categorization of a respondent as one who rounds to the nearest 10 percent (or to a more refined degree). Manski and Molinari assigned a respondent to this rounding type if all responses are multiples of 10 and at least one response is not a value in (0, 50, 100). We use here a tighter criterion that requires observation of at least two responses that are multiples of 10 other than (0, 50, 100), of which one must be in the domain under consideration and the other may be in a different domain and may also be a less rounded response.

Adding the new requirement reflects our desire for further credibility when assigning a person to a more refined rounding type. We want enhanced credibility because misclassification into an overly refined rounding category yields an inferential error, as the person's latent beliefs may not entirely lie within the overly refined interval. Misclassification of a person into a rounding category less refined than their actual one does not yield an inferential error, as the less refined interval includes the actual one as a subset.

The main criteria for classification of respondents are as follows:

- **Center rounding type** Define x_n in {1, 5, 10, 50}, with *n* = 1, . . . , 4. Respondent *j* is classified as rounding to the nearest x_n percent in the center within question domain *l* if one of the following two conditions holds: (i) they are observed to give at least two answers in the center that are multiples of x_n percent but not of $x_{n'}$ for any *n'* < *n* within domain *l*; or (ii) they are observed to give one answer in the center that is a multiple of x_n percent (but not of $x_{n'}$ for any *n'* < *n*) within domain *l* AND at least one answer in the center that is a multiple of $x_{n'}$ for any *n'* ≤ *n* within a second domain *l'* distinct from *l*.
- **Tail rounding type** Respondent *j* is classified as rounding to the nearest x_n percent in the tails within question domain *l* if one of the following two conditions holds: (i) they are observed to give at least two answers in the tails that are multiples of x_n percent but not of $x_{n'}$ for any *n'* < *n* within domain *l*; or (ii) they are observed to give one answer in the tails

that is a multiple of x_n percent (but not of $x_{n'}$ for any *n'* < *n*) within domain *l* AND at least one answer in the tails OR center that is a multiple of $x_{n'}$ for any *n'* ≤ *n* within a second domain *l'* distinct from *l*.

To illustrate, consider a respondent who has answered four expectations questions in the domain of personal finances, either within the same wave or over multiple waves. Two of the observed responses belong to the tails, {5, 85}, and two to the center, {30, 60}. As the set of responses includes two multiples of 5 percent in the tails and two multiples of 10 percent in the center, our algorithm classifies this respondent as one rounding to the nearest 5 percent, *or to a finer degree*, in the tails ($\mathcal{M}$5-T) and to the nearest 10 percent, *or to a finer degree*, in the center ($\mathcal{M}$10-C).

The Supplementary Appendix SA3.1 provides additional and more complex examples. It also presents the complete algorithm in a formal and compact way.

3.2. Empirical Distribution of Rounding Types and Association with Observable Characteristics

We apply the algorithm to all HRS respondents who responded to at least one expectations question in any question domain and in any wave between 2002 and 2014. Table 5 reports the empirical distribution of rounding types for each domain of questions. Depending on the domain, between 40.40% and 61.03% of respondents are inferred to apply finer rounding in the tails than in the center. Between 28.49% and 38.73% of respondents apply the same degree of rounding in the tails and in the center. Between 2.90% and 6.71% of respondents apply coarser rounding in the tails than in the center.

The rounding type of a small minority of respondents could not be determined either in the tails or in the center or both. Most undetermined cases occur when, for a given respondent, we do not observe any answer in the relevant domain and scale segment. Among respondents for whom we observe at least one answer in the relevant domain and scale segment, all cases of undetermined tail rounding type disappear and only a few cases of undetermined center rounding type remain. The latter are respondents for whom we only observe one answer in the center in the relevant domain and no answers in the center in the remaining two domains.

We now investigate how rounding types vary with observable respondent characteristics. We summarize the data using parametric bivariate ordered probit regression, which embodies the basic ordinal property that our rounding categories display across different degrees of rounding.

Table 6 presents estimated coefficients of three bivariate ordered probit regressions, one per question domain. The outcome variables are the respondent's bivariate vectors of tail and center rounding categories in each domain. As predictors, we use binary variables for respondent's gender (male, with female omitted), educational attainment (high school, some college, bachelor, and graduate, with less than high school omitted), and race (black and other, with white omitted).

We also include information on individual's age and cognitive functioning. While these variables are time-varying for each respondent, our analysis in Section 2.1.2 and 2.2 supports treating respondent rounding behavior as fixed over time. We therefore account for age variation across respondents by incorporating in our bivariate ordered probit regressions an indicator of whether each respondent's cross-wave average age lies in the categories 60-69, 70-79, and 80+ years, with 50-59 the omitted category. We account for variation in cognitive functioning across respondents by including each respondent's cross-wave average cognitive score. See Fisher et al.

(2012) and Crimmins et al. (2011) for a description and an empirical assessment of the HRS cognitive measures.

The cognitive score has a range of 0-35. In our data, the respondent-specific cross-wave average cognitive score has a mean of 23 and a standard deviation of 4.11 across respondents. The respondent-specific cross-wave standard deviation in cognitive score has a mean of 2.9 across respondents. The fact that the standard deviation of the cross-wave average score is larger than the average cross-wave standard deviation in the score lessens our concerns for using a time-fixed measure of cognitive functioning in our bivariate ordered probit regressions. Nonetheless, the time variation in cognitive score and its association with rounding warrant study in future research.

The model permits the error terms of the latent variables underlying the inferred tail and center rounding categories to be correlated with each other. The correlation parameter, ρ, is estimated along with the other coefficients. The rounding categories are ordered from least coarse to most coarse. Thus, positive associations indicate a tendency to round more coarsely.

We estimate the parameters by maximum likelihood using the Stata package described in Sajaia (2008). The bivariate probit model and other multivariate discrete outcome models are discussed, for example, in Amemiya (1981); see our Supplementary Appendix SA3.2 for a brief summary. Estimated coefficients with standard errors are reported in Table 6. Table 7 reports predicted probabilities of selected tail and center rounding types for persons with specified covariate values.

We find that higher levels of educational attainment and of person-specific average (cross-wave) cognitive score are associated with a tendency to give more refined responses across all scale segments and question domains. The patterns for the other predictors are more varied.

For example, respondents in the oldest age category (80+) tend to give more rounded responses than respondents belonging to the youngest one (50-59) across all scale segments and questions domains. On the other hand, respondents in the two intermediate age groups (i.e., 60-69 and 70-79) belong to rounding categories that may be more refined, coarser, or statistically indistinguishable from those characterizing younger respondents, depending on the specific domain or scale segment.

A potential interpretation of the observed age patterns is that individuals belonging to the intermediate age groups may have more direct experience and hence better knowledge of the topics covered by the questions than younger respondents, generating more refined responses among the middle groups. On the other hand, individuals of older age might already, on average, have lower cognitive functioning, leading to coarser responses. This pattern, however, continues to hold after conditioning on the respondent's average (across waves) cognitive score. Parameter estimates for a specification without cognitive score are shown in the Supplementary Appendix.

Male respondents tend to round more coarsely than female respondents in the personal health and personal finances domains, but only in the tails. On the other hand, male respondents tend to round less coarsely than women respondents in the center in the domain of general economic conditions. While respondents belonging to the residual race category (including Hispanic, Asian, and Pacific Islander) tend to round more coarsely than white respondents, the differential rounding tendencies of black respondents relative to white respondents vary across question domains and scale segments.

The large, positive, and statistically significant estimates of the correlation parameter ρ reveal that rounding tendencies are positively correlated across scale segments. Hence, respondents who give coarser responses in the tails are more likely to do so in the center.

3.3. Using Survey Responses and Rounding Types to Form Expectations Intervals

It is natural to wonder the extent to which failing to account for rounding might lead to inaccurate conclusions when analyzing data. A simple numerical illustration pertaining to the analysis of the effect of longevity expectations on hours worked shows that ignoring rounding may yield highly inaccurate conclusions.

Suppose that two respondents both round their response to the longevity expectation question to the closest multiple of 25. Suppose that one respondent views their probability to live past age 75 to be forty percent while the other respondent views it to be sixty percent, with the latter working significantly more hours as a consequence. With rounding, both respondents report their probability to live past age 75 as fifty percent. The notable difference in hours-worked outcomes with apparently the same expectations may be misinterpreted as caused by unobserved heterogeneity in labor-leisure preferences, when the actual cause is different longevity expectations.

Next, consider a scenario where the first respondent views their probability to live past age 75 to be thirty-seven percent while the other respondent views it to be thirty-eight percent, with the latter working slightly more hours. With rounding, the first respondent reports a probability to live past age 75 of twenty-five percent, and the second respondent reports fifty percent. The slight difference in outcomes with an apparent large difference in expectations may be misinterpreted as evidence of minimal effect of expectations on labor supply.

These examples, while stylized, illustrate that ignoring rounding might lead to “boundary mistakes;” that is, to significantly underestimating or overestimating an effect of interest. We therefore propose an algorithm that uses the information contained in each respondent’s reporting behavior across the survey, as analyzed in the preceding sections, to transform observed percent-chance point reports into intervals.

Here we present the construction of interval data within the context of the illustration introduced in Section 3.1. The Supplementary Appendix SA3.4 discusses more complex cases, presents the complete algorithm formally, and reports the distributions of interval width for the responses given to specific questions. Compared with MM10, our algorithm typically yields narrower intervals.[6]

In the example introduced in Section 3.1, the respondent is observed to answer with {5, 30, 60, 85} to four expectations questions concerning personal finances and is classified to be of rounding type ($\mathcal{M}$5-T, $\mathcal{M}$10-C) in that domain. Because the respondent is classified to round to the nearest 5 percent in the tails, the algorithm assigns to each of the respondent's point responses in the tails an interval of width 5 centered around the point response. Specifically, the algorithm assigns the interval [2.5, 7.5] to response 5 (i.e., 5 ∓ 2.5) and the interval [82.5, 87.5] to response 85 (i.e., 85 ∓ 2.5). Similarly, as the respondent is classified to round to the nearest 10 percent in the center, the algorithm assigns interval [25, 35] to the 30 percent response (i.e., 30 ∓ 5) and the interval [55, 65] to the 60 percent response (i.e., 60 ∓ 5).

In general, construction of intervals around point responses near the thresholds which separate the center from the tails (25 and 75 percent) requires specific "boundary conditions." Such conditions are not binding in this example. We explain them in the Appendix.

By construction, each interval contains the point response because the former is centered around the latter. Moreover, the interval is assumed to cover the unobserved true latent belief

[6] This may not always be the case, however, because our algorithm requires at least two instances or responses of a certain refinement level (e.g., multiples of 5) to establish a rounding type, rather than just one as in MM10.

with certainty. However, no assumption is made about the location of the true latent belief inside the interval.

Our algorithm relies on considerably weaker and hence more credible assumptions than inference using expectations reports at face value. At the opposite extreme, one could be ultraconservative, maintaining that each point response is consistent with any amount of rounding. One would then replace all reported expectations with a [0, 100] interval. Obviously, doing this empties the data of any information content.

Our choice of assumptions used to identify respondents' rounding types and bound their unobserved true beliefs strikes a balance between those two extremes and is informed by the respondents' response patterns across HRS questions and waves, which we have documented in this paper. A researcher entertaining a different set of assumptions about how survey respondents round their expectations reports could easily apply our framework by simply replacing our assumptions with theirs. In general, stronger and/or more numerous assumptions will yield (weakly) narrower intervals.

4. Illustrative Applications

This section demonstrates how interval data on subjective expectations can be employed as either an outcome variable or a covariate in prediction analyses of substantive interest. Section 4.1 presents an application where the objective is to predict the labor supply expectations of working HRS respondents conditional on specified covariates. Section 4.2 studies prediction of hours worked of male HRS respondents, using longevity expectations as a covariate. In both cases we examine how accounting for rounding in probabilistic expectations affects the conclusions that one can draw in empirical analysis.

4.1 Predicting Labor Supply Expectations of Older Workers

As the American population ages and a larger fraction of "baby boomers" approach retirement age, it of interest to analyze how subjective expectations of HRS respondents for working full-time past age 62 vary with several covariates, including age, gender, coupledness status, household wealth, race, and education.

In each of the HRS waves analyzed in this paper, respondents younger than 62 at the time of the interview were asked, *"Thinking about work in general and not just your present job, what do you think the chances are that you will be working full-time after you reach age 62?"*. See question P17 in Table S2 for the response distribution in each wave and in Table 4 for the response distribution with data pooled across waves. We compare the conclusions drawn when the elicited expectations are taken at face value, as is commonly done in the related literature (e.g., Honig, 1996, 1998), and when our algorithm is used to characterize rounding. We analyze data from each of the seven waves of the HRS from 2002 to 2014, pooling the data across waves. This yields a sample of size 24,052 after dropping respondents who are younger than fifty and those for whom we do not observe some covariates.

When we take the elicited expectations of working past age 62 at face value, we report the results of best linear prediction under square loss. In this case, we assume that nonresponse is random and drop respondents who answered "Don't know" or "Refuse" to the probability chance question posed in P017. The pooled sample has size 23,811.

When we use our algorithm to interpret the elicited expectations as intervals under the assumptions set forth in Section 3, we repeat the same exercise of best linear prediction under square loss, considering all points in the interval outcome variable of each respondent to be feasible values of the quantity of interest. In this case, the resulting best linear predictor's parameter vector is not point identified. Rather, it is *partially identified*, meaning that there is a *set* of values (rather than a single value) for the parameter vector that are consistent with the available data and maintained assumptions. This set of values is called the parameters' *identification region*. We estimate the identification region and report confidence intervals for it using the method proposed by Beresteanu and Molinari (2008) and the Stata package by Beresteanu et al. (2010). Beresteanu and Molinari (2008, Section 4) and Beresteanu et al. (2012, Section 3.2) give a detailed discussion of the method.

The results of our analysis are reported in Table 8. The first column shows the estimates and confidence intervals when elicited expectations are taken at face value. The results suggest an increased expectation to work full-time past age 62 for individuals who are closer to age 62, who are males, who have lower wealth, and who are more highly educated, while a reduced expectation to work past age 62 for wealthier individuals and for non-whites.

The second through fifth columns report set estimates and confidence intervals when elicited expectations are interpreted as interval data according to our algorithm. The only difference between the empirical exercises reported in the two sets of columns (2-3 and 4-5) is that the set estimation in columns 2-3 maintains the assumption that nonresponse to the expectation question is random. This is done exclusively to provide intermediate results based on the same sample as that used in column 1, but we consider the assumption unrealistic in the present application. Hence, we focus on the results in columns 4-5.

The results reveal that the strength of the conclusions that can be drawn is weaker when we interpret elicited expectations as intervals than when we take them at face value. This is to be expected, as there is an intrinsic trade-off between the strength and the credibility of inference. Despite this, our analysis –under considerably weaker assumptions– continues to find that males and individuals with higher education have higher expectations, while blacks have lower expectations, to work past age 62. Interestingly, the interval data that we construct remains sufficiently informative to allow us to learn the sign of several coefficients of the best linear predictor. We conjecture that this may not be the case, had we used the algorithm in Manski and Molinari (2010), which is based on a single wave of data and tends to give wider intervals.

4.2 Longevity Expectations and Hours Worked

Individuals' life horizon and the related mortality risk are key ingredients of economic models of life-cycle behaviors. This raises the question of whether life horizon and mortality risk as *perceived* by individuals are empirically important determinants of their labor supply, saving and investment decisions, etc. (e.g., Hamermesh (1985)). Previous work has examined the effect of subjective survival probabilities on retirement and Social Security claiming behaviors of older Americans (e.g., Hurd et al. (2004), Delavande et al. (2006)). Here we focus on the relationship between subjective survival probabilities and hours worked.

In all waves of the HRS, respondents under the age of 65 were asked to report their longevity expectations by means of the following question: *"What is the percent chance that you will live to be* 75 *or more?"* (question P28). The sample distribution of responses to P28 in each wave is displayed in Table S2. Table S8 reports the sample frequencies of the width of the algorithm intervals $[\upsilon_{jkt}^{L}, \upsilon_{jkt}^{U}]$, constructed around respondents' point responses to question P28 in the 2014 HRS wave.

We focus on working male individuals aged 50 through 64, who were asked to report their percent chance of living past 75. Our outcome variable is weekly hours worked. Hours worked were measured in question J172 as following: *"How many hours a week do you usually work [on* *this job/in this business]?"* This question was asked only of respondents who answered "yes" to question J20, *"Are you doing any work for pay at the present time?"*.

The predictors used are interval-valued longevity expectations, age, and coupledness status. As in the first application, the exercise is best prediction of the outcome variable given covariates. We again are interested in comparing the conclusions that can be drawn when rounding is addressed with those obtained when rounding is ignored. Econometrically, the key difference between this application and the earlier one is that now the interval-valued variable is used as a covariate. In this case, the inferential problem is more difficult than when the interval-valued variable is used as an outcome of a regression model, because the estimator is no longer linear in the interval data.

Manski and Tamer (2002) study the problem of inference on regressions with interval data on a regressor. That is, the problem is one of inferring, say $E(y\,|\,\upsilon, x)$, when υ is unobserved but is known to lie in some interval $[\upsilon^L, \upsilon^U]$ with probability 1. The latter assumption is called Interval (I). Under assumption (I), two additional ones – Monotonicity (M) and Mean Independence (MI), Manski and Tamer (2002) derive the identification region for $E(y\,|\,\upsilon, x)$ and discuss estimation methods.

We estimate the model using the inferential approach of Chernozhukov, Lee, and Rosen (2013) and the Stata package by Chernozhukov et al. (2015). We again present results for the pooled data, which yield a sample of size 13,717 after dropping respondents with missing covariates. As in the application of Section 4.1, when we take the elicited longevity expectations at face value, we drop respondents who answered "Don't know" or refused to answer the probability chance question posed.

In the interest of space, we present results graphically in Figure 1 rather than in a table. Each panel of Figure 1 reports on the x-axis the subjective percent-chance that a respondent will survive to age 75. The y-axis reports the mean weekly hours of work predicted in two ways. One uses a linear regression model estimated by least squares, taking the longevity expectations data at face value ("OLS"). The other ("Bounds") uses interval expectations to account for rounding, where the intervals are those described in Section 3. Additionally, the graphs display 95% confidence intervals for both the OLS and Bounds estimates. Different panels show estimates for different sub-samples, corresponding to different age-coupledness status combinations.

Taking the longevity expectations data at face value, we find that they have a positive but economically insignificant association with hours worked, while hours worked decrease substantially with age and if the respondent is not coupled. When we allow for rounding, as illustrated in the plots in Figure 1, we confirm that predicted mean hours worked increase quite weakly in the perceived likelihood of living past 75, while they decrease markedly as age increases, and for individuals who are not part of a couple.

5. **Conclusion**

We have studied the nature of rounding in numerical reports of probabilistic expectations, a type of survey measure that has become widely used in empirical economic analysis of individual and household decisions under uncertainty. Our analysis of the responses to all expectations questions asked in the HRS core questionnaire between 2002 and 2014 confirms earlier findings based on analysis of the 2006 waves of data and establishes important new findings. We propose an inferential approach that interprets expectations reports as interval data and that explicitly incorporates the documented patterns of responses across waves, question domains, and location within the measurement scale.

The main tenet of the analysis is that observed response patterns across questions and waves carry information about individual respondents' rounding practices. Observed response patterns, however, do not reveal whether individual respondents round their reports to simplify communication or to convey partial knowledge. Consistent with the first interpretation, we have assumed that respondents have well-formed latent point beliefs. If instead the relevant latent objects were sets or ranges of beliefs, the algorithm would still work as intended as long as the algorithm's interval completely includes the latent interval.

If respondents round to convey partial knowledge about the likelihood of future events of the kind HRS expectations questions refer to, it would be better to allow them to express their ambiguity directly. This could be achieved by allowing respondents to give either a single percent-chance value or a range as they see fit. Then range measures of subjective expectations may be analyzed using existing econometric tools for interval data. See Manski and Molinari (2010) and Giustinelli and Pavoni (2017) for exploratory data collection and analysis of this type.

References

Amemiya, T. (1981). Qualitative Response Models: A Survey. *Journal of Economic Literature* 19(4), 1483-1536.

Armantier, O., W. Bruine de Bruin, S. Potter, G. Topa, W. van der Klaauw, and B. Zafar (2013). Measuring Inflation Expectations. *Annual Review of Economics* 5, 273-301.

Attanasio, O. (2009). Expectations and Perceptions in Developing Countries: Their Measurement and Their Use. *American Economic Review* 99(2), 87-92.

Beresteanu, A. and F. Molinari (2008). Asymptotic Properties for a Class of Partially Identified Models. *Econometrica* 76(4), 763-814.

Beresteanu, A. and Molchanov, I. and Molinari, F. (2012). Partial Identification Using Random Set Theory. *Journal of Econometrics* 166(1): 17-32.

Beresteanu, A. and Molinari, F. and Steeg Morris, D. (2010), "Asymptotics for Partially Identified Models in STATA," Computer Program available at https://molinari.economics.cornell.edu/programs.html

Bruine de Bruin, W., P.S. Fischbeck, N.A. Stiber, and B. Fischhoff (2002). What Number is "Fifty-Fifty?": Redistributing Excessive 50% Responses in Elicited Probabilities. *Risk Analysis* 22(4), 713-723.

Bureau of Labor Statistics, U.S. Department of Labor (2018), *Occupational Employment Statistics*, www.bls.gov/oes/ [accessed 1/28/2018].

Chernozhukov, V. and Kim, W. and Lee, S. and Rosen, A.M. (2015). Implementing Intersection Bounds in Stata. *Stata Journal* 15(1):21-44.

Chernozhukov, V. and Lee, S. and Rosen, A.M. (2013). Intersection Bounds: Estimation and Inference. *Econometrica* 81(2): 667-737.

Crimmins, E.M., J.K. Kim, K.M. Langa, and D.R. Weir (2011). Assessment of Cognition Using Surveys and Neuropsychological Assessment: The Health and Retirement Study and the Aging, Demographics, and Memory Study. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* 66B(S1), i162–i171.

Delavande, A. (2014). Probabilistic Expectations in Developing Countries. *Annual Review of Economics* 6, 1-20.

Delavande, A. and Perry M. and Willis, R.J. (2006). Probabilistic Thinking and Social Security Claiming. MRRC Working Paper, 129.

Delavande, A. and C.F. Manski (2010). Probabilistic Polling and Voting in The 2008 Presidential Election: Evidence From The American Life Panel. *Public Opinion Quarterly* 74(3), 433-259.

Dominitz, J. and C.F. Manski (1997). Perception of Economic Insecurity: Evidence From The Survey of Economic Expectations. *Public Opinion Quarterly* 61(2), 261-287.

Fischhoff, B. and W. Bruine de Bruin (1999). Fifty–Fifty = 50%? *Journal of Behavioral Decision Making* 12, 149-163.

Fischhoff, B. and A.M. Parker and W. Bruine de Bruin and J. Downs and C. Palmgren and R. Dawes and C.F. Manski (2000). Teens Expectations For Significant Life Events. *Public Opinion Quarterly* 64(2), 189-205.

Fisher, G.G., H. Hassan, W.L. Rodgers, and D.R. Weir (2012). Health and Retirement Study – Imputation of Cognitive Functioning Measures: 1992–2010 Early Release. http://hrsonline.isr.umich.edu/modules/meta/xyear/cogimp/desc/COGIMPdd.pdf.

Gill, R.D., M.J. van der Laan, and J.M. Robins (1997). Coarsening at Random: Characterization, Conjectures, Counter-Examples. In Proceedings of the First Seattle Symposium in Biostatistics:

Survival Analysis. Eds. D.Y. Lin and T.R. Fleming. Lecture Notes in Statistics. New York: Springer-Verlag, pp. 255-294.

Giustinelli, P. and C.F. Manski (2018). Survey Measures of Family Decision Processes for Econometric Analysis of Schooling Decisions. *Economic Inquiry* 56, 81-99.

Giustinelli, P. and N. Pavoni (2017). The Evolution of Awareness and Belief Ambiguity in the Process of High School Track Choice. *Review of Economic Dynamics* 25, 93-120.

Gutsche, T.L. and A. Kapteyn and E. Meijer and B. Weerman (2014). The RAND Continuous 2012 Presidential Election Poll. *Public Opinion Quarterly* 78(S1), 233-254.

Hamermesh, D. (1985). Expectations, Life Expectancy, and Economic Behavior. *Quarterly Journal of Economics* 100: 389-408.

Heitjan, D.F. and D.B. Rubin (1991). Ignorability and Coarse Data. *Annals of Statistics* 19(4), 2244-2253.

Heitjan, D.F. (1994). Ignorability in General Incomplete-Data Models. *Biometrika* 81(4), 701-708.

Honig, M. (1996). Retirement Expectations by Race, Ethnicity and Gender. *The Gerontologist* 36.

Honig, M. (1998). Married Women's Retirement Expectations: Do Pensions and Social Security Matter? *American Economic Review Papers and Proceedings* 88, 202-206.

Hudomiet, P. and R.J. Willis (2013). Estimating Second Order Probability Beliefs from Subjective Survival Data. *Decision Analysis* 10(2), 152-170.

Hurd, M.D. (2009). Subjective Probabilities in Household Surveys. *Annual Review of Economics* 1, 543-562.

Hurd, M.D. and J.P. Smith and J.M. Zissimopoulos (2004). The Effects of Subjective Survival on Retirement and Social Security Claiming. *Journal of Applied Econometrics* 19: 761-775.

Juster, F.T. (1966). Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design. *Journal of the American Statistical Association* 61, 658-696.

Juster, F.T. and R. Suzman (1995). An overview of the Health and Retirement Study. *Journal of Human Resources* 30, S7–56.

Kleinjans, K.J. and A. van Soest (2014). Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions. *Journal of Applied Econometrics* 29, 567-585.

Lillard, L. and R.J. Willis (2001). Cognition and Wealth: The Importance of Probabilistic Thinking. Michigan Retirement Research Working Paper MRRC WP UM00-04, University of Michigan.

Manski, C.F. (1990). The Use of Intentions Data to Predict Behavior: A Best-Case Analysis. *Journal of the American Statistical Association* 85(412), 934-940.

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

Manski, C.F. (2018). Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise. *NBER Macroeconomics Annual*, 32, 411-471.

Manski, C.F. and F. Molinari (2010). Rounding Probabilistic Expectations in Surveys. *Journal of Business and Economic Statistics* 28(2), 219-231.

Manski, C.F. and E. Tamer (2002). Inference on Regressions with Interval Data on a Regressor or Outcome. *Econometrica* 70(2), 519-546.

Sajaia, Z. (2008). Maximum Likelihood Estimation of a Bivariate Ordered Probit Model: Implementation and Monte Carlo Simulations. *Stata Journal* 4, 1-18.

Schotter, A. and I. Trevino (2014). Belief Elicitation in the Laboratory. *Annual Review of Economics* 6, 103-128.

van der Klaauw, W. (2012). On the Use of Expectations Data in Estimating Structural Dynamic Models. *Journal of Labor Economics* 30(3), 521-554.

Tables and Figures Appendix

Table 1: Probabilistic Expectations Questions in the HRS (Section P, Waves 2002-2014)

#	Question	Wave						
		2002	**2004**	**2006**	**2008**	**2010**	**2012**	**2014**
	PERSONAL HEALTH (3-9 Qs in each wave, 10 across waves)							
P19	Health limit work during next 10 years	Y	-	-	-	-	-	-
P28	**Live to be 75 or more**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P29	**Live to be age X or more**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P32	**Move to nursing home ever (if age<65) / in the next 5 years (if age >= 65)**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P103	Live independently at 75	-	-	Y	Y	-	-	-
P104	Free of serious mental problems at 75	-	-	Y	Y	-	-	-
P106	Live independently at X	-	-	Y	Y	-	-	-
P107	Free of serious problems in thinking/reasoning at X	-	-	Y	Y	-	Y	Y
P108	Same health in 4 years	-	-	Y	Y	-	-	-
P109	Worse health in 4 years	-	-	Y	Y	-	-	-
	PERSONAL FINANCES (11-23 Qs in each wave, 31 across waves)							
P4	Income keep up inflation for next 5 years	Y	Y	Y	-	-	-	-
P5	**Leave inheritance >=$10,000**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P6	**Leave inheritance >=$100,000**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P7	**Leave any inheritance**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P8	Receive inheritance during next 10 years	Y	Y	Y	-	-	-	-
P14	Lose job next year	Y	Y	Y	-	Y	Y	Y
P15	Finding a job in few month in case of job-loss	Y	Y	Y	-	Y	Y	Y
P16	**Working for pay in the future**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P17	**Working full time after age 62**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P18	**Working full time after age 65**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P20	**Finding a job in few months if unemployed**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P30	Give $5,000 to others over next 10 years	Y	Y	Y	-	-	-	-
P31	Receive $5,000 from others over next 10 years	Y	Y	Y	-	-	-	-
P59	**Leave inheritance >=$500,000**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P70	Medical expenses use up savings in next 5 years	-	Y	Y	Y	-	-	-
P71	Give $1,000 to others during next 10 years	-	Y	Y	-	-	-	-
P72	Give $10,000 to others during next 10 years	-	Y	Y	-	-	-	-
P73	Give $20,000 to others during next 10 years	-	Y	Y	-	-	-	-
P74	Receive $2,500 from others over next 10 years	-	Y	Y	-	-	-	-
P75	Receive $1,000 from others over next 10 years	-	Y	Y	-	-	-	-
P76	Receive $10,000 from others over next 10 years	-	Y	Y	-	-	-	-
P111	Soc. Sec. will be worse over next 10 years - current own benefits	-	-	Y	Y	Y	Y	Y
P112	Soc. Sec. will be worse over next 10 years - future own benefits	-	-	Y	Y	Y	Y	Y
P166	Home worth more by next year	-	-	-	-	Y	Y	Y
P168	Home worth more/less by random "X" by next year	-	-	-	-	Y	Y	Y
P175	Out-of-pocket medical expense >$1,500 during next year	-	-	-	-	Y	Y	Y
P176	Out-of-pocket medical expense >$500 during next year	-	-	-	-	Y	Y	Y
P177	Out-of-pocket medical expense >$3,000 during next year	-	-	-	-	Y	Y	Y
P178	Out-of-pocket medical expense >$8,000 during next year	-	-	-	-	Y	Y	Y
P181	Any work after age 70	-	-	-	-	-	Y	Y
P182	Working full time after age 70	-	-	-	-	-	Y	Y
	GENERAL ECONOMIC CONDITIONS (2-7 Qs in each wave, 12 across waves)							
P34	U.S. have economic depression during next 10 years	Y	Y	Y	Y	-	-	-
P47	**Mutual funds increase in value by next year**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**
P110	Social Security in general will become worse in next 10 years	Y	-	Y	Y	Y	Y	-
P114	Mutual funds increase more than the cost of living over next 10 years	-	-	Y	-	-	-	-
P115	Mutual funds increase 8% more than the cost of living over next 10 years	-	-	Y	-	-	-	-
P116	Cost of living increases more than 5% over next 10 years	-	-	Y	Y	-	-	-
P150	Mutual funds increase by 20% (10%, or a random X%) by next year	Y	-	-	Y	Y	Y	Y
P180	Mutual funds decrease by 20% by next year	-	-	-	-	Y	Y	Y
P183	Medicare less generous in next 10 years	-	-	-	-	-	Y	Y
P190	Stock Market increase in value in 12 months of today	-	-	-	-	-	-	Y
P192	Stock Market increse by 20% (in 12 months)	-	-	-	-	-	-	Y
P193	Stock Market decrease by 20% (in 12 months)	-	-	-	-	-	-	Y
	Total N of Questions	**22**	**26**	**38**	**25**	**25**	**29**	**31**

Table 2: Response Tendencies in the 2002-2014 HRS

Wave	N	Response pattern						
		All NR	All 0 or 100	All 0, 50, or 100	Some $\mathbb{M}10$	Some $\mathbb{M}5$	Some 1-4 or 96-99	Some other
Based on the 12 questions asked in all waves								
2002	16032	0.022	0.101	0.101	0.392	0.320	0.054	0.011
2004	18250	0.015	0.062	0.084	0.418	0.353	0.056	0.013
2006	17191	0.027	0.072	0.077	0.409	0.336	0.065	0.014
2008	16060	0.021	0.068	0.063	0.417	0.340	0.072	0.018
2010	20400	0.010	0.053	0.050	0.426	0.350	0.092	0.020
2012	19360	0.015	0.051	0.058	0.445	0.328	0.083	0.020
2014	17647	0.012	0.065	0.062	0.458	0.295	0.090	0.018
Based on all questions asked in each wave								
2002	16032	0.014	0.023	0.039	0.324	0.459	0.119	0.022
2004	18250	0.010	0.019	0.032	0.337	0.467	0.108	0.026
2006	17191	0.025	0.019	0.023	0.263	0.513	0.117	0.039
2008	16060	0.021	0.025	0.019	0.290	0.511	0.101	0.033
2010	20400	0.009	0.029	0.022	0.316	0.442	0.144	0.038
2012	19360	0.014	0.027	0.021	0.317	0.443	0.139	0.038
2014	17647	0.012	0.026	0.022	0.329	0.427	0.142	0.042

NOTE: N = sample size, NR = nonresponse, $10 \equiv \{10, 20, 30, 40, 60, 70, 80, 90\}$, $\mathbb{M}5 \equiv \{5, 15, 25, 35, 45, 55, 65, 75, 85, 95\}$. The following 12 questions were asked in all HRS waves between 2002 and 2014: P47: mutual fund increase in value; P28: live to be 75 or more; P29: live to be X or more; P5: leave inheritance ≥ $10,000; P6: leave inheritance ≥ $100,000; P59: leave inheritance ≥ $500,000; P7: leave any inheritance; P16: work for pay in the future; P17: work full time after age 62; P18: work full time after age 65; P32: move to nursing home in 5 years; P20: finding a job in few months if unemployed.

Table 3: Transitions of Response Tendencies across Waves

Transition waves:	2002 to 2004	2004 to 2006	2006 to 2008	2008 to 2010	2010 to 2012	2012 to 2014	2002 to 2014
	Frequency (based on the 12 questions asked in all waves)						
% transitions to:							
same category	0.406	0.420	0.406	0.415	0.436	0.433	0.389
adjacent category	0.386	0.383	0.383	0.385	0.377	0.373	0.392
more distant category	0.209	0.197	0.212	0.201	0.187	0.194	0.218
N (100%)	14183	16126	15231	13732	18260	16923	8348
same or adjacent	0.792	0.803	0.788	0.800	0.813	0.806	0.782

NOTE: The percentages shown in the table are calculated from transition matrices of response tendencies defined in terms of the following categories: All NR, All (0, 100), All (0, 50, 100), Some $\mathbb{M}10$, Some $\mathbb{M}5$, Some 1-4 or 96-99, Some other. The following 12 questions were asked in all HRS waves between 2002 and 2014: P47: mutual fund increase in value; P28: live to be 75 or more; P29: live to be X or more; P5: leave inheritance ≥ $10,000; P6: leave inheritance ≥ $100,000; P59: leave inheritance ≥ $500,000; P7: leave any inheritance; P16: work for pay in the future; P17: work full time after age 62; P18: work full time after age 65; P32: move to nursing home in 5 years; P20: finding a job in few months if unemployed.

Table 4: Responses by Question and across Waves in the 2002-2014 HRS

Question: percent chance that…	N total obs.	Percentage of responses in: NR	$\mathbb{M}50$	$\mathbb{M}100$	$\mathbb{M}25$	$\mathbb{M}10$ T	$\mathbb{M}10$ C	$\mathbb{M}5$ T	$\mathbb{M}5$ C	$\mathbb{M}1$ T	$\mathbb{M}1$ C
Personal Health											
P19: Health limit work next 10 years	5475	0.044	0.311	0.153	0.087	0.217	0.144	0.031	0.007	0.005	0.001
P28: Live to be age 75 or more	56497	0.038	0.219	0.204	0.082	0.270	0.120	0.042	0.010	0.013	0.001
P29: Live to be age X or more	118404	0.050	0.211	0.191	0.075	0.236	0.156	0.049	0.013	0.018	0.001
P32: Move to nursing home in 5 y	74696	0.059	0.120	0.426	0.039	0.206	0.062	0.060	0.003	0.023	0.001
P103: Live independently at 75	7590	0.031	0.190	0.136	0.115	0.292	0.152	0.056	0.016	0.012	0.001
P104: Free of serious mental… at 75	7590	0.034	0.210	0.099	0.130	0.259	0.183	0.052	0.020	0.011	0.002
P106: Live independently at X	15291	0.060	0.219	0.144	0.100	0.234	0.166	0.046	0.015	0.015	0.001
P107: Free of serious think/reason…	33518	0.062	0.227	0.135	0.088	0.229	0.179	0.049	0.014	0.016	0.001
P108: Same health in 4 years	16253	0.048	0.226	0.151	0.097	0.263	0.151	0.044	0.009	0.010	0.001
P109: Worse health in 4 years	16232	0.069	0.228	0.146	0.077	0.272	0.143	0.043	0.008	0.014	0.001
General Economic Conditions											
P34: U.S. have economic depression	50661	0.069	0.234	0.148	0.083	0.228	0.170	0.041	0.014	0.011	0.001
P47: Mutual funds up /next y	105714	0.157	0.247	0.093	0.076	0.185	0.193	0.025	0.014	0.008	0.001
P110: SS in general will be worse	71770	0.054	0.212	0.200	0.087	0.235	0.151	0.035	0.011	0.014	0.001
P114: Mutual fund up /more than living	16680	0.281	0.182	0.096	0.063	0.178	0.157	0.026	0.010	0.006	0.001
P115: Mutual fund up 8% /more than…	16652	0.307	0.162	0.076	0.061	0.187	0.150	0.033	0.010	0.012	0.001
P116: Cost living up /more than 5%	32431	0.077	0.151	0.210	0.089	0.252	0.152	0.045	0.010	0.013	0.001
P150: Mutual funds up by 20/10/ X%	42092	0.034	0.156	0.090	0.070	0.314	0.237	0.063	0.017	0.018	0.002
P180: Mutual funds down by 20%	31658	0.019	0.179	0.098	0.061	0.318	0.225	0.064	0.017	0.016	0.002
P183: Medicare less generous in 10 y	36524	0.039	0.219	0.216	0.075	0.246	0.150	0.032	0.008	0.014	0.001
P190: Stock market up by next year	8615	0.077	0.335	0.090	0.058	0.185	0.202	0.026	0.011	0.016	0.001
P192: Stock market up by 20%	5430	0.021	0.151	0.108	0.054	0.342	0.199	0.084	0.012	0.028	0.001
P193: Stock market down by 20%	5306	0.013	0.183	0.115	0.048	0.314	0.210	0.076	0.012	0.026	0.002
Personal Finances											
P4: Income keep up inflation in 5 y	51559	0.066	0.196	0.226	0.069	0.249	0.136	0.036	0.007	0.015	0.001
P5: Leave inheritance ≥ $10K	116769	0.046	0.083	0.518	0.028	0.228	0.051	0.028	0.001	0.017	0.000
P6: Leave inheritance ≥ $100K	95625	0.014	0.100	0.490	0.037	0.228	0.072	0.035	0.002	0.022	0.000
P7: Leave any inheritance	19716	0.020	0.053	0.763	0.013	0.098	0.021	0.020	0.001	0.012	0.000
P8: Receive inheritance in 10 y	51559	0.032	0.043	0.755	0.016	0.091	0.024	0.023	0.001	0.014	0.000
P14: Lose job next year	32743	0.017	0.129	0.405	0.028	0.261	0.060	0.067	0.003	0.031	0.000
P15: Find job in few months/loss	32727	0.015	0.158	0.276	0.056	0.287	0.128	0.053	0.004	0.022	0.000
P16: Work for pay in the future	66855	0.018	0.055	0.672	0.021	0.139	0.037	0.035	0.001	0.021	0.000
P17: Work full time after age 62	36603	0.011	0.144	0.333	0.055	0.268	0.120	0.043	0.006	0.020	0.001
P18: Work full time after age 65	37062	0.011	0.144	0.280	0.058	0.282	0.130	0.057	0.008	0.028	0.001
P20: Find job in few months/unemployed	8206	0.012	0.211	0.184	0.061	0.277	0.174	0.050	0.012	0.019	0.001
P30: Give $5K to others in 10 y	50528	0.024	0.120	0.505	0.050	0.187	0.065	0.035	0.002	0.011	0.000
P31: Receive $5K… in 10 y	50528	0.023	0.047	0.674	0.020	0.143	0.026	0.047	0.001	0.019	0.000
P59: Leave inheritance ≥ $500K	73872	0.011	0.090	0.490	0.034	0.216	0.073	0.046	0.003	0.037	0.000
P70: Med expenses use up savings	50478	0.060	0.141	0.316	0.060	0.246	0.109	0.048	0.006	0.014	0.000
P71: Give $1K to others in 10 y	21024	0.007	0.097	0.551	0.044	0.186	0.060	0.041	0.002	0.013	0.000
P72: Give $10K to others in 10 y	12904	0.011	0.212	0.322	0.072	0.219	0.124	0.026	0.006	0.007	0.001
P73: Give $20K to others in 10 y	11155	0.011	0.152	0.334	0.061	0.265	0.100	0.057	0.005	0.015	0.000
P74: Receive $2.5K… in 10 y	30644	0.004	0.021	0.723	0.019	0.134	0.023	0.053	0.001	0.022	0.000
P75: Receive $1K… in 10 y	30397	0.003	0.042	0.686	0.024	0.141	0.031	0.051	0.001	0.021	0.000
P76: Receive $10K… in 10 y	3270	0.015	0.243	0.321	0.052	0.198	0.134	0.022	0.009	0.006	0.001
P111: SS worse/current own benefits	51023	0.036	0.246	0.197	0.080	0.246	0.138	0.037	0.007	0.012	0.001
P112: SS worse/future own benefits	26753	0.020	0.205	0.186	0.085	0.255	0.179	0.040	0.014	0.014	0.001
P166: Home worth more next year	28067	0.030	0.202	0.165	0.045	0.361	0.146	0.033	0.005	0.011	0.001
P168: Home worth more/less by X	26394	0.035	0.112	0.259	0.029	0.348	0.120	0.070	0.004	0.024	0.000
P175: OP med exp ≥ $1.5K next year	56760	0.031	0.143	0.340	0.051	0.261	0.109	0.043	0.004	0.017	0.000
P176: OP med exp ≥ $500 next year	10962	0.017	0.114	0.642	0.025	0.126	0.043	0.020	0.001	0.012	0.000
P177: OP med exp ≥ $3K next year	44022	0.012	0.132	0.235	0.058	0.318	0.126	0.082	0.006	0.033	0.000
P178: OP med exp ≥ $8K next year	36369	0.009	0.079	0.260	0.037	0.327	0.092	0.120	0.005	0.071	0.000
P181: Any work after age 70	17057	0.010	0.118	0.374	0.042	0.259	0.101	0.058	0.005	0.034	0.000
P182: Work full time after age 70	10384	0.003	0.100	0.264	0.038	0.323	0.108	0.097	0.007	0.060	0.000

NOTE: $\mathbb{M}50 \equiv \{50\}$, $\mathbb{M}100 \equiv \{0, 100\}$, $\mathbb{M}25 \equiv \{25, 75\}$, $\mathbb{M}10\text{-T} \equiv \{10, 20, 80, 90\}$, $\mathbb{M}10\text{-C} \equiv \{30, 40, 60, 70\}$, $\mathbb{M}5\text{-T} \equiv \{5, 15, 85, 95\}$, $\mathbb{M}5\text{-C} \equiv \{35, 45, 55, 65\}$, $\mathbb{M}1$-T ≡ non-round values in 1-24 or 76-99, $\mathbb{M}1$-C ≡ non-round values in 26-74.

Table 5: Distribution of Rounding Types by Domain

Rounding Type	Percent Personal Health	Percent Personal Finances	Percent General Economic Conditions
(M1-T, M1-C)	0.17	0.33	0.26
(M1-T, M5-C)	**1.07**	**3.03**	**1.22**
(M1-T, M10-C)	**6.08**	**15.84**	**5.73**
(M1-T, M25)	**1.33**	**1.72**	**0.80**
(M1-T, M50)	**1.27**	**1.31**	**0.86**
(M1-T, None/Undet.)	1.02	0.50	0.42
(M5-T, M1-C)	*0.07*	*0.08*	*0.11*
(M5-T, M5-C)	2.60	2.97	3.65
(M5-T, M10-C)	**16.05**	**23.47**	**16.98**
(M5-T, M25)	**3.20**	**2.95**	**2.29**
(M5-T, M50)	**2.53**	**1.75**	**1.35**
(M5-T, None/Undet.)	1.39	0.53	0.55
(M10-T, M1-C)	*0.13*	*0*	*0.16*
(M10-T, M5-C)	*1.84*	*0.73*	*2.47*
(M10-T, M10-C)	25.92	22.75	32.50
(M10-T, M25)	**5.91**	**5.09**	**5.24**
(M10-T, M50)	**7.98**	**5.88**	**5.93**
(M10-T, None/Undet.)	4.35	2.36	2.70
(M100, M1-C)	*0*	*0*	*0.01*
(M100, M5-C)	*0.16*	*0.03*	*0.14*
(M100, M10-C)	*2.89*	*1.04*	*1.96*
(M100, M25)	*1.62*	*1.01*	*1.08*
(*M100*, M50)	3.90	2.45	2.32
(*M100*, None/Undet.)	4.74	3.42	2.47
(None/Undet., M1-C)	0.01	0	0.01
(None/Undet., M5-C)	0.20	0.01	0.24
(None/Undet., M10-C)	1.27	0.01	2.50
(None/Undet., M25)	0.47	0.00	0.92
(None/Undet., M50)	0.92	0	2.06
(None/Undet., None/Undet.)	0.91	0.75	3.06
Total	100	100	100
Sample size	28044	28252	28172
Tails finer than center	**45.42**	**61.03**	**40.40**
Tails same as center	32.60	28.49	38.73
Tails coarser than center	*6.71*	*2.90*	*5.94*
No/Undet. T and/or C	15.27	7.58	14.93

NOTE: For each domain (T=tail and C= center), M1 denotes a respondent who rounds to the nearest 1 percent in that domain, M5 denotes a respondent who rounds to the nearest 5 percent or finer in that domain, and so on. *U*ndetermined denotes respondents who could not be classified to belong to any of the preceding types.

Table 6: Bivariate Ordered Probit Model Predicting Rounding Type

	Personal Health		**Personal Finances**		**Gen. Econ. Conditions**	
	Tail Type	Center Type	Tail Type	Center Type	Tail Type	Center Type
Male	0.0047	-0.0497	-0.0032	-0.0154	-0.0070	-0.0693
	(0.0149)	(0.0155)	(0.0142)	(0.0153)	(0.0151)	(0.0157)
Age 60-69	-0.1961	-0.1436	-0.0116	0.0145	-0.1090	-0.1049
	(0.0180)	(0.0194)	(0.0174)	(0.0189)	(0.0185)	(0.0195)
Age 70-79	-0.1639	0.0481	0.1466	0.1987	-0.0941	0.0232
	(0.0199)	(0.0206)	(0.0189)	(0.0204)	(0.0199)	(0.0208)
Age 80+	0.1092	0.4465	0.4934	0.5658	0.1718	0.3209
	(0.0266)	(0.0261)	(0.0246)	(0.0258)	(0.0266)	(0.0266)
High school	-0.0842	-0.0864	-0.1277	-0.1579	-0.0614	-0.1115
	(0.0224)	(0.0221)	(0.0208)	(0.0219)	(0.0226)	(0.0227)
Some college	-0.0642	-0.0758	-0.1688	-0.1948	-0.0588	-0.1487
	(0.0362)	(0.0379)	(0.0342)	(0.0372)	(0.0364)	(0.0389)
Bachelor	-0.2027	-0.2432	-0.2677	-0.3073	-0.1726	-0.2692
	(0.0288)	(0.0301)	(0.0277)	(0.0296)	(0.0292)	(0.0305)
Graduate	-0.2818	-0.3658	-0.3367	-0.3549	-0.2438	-0.3454
	(0.0319)	(0.0337)	(0.0307)	(0.0332)	(0.0320)	(0.0341)
Black	0.0188	0.1148	-0.1507	-0.0798	-0.0562	-0.0456
	(0.0220)	(0.0226)	(0.0203)	(0.0220)	(0.0219)	(0.0228)
Other race	0.1136	0.1374	0.0604	0.0173	0.0887	0.0477
	(0.0303)	(0.0322)	(0.0289)	(0.0310)	(0.0314)	(0.0322)
Avg. Cog.	-0.0261	-0.0339	-0.0368	-0.0373	-0.0202	-0.0370
	(0.0022)	(0.0023)	(0.0020)	(0.0022)	(0.0022)	(0.0023)
Rho	0.2595		0.3848		0.2897	
	(0.0081)		(0.0087)		(0.0093)	
N	22,447		24,541		22,593	

NOTES: (i) Respondents with undetermined tail or center rounding type are excluded from this analysis. (ii) Predictors are dummies for gender, age (averaged across waves), education, and race, plus average cognition score across waves. (iii) Omitted dummies are ‘Female,’ ‘Age in 50-59,’ ‘No degree,’ and ‘White.’ (iv) ‘Rho’ is the parameter capturing the correlation between the error terms of the tail and center latent equations. (v) Standard errors are in parentheses.

Table 7. Predicted Probabilities of Rounding Types for Selected Covariate Profiles

Panel A. Personal Health –(Female, White, Bachelor Degree) Respondents

Average Cognition Across Waves

Average Age Across Waves	Mean -1 SD	Mean	Mean +1 SD		Mean -1 SD	Mean	Mean +1 SD
	Prob. of Type ($\mathcal{M}$5-T, $\mathcal{M}$10-C)				**Prob. of Type ($\mathcal{M}$10-T, $\mathcal{M}$10-C)**		
50-59	0.1846	0.2036	0.2198	50-59	0.3118	0.3123	0.3064
60-69	0.2136	0.2289	0.2402	60-69	0.2971	0.2897	0.2767
70-79	0.2008	0.2194	0.2347	70-79	0.2784	0.2768	0.2696
80+	0.1433	0.1658	0.1878	80+	0.2494	0.2623	0.2701
	Prob. of Type ($\mathcal{M}$100-T, M25-C)				**Prob. of Type ($\mathcal{M}$100-T, $\mathcal{M}$50-C)**		
50-59	0.0199	0.0157	0.0121	50-59	0.0312	0.0221	0.0153
60-69	0.0135	0.0103	0.0077	60-69	0.0192	0.0133	0.0090
70-79	0.0151	0.0119	0.0091	70-79	0.0256	0.0180	0.0124
80+	0.0247	0.0207	0.0170	80+	0.0583	0.0433	0.0316

Panel B. Personal Finances –(Female, White, Bachelor Degree) Respondents

Average Cognition Across Waves

Average Age Across Waves	Mean -1 SD	Mean	Mean +1 SD		Mean -1 SD	Mean	Mean +1 SD
	Prob. of Type ($\mathcal{M}$5-T, M10-C)				**Prob. of Type ($\mathcal{M}$10-T, $\mathcal{M}$10-C)**		
50-59	0.2634	0.2724	0.2731	50-59	0.2483	0.2248	0.1976
60-69	0.2632	0.2722	0.2728	60-69	0.2440	0.2209	0.1942
70-79	0.2453	0.2621	0.2715	70-79	0.2583	0.2415	0.2191
80+	0.1887	0.2162	0.2402	80+	0.2665	0.2665	0.2586
	Prob. of Type ($\mathcal{M}$100-T, $\mathcal{M}$25-C)				**Prob. of Type ($\mathcal{M}$100-T, $\mathcal{M}$50-C)**		
50-59	0.0072	0.0049	0.0032	50-59	0.0107	0.0065	0.0038
60-69	0.0071	0.0048	0.0031	60-69	0.0107	0.0065	0.0038
70-79	0.0102	0.0071	0.0048	70-79	0.0175	0.0110	0.0067
80+	0.0196	0.0149	0.0109	80+	0.0443	0.0298	0.0194

Panel C. General Economic Conditions –(Female, White, Bachelor Degree) Respondents

Average Cognition Across Waves

Average Age Across Waves	Mean -1 SD	Mean	Mean +1 SD		Mean -1 SD	Mean	Mean +1 SD
	Prob. of Type ($\mathcal{M}$5-T, $\mathcal{M}$10-C)				**Prob. of Type ($\mathcal{M}$10-T, $\mathcal{M}$10-C)**		
50-59	0.2031	0.2170	0.2273	50-59	0.3733	0.3724	0.3647
60-69	0.2201	0.2315	0.2387	60-69	0.3625	0.3562	0.3435
70-79	0.2157	0.2298	0.2401	70-79	0.3509	0.3495	0.3415
80+	0.1671	0.1858	0.2027	80+	0.3524	0.3658	0.3725
	Prob. of Type ($\mathcal{M}$100-T, $\mathcal{M}$25-C)				**Prob. of Type ($\mathcal{M}$100-T, $\mathcal{M}$50-C)**		
50-59	0.0111	0.0088	0.0068	50-59	0.0165	0.0116	0.0080
60-69	0.0086	0.0067	0.0051	60-69	0.0119	0.0082	0.0056
70-79	0.0094	0.0074	0.0051	70-79	0.0145	0.0101	0.0070
80+	0.0166	0.0138	0.0112	80+	0.0317	0.0233	0.0167

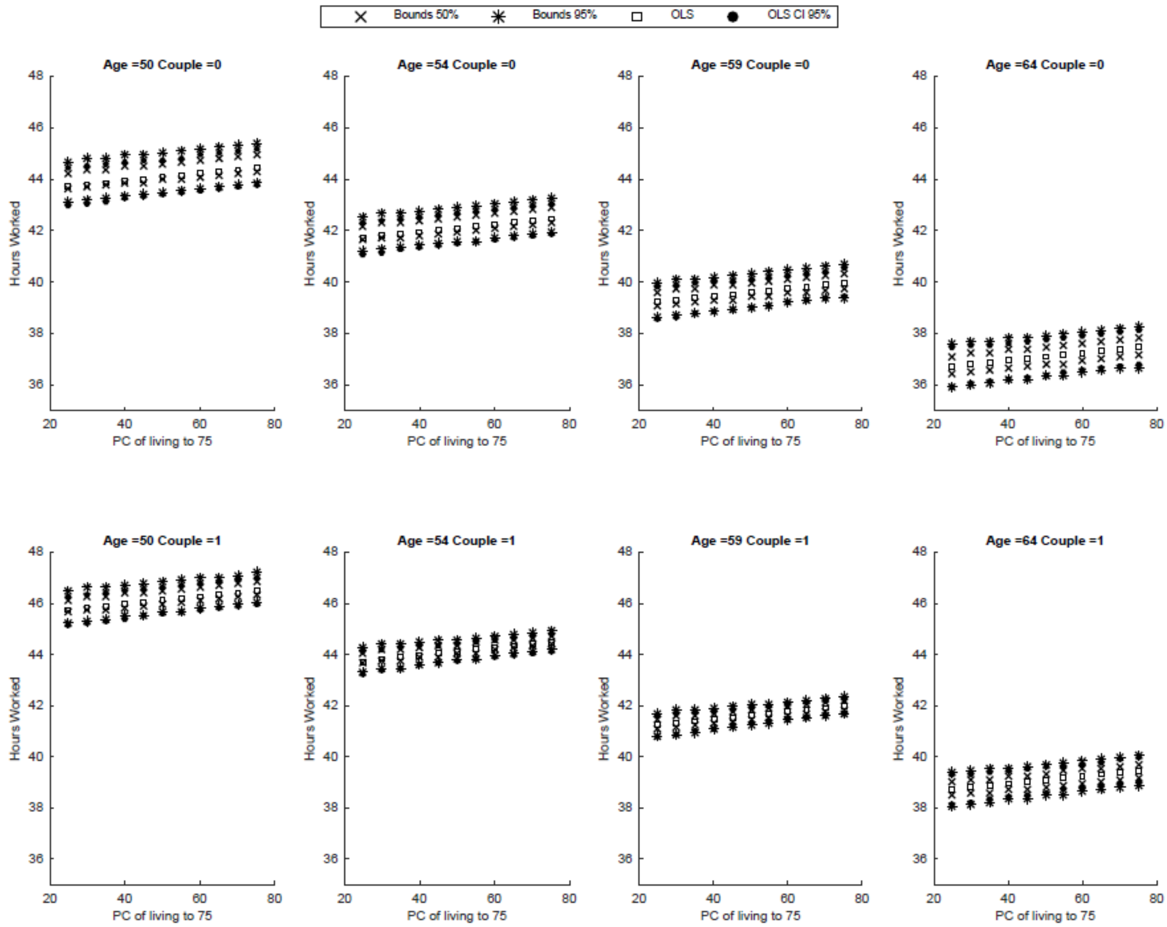
NOTES: (i) ($\mathcal{M}$ 5-T, $\mathcal{M}$10-C) denotes rounding to the nearest 5 percent or a finer degree in the tails and rounding to the nearest 10 percent or a finer degree in the center. ($\mathcal{M}$ 10-T, $\mathcal{M}$10-C) denotes rounding to the nearest 10 percent or a finer degree in both the tails and the center. ($\mathcal{M}$ 100-T, $\mathcal{M}$25-C) denotes rounding to any degree in the tails and to the nearest 25 percent or a finer degree in the center. ($\mathcal{M}$ 100-T, $\mathcal{M}$50-C) denotes rounding to any degree in both the tails and the center. (ii) Predicted probabilities are evaluated at the mean value of average cognition across waves (denoted Mean), at the mean minus one standard deviation value of average cognition across waves (denoted Mean – 1 SD), and as the mean plus one standard deviation Value of average cognition across waves. Predicted probabilities are evaluated at average age across waves falling in each of the categories 50-59, 60-69, 70-79, and 80+.

Table 8: BLP Prediction of Retirement Expectations: Point Estimates vs. Set Estimates with Pooled HRS 2002-2014 Data

	OLS Estimates I	**Set Estimates I**		**Set Estimates II**	
	(MCAR imposed)	**LB**	**UB**	**LB**	**UB**
Age	0.1638	-0.4036	0.7212	-0.4944	0.8110
	(0.0306, 0.2970)	(-0.5177, 0.8353)		(-0.5944, 0.9110)	
Coupled	-2.694	-8.5773	3.2792	-9.6014	4.4009
	(-4.1348, -1.2533)	(-9.6555, 4.3573)		(-10.6521, 5.4517)	
Male	8.2172	2.1835	13.9580	1.1365	14.7641
	(7.0017, 9.4327)	(1.2710, 14.8705)		(0.2024, 15.6982)	
Negative wealth	6.1812	-1.6447	13.5758	-4.1145	15.4530
	(4.3986, 7.9637)	(-3.2409, 15.1720)		(-5.7203, 17.0588)	
Below median wealth	6.2116	-1.5954	13.5862	-3.9164	15.2990
	(4.4898, 7.9333)	(-2.9980, 14.9888)		(-5.4242, 16.8065)	
Above median wealth	-0.4701	-9.3489	8.1918	-11.5634	9.8589
	(-2.5209, 1.5808)	(-10.9746, 9.8176)		(-13.1321, 11.4276)	
Black	-9.8655	-16.0655	-3.3521	-17.2459	-2.2001
	(-11.5115, 8.2196)	(-17.2151, -2.20253)		(-18.3527, -1.0933)	
Other race	-4.8209	-11.5792	2.1776	-13.2752	4.2223
	(-6.8371, -2.8046)	(-12.9955, 3.5940)		(-14.7696, 5.7167003	
High school	10.5356	3.0627	17.337	0.2633	19.1237
	(8.7016, 12.3696)	(1.5481, 18.8521)		(-1.1983, 20.5853)	
Some college	13.4775	4.7073	21.5118	1.9292	23.2121
	(10.7289, 16.2260)	(2.7421, 23.4770)		(0.0495, 25.0918)	
Bachelor degree	17.0926	7.9728	25.3006	5.2205	27.0224
	(14.6899, 19.4953)	(6.0970, 27.1764)		(3.5435, 28.6994)	
Graduate degree	19.1551	9.7651	27.6084	7.0036	29.3428
	(16.3555, 21.9546)	(7.8350, 29.5384)		(5.0635, 31.2829)	
Constant	26.0763	-5.8898	59.6645	-10.5647	65.9696
	(18.3266, 33.8259)	(-12.6411, 66.4158)		(-16.5846, 71.9895)	
N	23,811	23,811		24,052	

NOTE: OLS and SetBLP estimates I calculated after dropping DK/RF responses to the point PC question. SetBLP estimates II include DK/RF responses to the point PC question. 95% confidence intervals in parenthesis. OLS CIs clustered at the HH level. SetBLP estimates calculated using 501 bootstrap repetitions. Beresteanu and Molinari (2008)'s confidence sets based on directed Hausdorff. Omitted dummies are '0 wealth,' 'white,' and 'no degree.'

Figure 1: BLP Prediction of Hours Worked Per Week: Point Estimates vs. Set Estimates with Pooled HRS 2002-2014 Data



NOTE: OLS and SetBLP estimates of hours worked per week as a function of longevity expectations, age, and coupledness status. SetBLP estimates are obtained using Chernozhukov et al. (2013, 2015)'s inferential approach. Each graph plots the estimates as a function of longevity expectations for different age groups-coupledness status combinations.