Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Using Expectations Data to Study Subjective Income Expectations

Jeff Dominitz & Charles F. Manski

To cite this article: Jeff Dominitz & Charles F. Manski (1997) Using Expectations Data to Study Subjective Income Expectations, Journal of the American Statistical Association, 92:439, 855-867, DOI: 10.1080/01621459.1997.10474041

To link to this article: https://doi.org/10.1080/01621459.1997.10474041

Using Expectations Data To Study Subjective Income Expectations

Jeff DOMINITZ and Charles F. MANSKI

We have collected data on the one-year-ahead income expectations of members of American households in our survey of economic expectations (SEE), a module of a national continuous telephone survey conducted at the University of Wisconsin. The income-expectations questions take this form: "What do you think is the percent chance (or what are the chances out of 100) that your total household income, before taxes, will be less than Y over the next 12 months?" We use the responses to a sequence of such questions posed for different income thresholds Y to estimate each respondent's subjective probability distribution for next year's household income. We use the estimates to study the cross-sectional variation in income expectations one year into the future. We find that the estimated subjective interquartile range (IQR) of future income tends to rise with the estimated subjective median, but more slowly than proportionately. There is substantial variation in the estimated subjective IQR among respondents with the same estimated subjective median; thus respondents appear to have fairly heterogeneous perceptions of the one-year-ahead income uncertainty they face. Much of the cross-sectional variation in the central tendency of income expectations is associated with realized income, and some of the cross-sectional variation in income uncertainty is associated with realized income, age, and employment status.

KEY WORDS: Economic expectations; Elicitation; Subjective probability; Survey research.

1. INTRODUCTION

Economic theorizing about household behavior assigns a central role to income expectations as a determinant of consumption and savings decisions (Friedman 1957), schooling, labor supply, and occupation decisions (Becker 1964), and other major decisions. Yet economists engaged in empirical research on household behavior lack empirical knowledge of income expectations. Instead, the prevalent practice is to infer expectations from data on income realizations and assumptions about the process of expectations formation (see, e.g., Caballero 1990; Carroll 1992; Freeman 1971; Hall and Mishkin 1982; Skinner 1988; Willis and Rosen 1979; Zeldes 1989).

This article reports empirical evidence on the one-year-ahead household income expectations of respondents to our Survey of Economic Expectations (SEE), a module of a national continuous telephone survey conducted at the University of Wisconsin. The expectations questions take this form:

"What do you think is the percent chance (or what are the chances out of 100) that your total household income, before taxes, will be less than Y over the next 12 months?"

The responses to a sequence of such questions posed for different income thresholds Y enable estimation of each respondent's subjective probability distribution for next year's household income. We use the estimates to study the empirical distribution of income expectations across the sample of respondents.

The feasibility and utility of eliciting expectations in household surveys is a subject of some controversy among behavioral and social scientists. In particular, there are disparate perspectives on the relative merits of eliciting probabilistic and qualitative expectations. Section 2 explains our thinking in undertaking this study.

There are not many precedents for probabilistic elicitation in household surveys, so our work is necessarily exploratory. Section 3 describes our elicitation procedures at some length. We examine the response rate to the expectations questions and to questions asking respondents for their realized household income in the past 12 months. We explain how we use the expectations data to fit respondent-specific subjective distributions for next year's income. We assess the coherence of the responses and the meaningfulness of the fitted subjective distributions.

Section 4 analyzes the subjective uncertainty respondents report concerning next year's income. The findings enable us to assess the assumptions about income uncertainty imposed in economic studies using realizations to infer expectations. Lacking expectations data, economists have been able to only speculate about the uncertainty that persons perceive concerning their future incomes. It has been common to assume that the spread and central tendency of income expectations are proportional to one another. We find that the estimated subjective interquartile range (IQR) of future income tends to rise with the estimated subjective median, but more slowly than proportionately. Moreover, we find substantial variation in the estimated subjective IQR among respondents with the same estimated subjective median. Thus respondents appear to have fairly heterogeneous perceptions of the one-year-ahead income uncertainty they face.

Section 5 examines how income expectations vary with respondents' realized incomes and other attributes. We find that much of the cross-sectional variation in the central ten-

Jeff Dominitz is Instructor, Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125. Charles F. Manski is Professor, Department of Economics, University of Wisconsin, Madison, WI 53706. This research was supported by National Science Foundation grants SBR-9223220 and SBR-9321044. The authors are grateful to Bob Lee and Jim Sweet for enabling collection of the data by the University of Wisconsin Survey Center at the University of Wisconsin-Madison. They are grateful for the opportunity to present this work in seminars at the University of Wisconsin, the University of Michigan, Tilburg University, the Federal Reserve Board, and the University of Chicago, and in the Colloquium on Expectations Formation at the University of Paris. They have benefitted from the comments of the reviewers and an associate editor, and also from discussions with Baruch Fischhoff.

dency of income expectations is associated with realized income, and some of the cross-sectional variation in income uncertainty is associated with realized income, age, and employment status. We show how the SEE data may be used to estimate best empirical predictors of expectations conditional on realizations data of the type available in major household surveys. We suggest that such best predictors of expectations may improve on the practice of inferring expectations from realizations data alone. Section 6 gives conclusions.

2. BACKGROUND

The nature of subjective expectations is a shared concern of economists, psychologists, Bayesian statisticians, decision researchers, and many others. Yet there is a wide spectrum of views on how expectations should be conceptualized and studied. We decided to launch the SEE after reflecting on three matters: (1) the credibility problem confronting any attempt to infer expectations from realizations; (2) the limited comparability and informativeness of responses to qualitative expectations questions; and (3) the state of empirical knowledge about probabilistic elicitation of income and other economic expectations. This section discusses each in turn.

2.1 Inference on Expectations From Realizations

To credibly infer expectations from realizations, researchers must somehow know (and be able to persuade others that they know) what information on realizations persons possess and how they use the available information to form their expectations. Moreover, the available data on realizations must be rich enough for the researcher to simulate the assumed processes of expectations formation. Economists seeking to infer expectations from realizations are generally unable to meet these requirements for credible inference. They can only make assumptions about expectations formation and hope that these assumptions approximate reality.

The studies to date universally assume that expectations formation is homogenous; all persons condition their beliefs on the same variables and process their information in the same way. The hypothesized conditioning variables and information processing rule vary considerably across studies, however.

Consider, for example, the work of Caballero (1990), Carroll (1992), Hall and Mishkin (1982), Skinner (1988), and Zeldes (1989), all of which was concerned with unconditional income expectations. These authors assumed that persons use their own past incomes to forecast their future incomes. Perhaps so, but how do persons form expectations of future income conditional on past income? The authors assume that persons have *rational expectations;* that is, they know the actual stochastic process generating their income streams. Perhaps so, but what is this stochastic process? Each study a priori specifies the income process up to some parameters and uses available data on income realizations to estimate the parameters. The various authors differ in their specifications of the income process, however.

Or consider the work of Freeman (1971) and Willis and Rosen (1979), which was concerned with the income returns to schooling. These authors did not assume that persons use their own past incomes to forecast their future incomes. Rather, their models of expectations invoked cohort comparisons. Freeman (1971) assumed that expectations are myopic; each person believes that, should he obtain specified schooling, he will experience the mean income realized by the members of an earlier cohort who have this schooling. Willis and Rosen (1979) assumed rational expectations, each person somehow knowing the schooling-specific income distributions pertaining to his own cohort.

Economists have lacked empirical evidence that any of the expectations assumptions made in the studies to date is correct. Nor have they had reason to think that misspecifying expectations is innocuous. To the contrary, misspecification of expectations can easily generate flawed interpretations of observed behavior (see Manski 1993).

2.2 Qualitative Expectations Questions

If inference on income expectations from realizations is so challenging, then why do economists not collect and analyze data on expectations? The answer lies at least partly in the history of a scientific controversy that began in the 1940s and persisted until the 1960s. In the 1940s, the Federal Reserve Board began to fund annual surveys of consumer finances, conducted by the University of Michigan Survey Research Center (SRC), that elicited qualitative assessments of expected household finances. A typical question took this form:

> "How about a year from now—do you think you people will be making more money or less money than you are now, or what do you expect?"

The usefulness of responses to such vaguely worded questions was controversial, and the Board of Governors appointed a committee to assess their value. The Federal Reserve Consultant Committee on Consumer Survey Statistics (1955) called into question the value of the SRC data in predicting individual behavior. Katona (1957) defended SRC practices, asserting that the data were useful in predicting aggregate consumer behavior even if they were not useful in predicting individual behavior. A contentious conference on expectations data took place at the National Bureau of Economic Research (1960) and was followed by an intensive empirical study (Juster 1964). Juster found that data of the type collected at SRC have limited predictive value in regressions of behavior on realizations and expectations. By the mid-1960s, opinion among mainstream economists was firmly negative. Nonetheless, SRC has continued to collect qualitative expectations data and to publish aggregate findings monthly in its Index of Consumer Sentiment (see Curtin 1976 and Patterson 1991).

The Federal Reserve–SRC controversy may have soured most economists on the usefulness of qualitative expectations data, but collection and analysis of such data have continued to be important activities in other behavioral and social science disciplines (see Schuman and Johnson 1976 and Turner and Martin 1984). These applications notwith-

standing, we find qualitative expectations questions to be problematic on two grounds.

Perhaps the most basic problem concerns the interpersonal and intrapersonal comparability of responses. Respondents are often asked to report whether occurrence of a specified event is "expected," "likely," or "probable." Sometimes they are asked to report their strength of belief in the event on a numerical scale with undefined units. Do different respondents interpret these verbal phrases and numerical scales in the same way? Does a given respondent interpret them in the same way when asked about different events? As far as we are aware, cognitive and linguistic research does not give reason to think that responses should be or are comparable. Indeed, the available empirical evidence indicates that interpretations of qualitative expectations questions vary substantially between persons (see Wallsten, Budescu, Rapaport, Zwick, and Forsyth 1986).

A second problem is that the coarseness of the response options to common qualitative questions limits the information contained in the responses. Consider, for example, the fertility question asked of female respondents in the June 1987 supplement to the Current Population Survey:

> Looking ahead, do you expect to have any (more) children?
> Yes No Uncertain
> (U.S. Bureau of the Census 1988)

The three response options can express little of the richness of the uncertainty that women may perceive about their future childbearing. Savage (1971, p. 795) called attention to the problem when he observed that "Yes, no, or maybe is not enough." Juster (1966) hypothesized that a person facing a yes/no expectations question responds as would a statistician asked to make a best point prediction of a future random event. Manski (1990) formalized this hypothesis and derived an upper bound on the information about expectations that a researcher can extract from responses to yes/no questions.

2.3 Probabilistic Expectations Questions

Elicitation of probabilistic expectations has several a priori desirable features. One is that a researcher can use the algebra of probability (e.g., Bayes theorem, the law of total probability, etc.) to examine the internal consistency of a respondent's elicited expectations about different events. A second attraction—at least in those cases where probability has a frequentist interpretation—is that a researcher can compare elicited subjective probabilities with known event frequencies and reach conclusions about the correspondence between subjective beliefs and frequentist realities.

From the perspective of economic research, a third attraction is that probabilistic expectations provide empirical evidence in the form sought by modern economic theory. Economists analyzing decision making under uncertainty generally assume that individuals assign coherent subjective probabilities to future events. So economists can readily utilize probabilistic expectations data, whereas qualitative expectations data are difficult to utilize.

Contemplating initiation of the SEE, we needed to weigh the desirable a priori features of probabilistic expectations data against a set of potential problems. Perhaps the most basic concern is that there is no direct way to assess how well elicited income expectations reflect respondents' thinking. One may try to indirectly judge the meaningfulness of elicited income expectations by examining the internal consistency of the responses to questions about different events, the sensitivity of responses to nonsubstantive changes in question format, and the constancy of the responses in retest exercises. One cannot, however, externally validate income expectations in the same way that one can externally validate respondents' reports of their realized incomes. Of course, this concern is not specific to probabilistic elicitation of income expectations. Interpretation of responses to subjective questions of any type runs up against the generic problem that a researcher cannot directly observe a respondent's thinking.

A specific concern in the elicitation of probabilities is that respondents may not think probabilistically about uncertain events. A large and diverse literature, ranging from Keynes (1921) through Ellsberg (1961) to Walley (1991), among many others, suggests that persons may think about uncertain events using less than the full structure of modern probability theory. In particular, they may think in terms of upper and lower probabilities (see Walley 1991 for an extensive review). If so, then questions eliciting precise probabilities force respondents to give point responses when only interval responses are meaningful. A quite different position has been taken by Zimmer (1983, 1984), who hypothesized that humans internally process information using verbal rather than numerical modes of thinking. If so, then individuals may find it difficult to express their expectations numerically.

In personal communications, some survey researchers and cognitive psychologists have expressed the concern that whether or not persons can think probabilistically, they may be unwilling to articulate their thoughts when probabilistic expectations questions are posed in surveys. We are not aware of empirical research that directly supports this concern. A pair of recent studies report that a majority of their respondents prefer to communicate their own beliefs verbally and to receive the beliefs of others in the form of numerical probabilities (Erev and Cohen 1990; Wallsten, Budescu, Zwick, and Kemp 1993). This asymmetry is intriguing but only marginally relevant to the design of SEE. The relevant question for us was not what communication mode respondents prefer to use, but rather what modes they are willing to use.

The available empirical evidence regarding willingness to respond to probabilistic questions is relatively favorable. There is a long history of small, intensive studies eliciting probabilistic information from experts on specific technical subjects (see Hogarth 1975; Morgan and Henrion 1990; Wallsten and Budescu 1983). More germane are the findings on elicitation of personal beliefs from nonexperts.

Wallsten et al. (1993) reported that virtually all of their respondents were willing to communicate their beliefs in either probabilistic or verbal terms if the situation should warrant it. Willingness to respond has been borne out in those surveys that have elicited probabilistic expectations. Thirty

years ago, Juster (1966) reported positive findings eliciting purchase probabilities for durable goods from consumers. Purchase probabilities have since been elicited in a number of market research studies (see Jamieson and Bass 1989). Quadrel, Fischhoff, and Davis (1993) elicited from a broad sample of adults and adolescents their subjective probabilities of experiencing various illnesses and other health risks. Hurd and McGarry (1995) analyzed probabilistic expectations of mortality elicited in the National Health and Retirement Study.

We are aware of only two previous studies eliciting some form of probabilistic income expectations. Guiso, Jappelli, and Terlizzese (1992) analyzed expectations of one-year-ahead growth rates in labor earnings elicited in the 1989 edition of the Bank of Italy's biennial survey of the Italian population. And we have carried out an exploratory study eliciting high school and college students' expectations of the long-term income returns to schooling (Dominitz and Manski 1996).

3. ELICITING SUBJECTIVE DISTRIBUTIONS OF FUTURE INCOME

3.1 The Survey of Economic Expectations

From February–May 1993, we placed a set of probabilistic expectations questions concerning future income, earnings, and employment as a module in WISCON, an ongoing national computer-assisted telephone interview (CATI) survey conducted by the University of Wisconsin Survey Center at the University of Wisconsin-Madison. The WISCON core questions ask respondents about their labor market experiences, demographics, realized household income, and qualitative expectations (see Winsborough 1987). We refer to the edition of WISCON containing our expectations module as the survey of economic expectations SEE (see Sec. 1). This article analyzes the SEE income-expectations data. The data on earnings and employment expectations were examined in Dominitz (1994).

The SEE module follows the WISCON core questions. The core questions on realized income ask respondents to include income for all sources for each adult member of the household; see Appendix A.1 for the form of these questions. The SEE questions eliciting income expectations simply refer to "household income," but we think it reasonable to assume that respondents will have in mind the detailed definition of income given earlier in the survey.

In principle, income expectations might be elicited by asking each respondent to report quantiles of his or her subjective distribution of future income, moments of the distribution, or points on the cumulative distribution function (cdf). Morgan and Henrion (1990) discussed at length the practical pros and cons of different procedures for eliciting subjective distributions. Their recommendations formed the basis for our approach, with some tailoring of the procedures to fit the survey medium (telephone interview) and subject matter.

In a telephone survey it is infeasible to present visual aids that may help respondents to understand questions and to think probabilistically. Use of the telephone medium led us, after some pretests, to reject elicitation of quantiles of the subjective income distribution in favor of eliciting points on the cdf. We have elicited the medians of subjective income distributions in our study of the return-to-schooling expectations of high school and college students (Dominitz and Manski 1996). That study used an interactive computer program to elicit expectations, however. Questions are posed on-screen, and respondents key-in answers.

As indicated in Section 1, we asked each respondent to provide a sequence of points on his or her subjective cdf of household income over the next 12 months. Respondent i was asked about four income thresholds $Y_{i1}, \ldots, Y_{i4}$, posed in increasing order. The interviewer informed the respondent if a probability elicited at threshold Y_{i2}, Y_{i3}, or Y_{i4} was smaller than one elicited earlier. This ensured that the sequence of responses was always logically coherent. The only exception to the protocol occurs if a response of "100% chance" is given when the first, second, or third threshold is posed. In such cases, it is not necessary to elicit further responses as a coherent subjective distribution must give "100% chance" to all subsequent thresholds. See Appendix A.2 for the precise wording of the questions.

The thresholds about which a given respondent was queried were determined by the respondent's answer to a pair of preliminary questions asking for the lowest and highest possible incomes that the household might experience in the next year (see App. A.2). We did not interpret the answers to these preliminary questions literally as minimum and maximum incomes, because the phrases "lowest possible" and "highest possible" are too vague to warrant this formal interpretation (see the discussion in Sec. 3.3). Instead, we used the responses to suggest the general region of the respondent's subjective support of future income. Our reasoning was that responses to questions about a range of thresholds spanning the support of a respondent's subjective distribution should yield more information about the shape of the distribution than would the same number of questions asked about a narrower or wider range of thresholds.

Morgan and Henrion (1990) offered two other reasons for asking these preliminary questions. One is to decrease "overconfidence" problems wherein respondents focus too heavily on central tendencies, downweighting their uncertainty about outcomes. Another is to decrease "anchoring" problems wherein respondents' beliefs are influenced by the questions that interviewers happen to pose. Suppose, for example, that a respondent expects his household income to be no less than $30,000. Psychologists fear that if the first question asked concerns the probability that household income will be less than $15,000, then the respondent may be influenced to think that this amount is objectively reasonable and so may report a higher probability than believed a priori.

Our computer-programmed question-branching algorithm selected income thresholds Y in a region determined by the responses to the preliminary questions. In particular, the midpoint between the elicited lowest possible and highest possible incomes was used to determine the sequence of four thresholds about which the respondent would be queried. Respondents with midpoints in the interval [20,

25) thousand dollars, for example, were asked the percent chance that next year's income would be less than {15, 20, 25, 30} thousand dollars. Appendix A.3 gives the threshold sequences for all respondents.

Sections 3.2 and 3.3 describe our empirical experience with the question-branching algorithm and with the elicitation process more generally.

3.2 Response Rates

Using a random-digit dialing algorithm, the WISCON interviewers attempt to contact an adult respondent at about 20 households per day and obtain an interview at slightly more than 50% of the selected households. The nonresponse is fairly evenly divided between refusals to be interviewed and cases in which repeated phone calls do not find an appropriate respondent at home.

Our expectations module was administered to 622 WISCON respondents. Of these 622 respondents, 509 answered the preliminary questions eliciting their lowest and highest possible incomes in the next year. Of these 509 respondents, 489 answered the subsequent questions eliciting points on the subjective cdf. We are able to fit a subjective income distribution only to the responses of respondents whose elicited probabilities vary with the income threshold posed (see Sec. 3.4). Of the 489 respondents who completed the SEE, 437 gave responses in which the elicited probabilities do so vary. Hence the effective response rate for our analysis is 437/622, or about .70.

The 52 cases out of 489 in which respondents reported the same probability values at all four of the income thresholds posed have this breakdown: 22 respondents answered "100%" to the first threshold posed (implying the same answer for all subsequent thresholds), 15 answered "0%" four times, and 15 reported a single value between 0 and 100 four times. The constant-probability responses given by these 52 respondents may reflect an inability to articulate their expectations probabilistically or may legitimately express their thinking. The 22 respondents answering "100%" to the first threshold, for example, may have subjective income distributions concentrated to the left of this threshold. Refinement of our question-branching algorithm to permit dynamic determination of thresholds may yield question sequences that succeed better in spanning the supports of these respondents' subjective distributions.

The .70 effective response rate to the income-expectations module compares favorably with the response rate to the WISCON core questions asking respondents to report their household income in the past 12 months. A total of 390 of the 622 respondents answered the questions asking for their own income and, if applicable, the incomes of other adult members of the household. Thus the WISCON response rate for realized household income is about .63.

Response to the expected income and realized income questions was positively associated, as follows:

- Effective response to both expected and realized income: 331 respondents (.53)
- Effective response to expected income only: 106 respondents (.17)
- Effective response to realized income only: 59 respondents (.09)
- Effective response to neither expected nor realized income: 126 respondents (.20)

Response rates to both the expected and the realized income questions vary substantially with respondent attributes. Table 1 reports a number of patterns. Males responded to each set of questions more often than females (.78 versus .65 and .69 versus .58). Response rates first rise with age and then fall: Sample members age 40–49 responded most frequently (.82 and .77); those age 60 and over responded least frequently (.52 and .52). Response rates increase with education: sample members with college degrees responded much more frequently (.77 and .72) than those with less than a high school diploma (.44 and .49).

These response rate statistics corroborate the common finding that income questions draw substantial nonresponse on surveys. Yet the statistics also carry an important positive message—the SEE questions eliciting probabilistic income expectations draw less nonresponse than the realized-income questions administered in the same survey. Some of the nonresponse that we experience derives not from respondent failure to answer the expectations questions, but rather from the constancy of the answers across the thresholds posed. Refinement of the protocol for elicitation of ex-

Table 1. Response Rates for Respondents With Various Attributes

	Number of respondents	Response rate: Expected income	Response rate: Realized income
Total sample	622	.70	.63
Demographics			
Female	358	.65	.58
Male	264	.78	.69
White	535	.67	.62
Nonwhite	79	.63	.68
Single	237	.65	.69
Married	353	.74	.60
Cohabiting	30	.77	.47
Age < 30	104	.78	.62
30–39	145	.79	.68
40–49	137	.82	.77
50–59	74	.68	.59
≥60	151	.52	.52
Education			
<12 years	61	.44	.49
High school diploma	124	.60	.48
Some postsecondary	254	.77	.67
≥Bachelor's degree	181	.77	.72
Employment status			
Unemployed	21	.62	.67
Employed	401	.79	.67
Out of labor force	176	.50	.52
Temporary absence	10	.80	.70

NOTE: A few sample members did not respond to some questions eliciting their demographic, education, and employment attributes. The numbers of respondents reporting different values of a given attribute sum to correspondingly less than 622.
Source: Survey of Economic Expectations

pectations may yield improved selection of thresholds and, as a result, higher effective response rates than those found in this study.

3.3 Are the Expectations Responses Coherent?

Willingness to respond to the SEE questions does not imply that the responses are meaningful. Before using the responses to fit respondent-specific subjective income distributions, we attempt to judge whether respondents report their expectations coherently. We draw on three types of evidence: the rate of logical response errors, the extent of bunching of responses, and the association between the elicited probabilities and the responses to the preliminary questions asking for lowest possible and highest possible future incomes.

3.3.1 Violations of Monotonicity. Consistency checks in the CATI software ensure that the probabilities elicited from each respondent must ultimately be logically consistent, in the sense that the response increases weakly across the four thresholds posed. Because the CATI software captures as output the entire sequence of a respondent's responses and revisions, we are able to determine how often respondents must be prompted to change initially nonmonotone response patterns.

We find that 415 of the 437 respondents with usable responses to the percent chance questions initially report a monotone response sequence. Only 22 respondents initially report a sequence of values that violates monotonicity of the cdf and so need to be prompted for revision. The low rate of violation of monotonicity in the income-expectations section of SEE may be due in part to the training respondents receive answering earlier sections of the questionnaire, which elicit expectations about individual earnings and unemployment durations. We see that 51 of the 437 respondents violate monotonicity in one of these earlier sections of SEE but give monotone response sequences in the section eliciting household income expectations.

3.3.2 Bunching of The Percent Chance Responses. It is of interest to examine the extent of bunching of the percent chance responses at 0%, 50%, and 100%. Asked for the percent chance that next year's income will be less than the first (lowest) threshold, 144 respondents say 0% chance, 52 respondents say 50% chance, 22 say 100% chance, and the remaining 271 give some other value. Here are the statistics for all four thresholds:

	Percent chance			
	0	50	100	Other
Threshold 1	144	52	22	271
Threshold 2	78	94	40	277
Threshold 3	24	56	114	295
Threshold 4	15	24	236	214

The pattern of responses across thresholds is sensible. As the threshold increases, the number of respondents reporting 0% chance falls sharply, the number reporting 50% chance rises and then falls, and the number reporting 100% chance rises sharply. Most of the responses are at values other than 0%, 50%, and 100%. Examination of the raw data shows that respondents tend to round their responses to the nearest 5% (e.g., 30%, 35%, 40%), except at the tails where they round to the nearest 1% (e.g., 97%, 98%, 99%).

3.3.3 Reported Lowest and Highest Possible Incomes. As described earlier, we used preliminary questions eliciting the lowest and highest possible incomes that the household might experience in the next year to suggest the region of the respondent's subjective support of future income. It is tempting to interpret the responses to these questions as literally giving the support of the subjective income distribution. But this interpretation is at odds with the responses to the percent chance questions.

Comparison of the responses is most straightforward when the question-branching algorithm yields thresholds $Y_1, \ldots, Y_4$ that include a respondent's reported lowest possible income $Y_{\min}$ or highest possible income $Y_{\max}$. It happens that $Y_{\min} \in (Y_1, \ldots, Y_4)$ for 200 of the 437 respondents and $Y_{\max} \in (Y_1, \ldots, Y_4)$ for 221 respondents. Among the 200 respondents asked for the percent chance that household income will be less than $Y_{\min}$, the median response was 20%. Among the 221 respondents asked for the percent chance that income will be less than $Y_{\max}$, the median response was 80%. Here are more detailed findings:

	Reported percent chance that income will be less than	
Empirical quantiles	$Y_{\min}$	$Y_{\max}$
.10 quantile	0	20
.25 quantile	5	50
.50 quantile	20	80
.75 quantile	50	90
.95 quantile	60	100
Number of respondents	200	221

These findings indicate that most respondents associate the phrases "lowest possible" and "highest possible" with low and high probabilities, but do not interpret these phrases as defining the support of a subjective distribution. See Section 3.5 for further analysis.

3.4 Fitting Subjective Income Distributions

After division by 100, we interpret a respondent's answers to the four percent chance questions as points on his or her subjective cdf of household income over the next 12 months. Thus for each respondent i, we observe $F_{ik} \equiv \Pr(y < Y_{ik} | \psi_i), k = 1, 2, 3, 4$. Here y denotes future income, ψ_i denotes the information currently available to respondent i, and $Y_{i1} < Y_{i2} < Y_{i3} < Y_{i4}$ are the income thresholds about which the respondent is queried.

The subjective probabilities $(F_{ik}, k = 1, \ldots, 4)$ elicited from respondent i imply bounds on his or her subjective income distribution but do not identify the distribution. It is possible, but cumbersome, to analyze the expectations data without imposing auxiliary assumptions (see Dominitz 1994 chaps. 2 and 3). It facilitates analysis if we use the ex-

pectations data to fit a respondent-specific parametric distribution.

Let $F(Y; m, r)$ denote the log-normal cdf with median m and interquartile range r, evaluated at any point Y. For each respondent i, we find (m_i, r_i) that solves the least squares problem

$$\inf_{m,r} \sum_{k=1}^{4} [F_{ik} - F(Y_{ik}; m, r)]^2$$

and analyze the data as if we observe i's subjective income distribution to be $F(Y; m_i, r_i)$. In particular, we use the median m_i to characterize the central tendency of respondent i's subjective income distribution and the interquartile range r_i to characterize its spread. A reader who prefers to use the mean and variance to measure central tendency and spread can derive these quantities from the log-normal median and interquartile range.

Observe that our formulation of the least squares problem seeks the "inf" rather than the "min" of the sum of squares. This is because the least squares solution is a degenerate log-normal distribution whenever at least three of the four elicited probabilities $(F_{ik}, k = 1, \ldots, 4)$ take the value 0 or 1. For example, if the responses are (0, 0, 0.8, 1), then the best fitting distribution has all its mass at the single point Y_{i3}; hence $m_i = Y_{i3}$ and $r_i = 0$. This distribution gives a perfect fit to the four observed responses. Thus our fitting criterion interprets some respondents as having no subjective uncertainty about their future income. Of the 437 respondents, 85 gave responses that imply such degenerate solutions to the least squares problem.

3.5 Are the Fitted Subjective Income Distributions Meaningful?

There is inevitably some arbitrariness in using any specific criterion (here least squares) to fit the expectations data to any specific family of distributions (here the log-normal distributions). The four probabilities elicited from each respondent do not suffice to determine whether log normality, which is a common assumption in studies of realized income distributions, best describes the shape of respondents' subjective income distributions. We can say that the respondent-specific log-normal distributions fit well in the sense of yielding small deviations between elicited and fitted probabilities. The average absolute deviation

$$\frac{1}{4} \sum_{k=1}^{4} |F_{ik} - F(Y_{ik}; m_i, r_i)|$$

between elicited and fitted probabilities is less than .05 for 360 of the 437 respondents and is less than .17 for all of the respondents.

Table 2 tabulates the medians m and IQRs r of the log-normal distributions fitted to the expectations data elicited from the 437 respondents. We have no direct way to assess how well these fitted subjective medians and IQRs reflect respondents' thinking. But we can compare the fitted (m, r) values to other measures of the central tendency and spread of respondents' expectations obtained from the reported lowest and highest possible incomes $(Y_{\min}, Y_{\max})$.

In particular, consider using the midpoint $M_i \equiv (Y_{i,\min} + Y_{i,\max})/2$ to characterize the central tendency of respondent i's expectations and using the range $R_i \equiv Y_{i,\max} - Y_{i,\min}$ to characterize the spread. Then the (M_i, R_i) pair is loosely analogous to the (m_i, r_i) pair. We say loosely because, as described in Section 3.3, the interval $(Y_{i,\min}, Y_{i,\max})$ should not be interpreted as giving the support of a subjective probability distribution.

Table 2 also tabulates (M, R) in the same way as it tabulates (m, r). It is striking how similar the two tables appear. Even more salient is Table 3, which tabulates (m, M) and (r, R). The first part of Table 3, which segments income into \$10,000 intervals up to \$70,000, shows a very strong positive association between m and M; 285 of the 437 entries are on the diagonal, and 408 are within one cell of the diagonal. The second part of Table 3, which segments income into \$5,000 intervals up to \$25,000, shows a less strong but nevertheless clear positive association between r and R, 141 entries are on the diagonal, and 275 are within one cell of the diagonal.

We take Tables 2 and 3 as evidence that the fitted subjective income distributions do meaningfully express the main features of respondents' income expectations. Recall that (m, r) and (M, R) are separate measures of respondents' expectations, the former being fitted functions of the raw percent chance data and the latter being derived from the data on lowest and highest possible incomes. The similarity of the values of these measures suggests that most respondents do have coherent expectations and that they can articulate their expectations in numerical terms.

The analysis in the remainder of this article uses (m, r), which has a formal probabilistic interpretation, to characterize expectations rather than (M, R), which lacks such an interpretation. It may be that the (m, r) and (M, R) data can

Table 2. Frequencies of (m,r) and (M,R)

Frequencies of fitted subjective (median, IQR)

Subjective median (m)	*Subjective interquartile range (r)*						
	[0]	*(0, 5)*	*[5, 10)*	*[10, 15)*	*[15, 25)*	*[25, ∞)*	*Totals*
[0, 20)	23	20	19	9	7	3	81
[20, 40)	28	25	32	30	24	13	152
[40, 60)	13	7	15	17	23	20	95
[60, 80)	11	4	19	13	9	8	64
[80, 100)	2	1	0	2	5	3	13
[100, ∞)	8	0	1	0	5	18	32
Totals	85	57	86	71	73	65	437

Frequencies of subjective (midpoint, range) data

Subjective midpoint (M)	*Subjective range (R)*						
	[0]	*(0, 5)*	*[5, 10)*	*[10, 15)*	*[15, 25)*	*[25, ∞)*	*Totals*
[0, 20)	22	19	16	5	3	1	66
[20, 40)	23	25	45	26	19	9	147
[40, 60)	13	20	19	24	17	16	109
[60, 80)	2	6	9	13	13	14	57
[80, 100)	2	2	3	5	8	13	33
[100, ∞)	5	0	0	2	2	16	25
Totals	67	72	92	75	62	69	437

NOTE: The units of *m*, *r*, *M*, and *R* are thousands of dollars.

Table 3. Frequencies of (m, M) and (r, R)

Frequencies of (m,M)

Subjective median (m)	*Subjective midpoint (M)*							
	[0, 10)	*[10, 20)*	*[20, 30)*	*[30, 40)*	*[40, 50)*	*[50, 60)*	*[60, 70)*	*[70, ∞)*
[0, 10)	3	1	4	0	0	0	1	1
[10, 20)	6	45	18	0	1	1	0	0
[20, 30)	2	7	52	18	4	1	0	0
[30, 40)	1	0	9	37	18	1	1	1
[40, 50)	0	0	2	4	34	12	1	0
[50, 60)	1	0	1	1	7	23	8	1
[60, 70)	0	0	0	0	0	5	15	6
[70, ∞)	0	0	1	0	0	2	4	76

Frequencies of (r,R)

Subjective IQR (r)	*Subjective range (R)*						
	[0]	*(0, 5)*	*[5, 10)*	*[10, 15)*	*[15, 20)*	*[20, 25)*	*[25, ∞)*
[0]	33	20	15	5	3	4	5
(0, 5)	10	17	14	13	3	0	0
[5, 10)	9	17	27	17	2	6	8
[10, 15)	9	8	17	18	9	7	3
[15, 20)	1	2	12	10	8	5	12
[20, 25)	3	0	3	3	3	4	7
[25, ∞)	2	8	4	9	3	5	34

NOTE: The units of *m*, *r*, *M*, and *R* are thousands of dollars.
Source: Survey of Economic Expectations

usefully be combined to yield expectational measures that improve on each separately. This possibility seems worth exploring in future research.

4. SUBJECTIVE INCOME UNCERTAINTY

4.1 Assumptions in Studies Inferring Expectations from Realizations

Studies inferring income expectations from income realizations have typically assumed a fixed relationship between the central tendency and the spread of expectations. Hall and Mishkin (1982) assumed the subjective distribution of next year's income to be normal with household-specific mean μ_i and constant variance σ^2 or, equivalently, with constant IQR $1.349 \cdot \sigma$. Hall and Mishkin did not state what year's price level they used in converting nominal income to real terms, but it appears that they used 1967 dollars. Prices increased by approximately a factor of 3.5 between 1967 and 1993, so we rescale their estimate of σ by this factor and obtain an estimate for σ of \$8,100 in 1993 dollars. Thus Hall and Mishkin estimated the subjective IQR of next year's income to be \$10,900 for all households.

Skinner (1988) and Zeldes (1989) made the subjective distribution of next year's log-income normal with household-specific mean $\log(m_i)$ and constant variance δ^2. Equivalently, the subjective distribution of income is log-normal with median m_i and IQR $m_i[\exp(.6745 \cdot \delta) - \exp(-.6745 \cdot \delta)]$. Thus the IQR in these studies is proportional to the subjective median m_i.

Carroll (1992) assumed the same form for expectations as Skinner and Zeldes, except that he superimposed a .005 chance of receiving no income at all. This slight modification of the log normality assumption has negligible effect on the median and IQR of the subjective income distribution. Using data on family income from 1968–1985, Carroll estimated δ to be .192. Thus he estimated household i's subjective IQR of next year's income to be about $.26 \cdot m_i$.

4.2 Empirical Findings

Table 4 presents selected quantiles of the empirical distribution of the subjective medians m and IQRs r, and kernel-smoothed quantiles of the empirical distribution of r conditional on m. Here and elsewhere in the article, the estimates use the standard normal density as kernel and a bandwidth of \$10,000. Table 4 also presents bootstrapped confidence intervals that reflect our choice of kernel and bandwidth.

Be aware that Table 4 refers to three distinct probability distributions. First, each respondent i has a fitted *subjective income distribution* indexed by (m_i, r_i). Second, there is an *empirical distribution* of (m, r) across the 437 respondents. Third, considering the 437 respondents to be a random sample from a population of potential respondents with usable expectations data, one may view the entries in the table as estimates of quantiles of the *population distribution* of (m, r).

For example, the entry in the first row of the column titled "Subjective median (m)" shows that 10% of the respondents believed there to be at least a 50-50 chance that their household income in the next year would be no greater than \$15,000. The associated confidence interval (13.7, 15.6) is an interval estimate for the unknown value $Y_{.10}$ appearing in this statement: 10% of potential respondents believe there to be at least a 50-50 chance that their income in the next year would be no greater than $Y_{.10}$ thousand dollars.

Table 4 shows that our subjective IQR estimates based on expectations data have the same order of magnitude as the estimates implied by recent studies using income realizations to infer expectations. The empirical median of r

Table 4. Quantiles of the Empirical Distribution of Income Expectations

Empirical quantile	Subjective median (m)	Subjective IQR (r)	Subjective IQR conditional on median		
			m = 20	m = 40	m = 60
.10	15.0	0	0	0	0
	(13.7, 15.6)	(0, 0)	(0, 0)	(0, 0)	(0, 0)
.25	22.8	3.1	1.4	4.2	5.3
	(21.3, 25.0)	(2.3, 4.2)	(0, 2.6)	(3.1, 4.8)	(4.1, 7.2)
.50	37.9	9.6	6.7	9.9	11.7
	(36.2, 40.2)	(8.3, 10.5)	(5.7, 8.5)	(8.5, 11.3)	(10.6, 14.8)
.75	59.7	17.4	13.6	17.0	19.0
	(55.0, 65.7)	(16.4, 18.2)	(11.7, 14.4)	(15.1, 17.7)	(17.7, 23.8)
.90	8.1	31.3	22.6	28.0	32.1
	(77.8, 96.0)	(27.9, 34.9)	(17.4, 23.7)	(23.7, 29.5)	(26.3, 32.6)

NOTE: The top entries are the empirical quantiles of m and r and kernel-smoothed empirical quantiles of r conditional on m. The bottom entries are bootstrapped 90% confidence intervals interpreting the SEE respondents as a random sample from a population of potential respondents. The units of m and r are thousands of dollars.
Source: Survey of Economic Expectations.

presented in the second column of Table 4 is \$9,600. The Hall and Mishkin estimate of \$10,900 is close to this figure, as is the Carroll estimate of $.26 \cdot m_i$ when computed at the empirical median of m presented in the first column of Table 4. Setting $m = 37.9$, the Carroll estimate of IQR is \$9,900.

Although the estimates based on realizations data and expectations data have the same order of magnitude, Table 4 clearly indicates that the IQR of income expectations is neither constant across households nor proportional to the subjective median. Conditioning on m, we find that r varies substantially across respondents. The value of r tends to increase with m, but more slowly than proportionately. The kernel-smoothed empirical median of r increases from \$6,700 to \$9,900 to \$11,700 as m increases from \$20,000 to \$40,000 to \$60,000. Observe that the .10 quantile of r is 0; that is, r equals 0 for at least 10% of the respondents. These are the respondents discussed in Section 3.4, whose fitted log-normal distributions are degenerate.

Table 5. Best Linear LAD Prediction of Expectations Conditional on Realizations

	Subjective median (m)	Subjective IQR (r)
Predictor		
Household income in past 12 months (10^3 dollars)	.896	.172
	(.036)	(.032)
Labor force participation by respondent and spouse (2 if both, 1 if either)	−1.088	−.106
	(1.465)	(1.339)
Unemployment indicator (1 if respondent or spouse is unemployed)	−4.439	10.193
	(2.069)	(6.220)
Gender of respondent (1 if female)	−1.661	−1.083
	(1.013)	(1.244)
Marital status of respondent (1 if married or cohabiting)	−.999	.266
	(1.132)	(1.149)
Age of respondent (years)	−.014	−.138
	(.047)	(.050)
Respondent has postsecondary schooling but no Bachelor's degree (1 if yes)	−.915	−1.105
	(1.172)	(1.712)
Respondent has Bachelor's degree (1 if yes)	−1.077	.517
	(1.263)	(1.621)
Constant	7.333	8.316
	(3.710)	(3.758)
Average absolute deviation between outcome and its empirical median	23.924	10.677
Average absolute deviation between outcome and its BLP	10.184	9.482

NOTE: LAD estimation was performed using the STATA software. The top entries are the coefficient estimates. The bottom entries are bootstrapped standard errors. The units of m and r are thousands of dollars.
Source: Survey of Economic Expectations

Our finding of extensive cross-sectional variation in r conditional on m calls into question the conventional assumption of a fixed relationship between the central tendency and spread of expectations. Hall and Mishkin (1982), for example, identified the responsiveness of consumption to transitory income by assuming that the variance of income is constant across households and over time. In a study using consumption and income realizations data to measure permanent income inequality, Blundell and Preston (1994) assumed that the variance of income is constant within age cohorts in a given year. If these identifying assumptions are substantially incorrect, then it is problematic to interpret the analyses that rest on them.

Studies inferring income expectations from realizations data often use their estimates to calibrate models of precautionary savings behavior in the presence of income uncertainty (see, e.g., Caballero 1990, Carroll 1992, Hubbard, Skinner, and Zeldes 1995, Skinner 1988, and Zeldes 1989.) These studies attempt to explain the substantial cross-sectional variation in savings, conditional on observable attributes, documented by Avery and Kennickell (1989) and others. Hubbard et al. (1995), for example, argued that some cross-sectional variation in savings reflects the incentive effects of asset-based, means-tested social insurance programs on precautionary savings. Other studies attribute cross-sectional variation in savings to heterogeneous preferences or to liquidity constraints. Our empirical findings raise the possibility that the cross-sectional variation in subjective uncertainty may account for at least some of the cross-sectional variation in savings.

5. USING SEE DATA TO PREDICT EXPECTATIONS CONDITIONAL ON REALIZATIONS

Major household surveys have not regularly asked probabilistic questions eliciting income and other economic ex-

Table 6. Descriptive Statistics for the Variables in Table 5

Variable	Mean	Std. Dev.	Min	Max
Subjective median	47.7	32.8	0	180.1
Subjective IQR	13.8	22.0	0	263.8
Realized household income	49.2	33.0	1.7	205.0
Number in labor force	1.27	.67	0	2
Unemployment indicator	.04	.19	0	1
Respondent gender	.51	.50	0	1
Respondent marital status	.62	.49	0	1
Respondent age	43.6	14.6	18	85
Postsecondary schooling	.44	.50	0	1
Bachelor's degree	.35	.48	0	1

pectations. Hence researchers have had to learn about expectations in other ways. Surveys such as the SEE make it possible to improve on the conventional approach of inferring expectations from realizations data alone. In particular, the SEE data may be used to estimate best empirical predictors of expectations conditional on realizations data of the type available in major household surveys.

Table 5 presents least-absolute-deviations estimates of linear functions using the realized household income data and other respondent attributes collected in the WISCON core questionnaire to predict respondents' fitted subjective medians and IQRs. The attribute data include the respondent's gender, marital status, age, education, and employment status. The data also include the employment status of the respondent's spouse/partner, when one exists.

As indicated in Section 3.2, 331 respondents provided usable data on income realizations as well as expectations. Of these 331 respondents, 7 provided incomplete responses about some of the attributes used as predictors in Table 5. Hence the estimates in Table 5 are based on the 324 respondents with complete realizations, expectations, and predictor data. Table 6 presents statistics describing the outcome and predictor variables of these respondents. Sections 5.1 and 5.2 examine the findings.

5.1 Predicting the Medians of Subjective Income Distributions

Table 5 shows striking empirical findings on prediction of respondents' fitted subjective medians m. Consider first the overall fit between m and its best linear predictor (BLP). Whereas the average absolute deviation between m and its empirical median is \$23,900, the average absolute deviation between m and its BLP is just \$10,200. Thus the BLP "explains" more than half the empirical variation in m.

Realized household income, henceforth denoted by y, is the dominant predictor variable. The estimated BLP of m increases \$896 with every \$1,000 increase in y. Other respondent attributes have modest or negligible effects on the BLP. The predicted value of m in a household where the

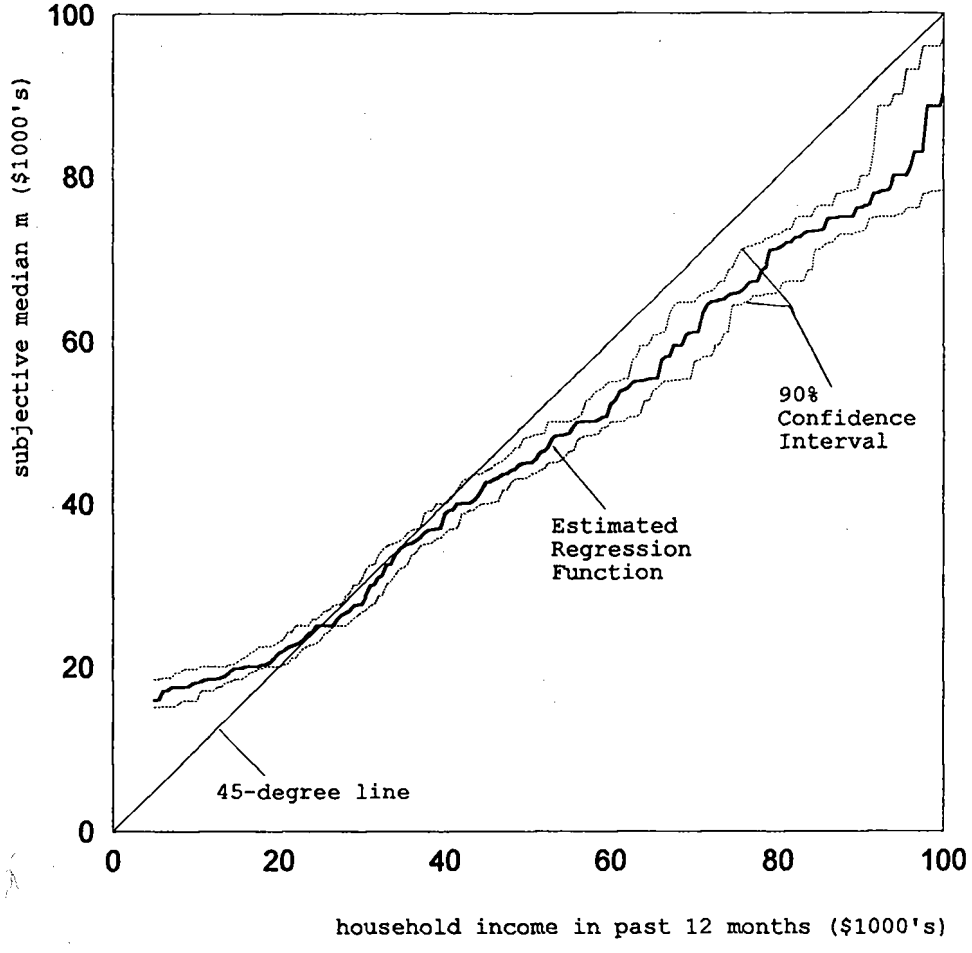

Figure 1. Best Prediction of Expectations Conditional on Realizations med(m|y).

respondent or spouse is unemployed is \$4,400 lower than in a household where neither is unemployed, ceteris paribus. The predicted value of m varies little with the labor force participation of the respondent and spouse, and with the respondent's gender, marital status, age, and education, ceteris paribus.

Given the dominance of realized household income as a linear predictor of m, it is natural to ask how m varies with y if the predictor function is not constrained to be linear. Figure 1 graphs med$(m|y)$, a kernel-smoothed estimate of the population median of m conditional on y. A bootstrapped 90% confidence interval is also shown.

The graph shows med$(m|y)$ to be close to a linear function of y. The confidence interval is quite tight at all but the highest income levels. Observe that med$(m|y) > y$ at low values of y and med$(m|y) < y$ at high values of y. This pattern is consistent with the hypothesis that respondents believe current income to have permanent and transitory components. Under this hypothesis, respondents with low current income would expect next year's income to be higher, and those with high current income would expect next year's income to be lower.

5.2 Predicting the IQRs of Subjective Income Distributions

The findings on prediction of respondents' fitted subjective IQR's r may be less dramatic but still are interesting. The average absolute deviation between r and its empirical median is \$10,700, whereas the average absolute deviation between r and its BLP is \$9,500. Thus the BLP "explains" about 11% of the empirical variation in r.

Table 7. Best Prediction of Expectations Conditional on Realizations med(r|y, age)

Realized income	Age 30	40	50	60
10	6.9	6.7	3.9	1.3
	(5.7, 10.0)	(4.4, 7.3)	(1.4, 5.7)	(0, 2.4)
20	6.9	6.8	5.7	3.6
	(5.7, 8.7)	(5.2, 8.5)	(3.8, 7.1)	(1.3, 8.2)
30	7.8	8.6	6.9	7.8
	(6.0, 10.9)	(6.4, 10.9)	(5.5, 8.4)	(4.0, 8.3)
40	10.0	10.1	7.1	7.2
	(7.2, 12.1)	(6.9, 12.6)	(6.2, 9.6)	(4.0, 9.1)
50	11.1	10.9	7.6	6.2
	(8.2, 12.5)	(7.6, 13.0)	(6.9, 11.1)	(2.5, 8.3)
60	10.8	11.1	9.3	7.2
	(8.6, 13.4)	(8.0, 14.2)	(7.2, 14.2)	(3.1, 9.3)
70	10.6	13.6	9.3	9.3
	(8.8, 14.0)	(8.8, 14.8)	(7.6, 11.1)	(3.1, 14.0)
80	10.6	14.8	11.1	14.0
	(8.6, 16.0)	(8.8, 17.0)	(7.6, 14.2)	(8.6, 32.1)
90	10.6	15.6	17.0	17.7
	(0, 17.0)	(10.0, 17.7)	(9.1, 19.0)	(14.0, 32.1)
100	15.9	17.7	19.0	20.0
	(0, 38.8)	(10.0, 27.4)	(10.1, 27.4)	(8.6, 57.4)

NOTE: The top entries are kernel-smoothed empirical medians of r conditional on (y, age). The bottom entries are bootstrapped 90% confidence intervals interpreting the SEE respondents as a random sample from a population of potential respondents. The units of m and r are thousands of dollars.

Source: Survey of Economic Expectations.

Realized household income is an important predictor variable but is not the only important one. The estimated BLP of r increases \$172 with every \$1,000 increase in y. The BLP decreases \$138 with every year increase in the age of the respondent. The predicted value of r in a household where someone is unemployed is fully \$10,200 higher than in a household where no one is unemployed, ceteris paribus. This last effect is enormous, but it should be kept in mind that only 4% of the SEE respondents report having someone unemployed in the household (see Table 6).

Table 7 focuses more closely on the predictors (y, age). The table reports med$(r|y$, age), a kernel-smoothed estimate of the population median of r conditional on (y, age). Conditional on age, med$(r|y$, age) seems always to be an increasing, or at least nondecreasing, function of realized income y. The behavior of med$(r|y$, age) as a function of age seems to vary with the value of y. The confidence intervals on these nonparametric estimates are too wide for us to draw firm conclusions, however.

6. CONCLUSION

We began this project cognizant that eliciting probabilistic expectations in household surveys is a controversial undertaking. Yet we felt that launching SEE must be worthwhile given the absence of any good alternative way to learn about economic expectations. Severe credibility problems confront attempts to infer expectations from realizations. The information yielded by qualitative expectations questions is too limited to serve the needs of economic analysis. The absence of well-defined scales for responses seems to pose a fundamental problem for qualitative elicitation, manifesting itself in interpersonal incomparability of the elicited responses.

We conclude with a sense that probabilistic elicitation is broadly feasible and useful. The way we have posed questions and fit subjective distributions appears to have worked well enough to yield suggestive findings about basic features of respondents' income expectations, as shown in Sections 4 and 5. At the same time, we have no doubt that our elicitation procedures may be improved. Refining the process of selecting thresholds may yield improved coverage of the supports of respondents' subjective distributions and higher usable response rates. Further study of the association between the responses to the percent chance questions and the questions eliciting lowest and highest possible incomes may yield improved ways of characterizing the central tendency and spread of respondents' expectations. There is much to be done to yield mature elicitation procedures with well-understood properties.

There is also much to be done to understand the process by which persons form their income expectations. The SEE data reveal only what expectations people have, not how they form those expectations. There is a need to join data of the type collected in SEE with data revealing the information sources that people utilize and how they transform the available information into expectations.

We expect that elicited income expectations will eventually prove highly valuable in the study of consumption/savings choices, job search activities, and other eco-

nomic behaviors. The current version of SEE does not question respondents about these behaviors. In the future, we hope to add such questions to the survey.

APPENDIX: QUESTIONS ELICITING REALIZED AND EXPECTED INCOME

A.1 Realized Income

Did you have any income, from any source, in the past 12 months? Be sure to include income from work, government benefits, pensions, and all other sources.

Roughly, what was your *own* total income, from all sources, in the past 12 months, *before taxes?* Be sure to include income from work, government benefits, pensions, and all other sources. (The respondent is then asked, using the same question format, about the incomes of his or her spouse/partner and other adults in the household.)

A.2 Expected Income

Now I would like to ask you some final questions about your household income prospects over the next 12 months.

What do you think is the *lowest* amount that your total household income, from all sources, could possibly be over the next 12 months, *before taxes?*

What do you think is the *highest* amount that your total (household) income, from all sources, could possibly be over the next 12 months, *before taxes?*

Still thinking about your total household income, *before taxes,* over the next 12 months . . .

What do you think is the *percent chance* (or what are the *chances out of 100*) that your total (household) income, *before taxes,* will be less than Y? (This question is posed for each of the income thresholds $Y_{i1}, \ldots, Y_{i4}$.)

A.3 Income Thresholds

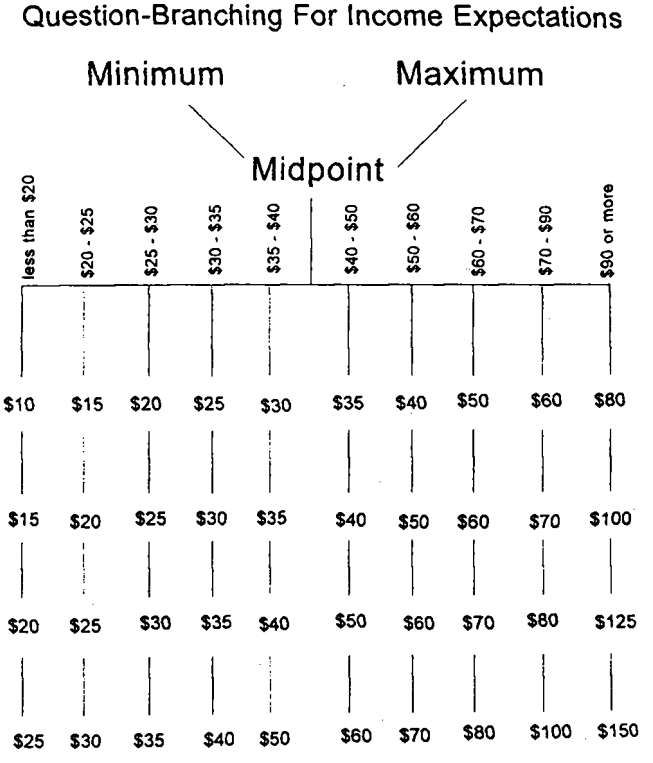

Figure A.1. Algorithm for Selection of Income Thresholds. All amounts correspond to thousands of dollars.

[Received August 1995. Revised January 1997.]

REFERENCES

Avery, R., and Kennickell, A. (1991), "Household Saving in the U.S.," *The Review of Income and Wealth*, 37, 409–432.

Becker, G. (1964), *Human Capital*, New York: Columbia University Press.

Blundell, R., and Preston, I. (1994), "Income or Consumption in the Measurement of Inequality and Poverty," Institute for Fiscal Studies Working Paper No. W94/12, University College London.

Caballero, R. (1990), "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics*, 25, 113–136.

Carroll, C. (1992), "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity*, 2, 61–156.

Curtin, R. (1976), "Survey Methods and Questionnaire," in *Surveys of Consumers: 1974–75*, ed. R. Curtin, Ann Arbor, MI: Institute for Social Research.

Dominitz, J. (1994), "Subjective Expectations of Unemployment, Earnings, and Income," Ph.D. dissertation, University of Wisconsin-Madison.

Dominitz, J., and Manski, C. (1996), "Eliciting Student Expectations of the Returns to Schooling," *Journal of Human Resources*, 31, 1–26.

Ellsberg, D. (1961), "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

Erev, I., and Cohen, B. (1990), "Verbal Versus Numerical Probabilities: Efficiency, Biases, and the Preference Paradox," *Organizational Behavior and Human Decision Processes*, 45, 1–18.

Federal Reserve Consultant Committee on Consumer Survey Statistics (1955), Smithies Committee report in *Reports of the Federal Reserve Consultant Committees on Economic Statistics*, Hearings of the Subcommittee on Economic Statistics of the Joint Committee on the Economic Report, 84th U.S. Congress.

Freeman, R. (1971), *The Market for College-Trained Manpower*, Cambridge, MA: Harvard University Press.

Friedman, M. (1957), *A Theory of the Consumption Function*, Princeton, NJ: Princeton University Press.

Guiso, L., Jappelli, T., and Terlizzese, D. (1992), "Earnings Uncertainty and Precautionary Saving," *Journal of Monetary Economics*, 30, 307–337.

Hall, R., and Mishkin, F. (1982), "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50, 461–477.

Hogarth, R. (1975), "Cognitive Processes and the Assessment of Subjective Probability Distributions," *Journal of the American Statistical Association*, 70, 271–289.

Hubbard, G., Skinner, J., and Zeldes, S. (1995), "Precautionary Saving and Social Insurance," *Journal of Political Economy*, 103, 360–399.

Hurd, M., and McGarry, K. (1995), "Evaluation of Subjective Probabilities of Mortality in the HRS," *Journal of Human Resources*, 30, S268–S292.

Jamieson, L., and Bass, F. (1989), "Adjusting Stated Intentions Measures to Predict Trial Purchase of New Products: A Comparison of Models and Methods," *Journal of Marketing Research*, 26, 336–345.

Juster, T. (1964), *Anticipations and Purchases*, Princeton, NJ: Princeton University Press.

——— (1966), "Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design," *Journal of the American Statistical Association*, 61, 658–696.

Katona, G. (1957), "Federal Reserve Board Committee Reports on Consumer Expectations and Savings Statistics," *Review of Economics and Statistics*, 39, 40–46.

Keynes, J. (1921), *A Treatise on Probability*, London: Macmillan.

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

——— (1993), "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?," in *Studies of Supply and Demand in Higher Education*, eds. C. Clotfelter and M. Rothschild, Chicago: University of Chicago Press, pp. 43–57.

Morgan, G., and Henrion, M. (1990), *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*, New York: Cambridge University Press.

National Bureau of Economic Research (1960), *The Quality and Economic*

Significance of Anticipations Data, Special Conference Series, Princeton: Princeton University Press.

Patterson, G. (1991), "Consumer-Confidence Surveyor Has Become an Economic Guru," *The Wall Street Journal*, May 2.

Quadrel, M., Fischhoff, B., and Davis, W. (1993), "Adolescent (In)vulnerability," *American Psychologist*, 48, 102–116.

Savage, L. (1971), "Elicitation of Personal Probabilities and Expectations," *Journal of the American Statistical Association*, 66, 783–801.

Schuman, H., and Johnson, M. (1976), "Attitudes and Behavior," *Annual Review of Sociology*, 2, 161–207.

Skinner, J. (1988), "Risky Income, Life Cycle Consumption and Precautionary Savings," *Journal of Monetary Economics*, 22, 237–255.

Turner, C., and Martin, E. (Eds.) (1984), *Surveying Subjective Phenomena*, Vol. 1, New York: Russell Sage Foundation.

U.S. Bureau of the Census (1988), Current Population Reports, Series P-20, No. 427, *Fertility of American Women: June, 1987*, Washington, DC: U.S. Government Printing Office.

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

Wallsten, T., and Budescu, D. (1983), "Encoding Subjective Probabilities: A Psychological and Psychometric Review," *Management Science*, 29, 151–173.

Wallsten, T., Budescu, D., Rapaport, A., Zwick, R., and Forsyth, B. (1986), "Measuring the Vague Meanings of Probability Terms," *Journal of Experimental Psychology: General*, 115, 348–365.

Wallsten, T., Budescu, D., Zwick, R., and Kemp, S. (1993), "Preferences and Reasons for Communicating Probabilistic Information in Verbal or Numerical Terms," *Bulletin of the Psychonomic Society*, 31, 135–138.

Willis, R., and Rosen, S. (1979), "Education and Self-Selection," *Journal of Political Economy*, 87, S7–S36.

Winsborough, H. (1987), "The WISCON Survey: Wisconsin's Continual Omnibus, National Survey," University of Wisconsin-Madison, Center for Demography and Ecology.

Zeldes, S. (1989), "Optimal Consumption With Stochastic Income: Deviations From Certainty Equivalence," *Quarterly Journal of Economics*, 104, 275–298.

Zimmer, A. (1983), "Verbal vs. Numerical Processing of Subjective Probabilities," in *Decision Making Under Uncertainty*, ed. R. Scholz, Amsterdam: North-Holland.

——— (1984), "A Model for the Interpretation of Verbal Predictions," *International Journal of Man-Machine Studies*, 20, 121–134.