Using Expectations Data to Study Subjective Income Expectations

Jeff Dominitz
Institute for Social Research
University of Michigan

and

Charles F. Manski
Department of Economics
University of Wisconsin-Madison

October 1994

Abstract

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.

1. Introduction

Economic thinking about household behavior assigns a central role to expectations as a determinant of decisions. In particular, the allocation of current income between consumption and savings is assumed to reflect expectations of future income. Yet studies of consumption and other behaviors are unable to draw on empirical knowledge of income expectations. Skeptical of subjective data of all kinds, economists do not ordinarily collect data on income expectations. Instead, the standard approach is to infer expectations from panel data on income realizations. See, for example, Hall and Mishkin (1982), Skinner (1988), Zeldes (1989), Caballero (1990), and Carroll (1992).[1]

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. Our 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?"

---

[1] In an often cited article, MaCurdy (1982) estimated various models of the stochastic process generating income realizations in panel data. He did not assert that income expectations coincide with this process but Skinner (1988), Zeldes (1989), and Caballero (1990) interpret his findings in this manner.

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 our estimates to study the cross-sectional variation in income expectations one year into the future and report the findings in this paper.

We decided to collect income-expectations data after reflecting on two matters. First, the conventional alternative of inferring expectations from income realizations is not very attractive. A researcher seeking to learn expectations from realizations must assume that he or she knows what information households possess and how they use the available information to form expectations. Moreover, the available data on realizations must be rich enough for the researcher to emulate the assumed processes of expectations formation. These are strong requirements.[2] [3]

Second, examination of the history of economic thinking about expectations

---

[2] Recent studies inferring income expectations from panel data on income realizations share other features that are not fundamental but that facilitate empirical analysis. Each agent is assumed to know that his income stream is a realization of a specified stochastic process. Each agent is assumed to use his knowledge of this process and his observation of past income realizations to form expectations of future income conditional on past income. The analyst assumes that he or she knows the income-generating process up to some parameters and uses the available panel data on income realizations to estimate the parameters.

[3] The present discussion focuses on the use of realizations to infer unconditional income expectations. Distinct empirical literatures use income realizations to infer income expectations conditional on schooling, occupation, marriage and other decision variables. What is said here applies to these literatures as well. See Manski (1993) for a discussion focusing on the problem of inference on the subjective returns to schooling.

data shows that the longstanding general skepticism of such data is based on only a narrow foundation.   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 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) that found data of the type collected at SRC to have limited predictive value.   By the mid-1960s, opinion among mainstream economists was firmly negative.[4]

---

[4] Nevertheless, SRC has continued to collect qualitative expectations data and publishes aggregate findings monthly in its Index of Consumer Sentiment.   See Curtin (1976) and Patterson (1991).

We find compelling the criticisms of expectations data made by economists forty years ago (see Manski, 1990).   But the weakness of vaguely worded qualitative questions in measuring expectations implies nothing about the usefulness of more tightly worded probabilistic questions such as those we ask. There is no empirical evidence supporting condemnation of all expectations data.[5]

Section 2 describes the questions we ask to elicit income expectations. We examine the response rate to these questions and to questions asking respondents for their realized household income in the past twelve months.   We explain how we use our expectations data to fit respondent-specific subjective distributions for next year's income.

Section 3 analyzes the uncertainty respondents report concerning next year's income.   We use our empirical findings to assess the assumptions about income uncertainty imposed in studies using realizations to infer expectations. We also compare our findings with those of Guiso et al. (1992), who have recently analyzed the income expectations of Italian households.

Section 4 examines how respondents' income expectations vary with their realized incomes and other attributes.   We use the SEE data to estimate best

---

[5] Lacking empirical evidence, some economists dismiss subjective data a priori by asserting that respondents to surveys have no incentive to answer questions carefully or honestly.   Hence, they conclude, there is no reason to think that subjective responses reliably reflect respondents' thinking.   If this reasoning is to be taken seriously, it should be applied to survey data on realizations and not just to subjective data.   But economists do not dismiss self-reports of realizations a priori. Empirical economic analyses of household behavior routinely use self-reports of realized income, assets, employment, and other variables.

empirical predictors of expectations conditional on realizations data of the type available in major household surveys.   Section 5 gives conclusions.

The reader interested in other work using probabilistic questions to learn about household economic expectations would do well to begin with Juster (1966), who found that elicited subjective probabilities of consumer durable purchases predict subsequent purchase behavior better than do the responses to binary (i.e., yes/no) purchase intentions questions.   Juster's work was long ignored by economists but it has influenced empirical practices among market researchers (see Morrison, 1979, Urban and Hauser, 1980, and Jamieson and Bass, 1989) and was among the factors affecting our decision to undertake the work reported here. A recent descendant of Juster's early work is a set of probabilistic questions on retirement and longevity expectations in the Health and Retirement Survey. See Juster and Suzman (1993) for a description of the data and Hurd and McGarry (1993) for an initial analysis.

2. Eliciting Subjective Distributions of Future Income

2.1. The Survey of Economic Expectations

From February to May 1993, we placed a set of probabilistic expectations questions concerning future income, earnings, and employment as a module in WISCON, a national continuous telephone survey conducted by the Letters and Science Survey Center at the University of Wisconsin-Madison.[6] The WISCON core questions ask respondents about their labor market experiences, demographics, household income, and qualitative expectations (see Winsborough, 1987). We refer to the edition of WISCON containing our probabilistic expectations module as the Survey of Economic Expectations (SEE).

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) discuss 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

---

[6] Two sets of pretests of the expectations questions were conducted in late 1992. Each pretest was followed by revisions of procedures, yielding those described in this section. The present paper focuses on the income-expectations data collected in SEE but the data on earnings and employment expectations are also of interest. The latter data are analyzed in Dominitz (1994).

procedures to fit the survey medium (telephone interview) and subject matter.[7]

As indicated in the introduction, we asked each respondent to provide a sequence of points on his or her subjective CDF of household income over the next twelve months. In particular, respondent i was asked about four income thresholds $Y_{i1}, \ldots, Y_{i4}$, posed in increasing order.[8] The interviewer informed the respondent if a probability elicited at threshold Y_{i2}, Y_{i3}, or Y_{i4} was smaller than one elicted earlier. In this way, we ensured that the sequence of responses was always logically coherent.

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 possible" and "highest possible" incomes that the household might experience in the next year. We did not interpret the answers to these preliminary questions literally as minimum and maximum incomes. Rather, we used them to indicate the general region of the respondent's subjective support of

---

[7] Eliciting expectations by telephone makes it 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 successfully elicited the medians of subjective income distributions in our study of the returns-to-schooling expectations of high school students (Dominitz and Manski, 1994). That study uses the medium of an interactive computer program to elicit expectations. Questions are posed on-screen and respondents key-in their answers.

[8] The only exception occurs if a response of "100 percent 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 percent chance" to all subsequent thresholds.

future income. Our reasoning was that the 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.[9] Our computer-programmed question-branching algorithm selected income thresholds Y in this region.[10] See Appendix Section A.1 for the precise wording of the questions we posed.

2.2. Response Rates

The WISCON interviewers obtain a telephone interview from slightly over fifty percent of the households with whom contact is attempted. Our expectations module was administered to 622 WISCON respondents. Of these, 437 gave usable

---

[9] Morgan and Henrion (1990) offer another reason for asking these preliminary questions, namely 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, the respondent may be influenced to think that this amount is objectively reasonable and so may report a higher probability than believed a priori.

[10] In particular, each respondent was asked about four consecutive income thresholds selected from this set of candidate values: {$5000, $10000, $15000, $20000, $25000, $30000, $35000, $40000, $50000, $60000, $70000, $80000, $100000, $125000, $150000}. The midpoint between the respondent's elicited "lowest possible" and "highest possible" incomes was used to determine the thresholds about which the respondent was queried.

responses to the entire sequence of income expectations questions.[11] Thus, the

effective item-response rate for the expectations module was about 0.70.

It is of interest to compare response to the income expectations module

with response to the set of WISCON core questions eliciting household income in

the past twelve months.[12] 390 of the 622 respondents answered the questions

asking for their income and, if applicable, the incomes of their spouse and of

other adult members of the household. Thus, the response rate for realized

household income was about 0.63. Response to the expected income and realized

income questions was positively associated, as shown below:

response to both expected and realized income   : 331 respondents (.53)

response to expected income only   : 106 respondents (.17)

response to realized income only   : 59 respondents (.09)

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 noteworthy

---

[11] 509 of the 622 respondents answered the preliminary questions eliciting
their lowest and highest possible incomes in the next year. 489 of the 509
answered the subsequent questions eliciting points on the subjective CDF. Our
analysis excludes 52 respondents who reported the same probability values at all
four of the income thresholds posed. (22 respondents answered "100 percent" to
the first threshold posed, implying the same answer for all subsequent
thresholds, 15 answered "0 percent" four times, and 15 reported a single value
between 0 and 100 four times.) We cannot use these 52 observations to fit
respondents' subjective distributions in the manner to be described in Section
2.3.

[12] See Appendix Section A.2 for the form of these questions.

patterns.   Males responded to each set of questions more often than did females (.78 versus .65 and .69 versus .58).   Response rates first rise with age and then fall: sample members aged 40-49 responded most frequently (.82 and .77), those aged 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 did those with less than a high school diploma (.44 and .49).

Table 1: Response Rates for Respondents with Various Attributes

	Number of Respondents	Response Rate Expected Income	Realized Income
Total Sample	622	.70	.63
Demographics			
Female	358	.65	.58
Male	264	.78	.69
White	535	.67	.62
Non-White	79	.63	.68
Single	237	.65	.69
Married	353	.74	.60
Cohabit	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
$\geq$ 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

2.3. Fitting Subjective Income Distributions

After division by 100, we interpret a respondent's answers to the four expectations questions as points on his or her subjective CDF of household income over the next twelve 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,...,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, Chapters 2 and 3). It facilitates analysis if we use the expectations 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 $F(Y; m_i, r_i)$.[13] In particular, we use the median m_i to characterize the central tendency of respondent i's subjective income distribution and the interquartile range (henceforth, IQR) r_i to characterize its spread.

2.4. The Empirical Distribution of Fitted Income Expectations

There is inevitably some arbitrariness in using any specific criterion (here least squares) to fit the expectations data to any specific parametric family of distributions (here the log-normal distributions). The most compelling evidence we can offer for the success of our approach to eliciting income expectations is the reasonableness of our findings. Table 2 tabulates the medians m and IQRs r of the log-normal distributions fitted to the expectations data elicited from the 437 respondents with usable responses.

---

[13] We say "inf" rather than "min" because the least squares solution is a degenerate log-normal distribution in some cases. In particular, this occurs whenever at least three of the four elicited probabilities (F_{ik}, k = 1,...,4) take the value zero or one. 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}. 85 of the 437 respondents gave responses that imply degenerate solutions to the least squares problem (see Table 2 below).

Table 2: Income Expectations of the Respondents
_____

	Subjective Interquartile Range (r)							Totals
	[0]	(0, 5)	[5, 10)	[10, 15)	[15, 25)	[25, ∞)		
Subjective Median (m)								
[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

Note: The entry in each cell is the number of respondents whose fitted log-normal distribution has median m and interquartile range r. The units of m and r are thousands of dollars.

Source: Survey of Economic Expectations
_____

Our analysis of the fitted subjective income distributions is in two parts. Section 3 examines respondents' subjective income uncertainty. Section 4 investigates how respondents' expectations vary with their realized income and other attributes.

3. Subjective Income Uncertainty

3.1. 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. In Hall and Mishkin (1982), the subjective distribution of next year's income is assumed normal with household-specific mean μ_i and constant variance σ^2 or, equivalently, with constant IQR $1.349 \cdot \sigma$. Hall and Mishkin estimate σ to be 8.1 thousand dollars (in 1993 dollars).[14] Thus, they estimate the subjective IQR of next year's income to be 10.9 thousand dollars for all households.

In Skinner (1988) and Zeldes (1989), the subjective distribution of next year's log-income is 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 $m_i[\exp(0.6745 \cdot \delta) - \exp(-0.6745 \cdot \delta)]$. Thus, the IQR in these studies is proportional to the subjective median m_i.

Carroll (1992) assumes the same form for expectations as do Skinner and Zeldes except that he superimposes a .005 chance of receiving no income at all.

---

[14] Hall and Mishkin did not state what year's price level they used in converting nominal income to real terms. 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.

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 to 1985, Carroll estimates δ to be 0.192. Thus, he estimates household i's subjective IQR of next year's income to be about $0.26 \cdot m_i$.

3.2. Empirical Findings

Table 3 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.[15] The reader should be aware that the table 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)*.[16]

---

[15] These estimates use the standard normal density as kernel and a bandwidth of 10 thousand dollars. See Härdle (1990) and Ullah and Vinod (1993) for expositions of kernel and other nonparametric methods for estimation of conditional quantiles.

[16] For example, the entry in the first row of the column headed "Subjective Median (m)" shows that ten percent 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 thousand dollars. The associated confidence interval (13.7, 15.6) is an interval estimate for the unknown value $Y_{.10}$ appearing in this statement: ten percent of potential respondents believe there to be at least a

* Table 3 shows that our subjective IQR estimates based on expectations data

Table 3: 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
0.10	15.0 (13.7, 15.6) (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0
0.25	22.8 (21.3, 25.0) (4.1, 7.2)	3.1 (2.3, 4.2)	1.4 (0, 2.6)	4.2 (3.1, 4.8)	5.3
0.50	37.9 (36.2, 40.2) (10.6, 14.8)	9.6 (8.3, 10.5)	6.7 (5.7, 8.5)	9.9 (8.5, 11.3)	11.7
0.75	59.7 (55.0, 65.7) (17.7, 23.8)	17.4 (16.4, 18.2)	13.6 (11.7, 14.4)	17.0 (15.1, 17.7)	19.0
0.90	80.1 (77.8, 96.0) (26.3, 32.6)	31.3 (27.9, 34.9)	22.6 (17.4, 23.7)	28.0 (23.7, 29.5)	32.1

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 percent 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

50-50 chance that their income in the next year would be no greater than $Y_{.10}$ thousand dollars.

have the same order of magnitude as the estimates implied by recent studies using income realizations to infer expectations. The empirical median of r presented in the second column of Table 3 is 9.6 thousand dollars. The Hall and Mishkin estimate of 10.9 thousand dollars is close to this figure. So is the Carroll estimate of 0.26·m when computed at the empirical median of m presented in the first column of Table 3. Setting m = 37.9, the Carroll estimate of IQR is 9.9 thousand dollars.

Although the estimates based on realizations data and expectations data have the same order of magnitude, Table 3 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. r tends to increase with m, but more slowly than proportionately. The kernel-smoothed empirical median of r increases from 6.7 to 9.9 to 11.7 thousand dollars as m increases from 20 to 40 to 60 thousand dollars.[17]

3.3. Income Uncertainty in Italy

We are aware of only one other household survey using probabilistic questions to elicit income expectations, that being the 1989 edition of the Survey of Household Income and Wealth (SHIW), the Bank of Italy's biennial survey

---

[17] Observe that the .10-quantile of r is zero; that is, r equals zero for at least ten percent of the respondents. These are the respondents discussed in note 13, whose fitted log-normal distributions are degenerate.

of the Italian population. As described by Guiso et al. (1992), the SHIW
elicited subjective probability distributions for the growth rate of nominal
labor earnings and pensions and for the rate of inflation over the next twelve
months. In particular, respondents were asked to report the subjective
probability that these rates would fall in each of the following 12 intervals
(numbers correspond to percentage points):

    <0, 0-3, 3-5, 5-6, 6-7, 7-8, 8-10, 10-13, 13-15, 15-20, 20-25, >25.
Guiso et al. use the responses to estimate respondents' subjective distributions
of real head-of-household earnings over the next twelve months. In particular,
they use the ratio σ/μ of the standard deviation to the mean of the subjective
distribution to measure subjective earnings uncertainty.

The values of σ/μ found in the Italian study are much smaller than those
found in our study of American households (and also much smaller than those found
in American studies using income realizations to infer expectations). Examine
Table 4 which compares the Guiso et al. empirical distribution of σ/μ with ours
derived from the fitted log-normal distributions. Whereas the value of σ/μ is
less than .025 for 88 percent of the SHIW respondents, it is less than .025 for
only 20 percent of the SEE respondents. Whereas the value of σ/μ is less than
.100 for all of the SHIW respondents, it is less than .100 for only 34 percent
of the SEE respondents.

Table 4: Empirical Distribution of σ/μ in SHIW and SEE

	Italian SHIW	American SEE
$Pr[\sigma/\mu = 0.000]$	0.34	0.20
$Pr[\sigma/\mu \leq 0.005]$	0.44	0.20
$Pr[\sigma/\mu \leq 0.015]$	0.70	0.20
$Pr[\sigma/\mu \leq 0.025]$	0.88	0.20
$Pr[\sigma/\mu \leq 0.035]$	0.94	0.21
$Pr[\sigma/\mu \leq 0.045]$	0.97	0.22
$Pr[\sigma/\mu \leq 0.065]$	0.99	0.24
$Pr[\sigma/\mu \leq 0.100]$	1.00	0.34
$Pr[\sigma/\mu \leq 0.150]$	1.00	0.44
$Pr[\sigma/\mu \leq 0.200]$	1.00	0.53
$Pr[\sigma/\mu \leq 0.300]$	1.00	0.70
$Pr[\sigma/\mu \leq 0.400]$	1.00	0.78
$Pr[\sigma/\mu \leq 0.500]$	1.00	0.85
$Pr[\sigma/\mu \leq 1.000]$	1.00	0.94
$Pr[\sigma/\mu \leq 2.000]$	1.00	0.98
$Pr[\sigma/\mu \leq 5.000]$	1.00	0.99

Sources: Survey of Household Income and Wealth (SHIW) and Survey of Economic Expectations (SEE)

There are too many differences between the SHIW and SEE instruments and sample designs for us to be willing to engage in any refined comparison of the two sets of findings. Nevertheless, the differences in the empirical distribution of σ/μ are so large that we are tempted to draw the obvious conclusion that American households perceive far more income uncertainty than do Italian ones.

4. Using SEE data to Predict Expectations Conditional on Realizations

We think that major household surveys should regularly ask questions eliciting income and other expectations thought to be important determinants of decision making. Until that happens, researchers will continue to have to learn about expectations in less direct 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 m and IQRs r. The available 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.

The estimates in Table 5 are based on the 324 respondents reporting complete realizations and expectations data.[18] Statistics describing the outcome

---

[18] As indicated in note 11, there were 331 respondents who provided usable data on income realizations and expectations. The present analysis excludes 7 of these 331 who provided incomplete data about some of the attributes used as predictors in Table 5.

and predictor variables of these 324 respondents are given in Table 6. Sections

4.1 and 4.2 examine the findings.

Table 5: Best Linear (LAD) Prediction of Expectations Conditional on Realizations

	Subjective Median (m)	Subjective IQR (r)
Predictor		
Household income in past twelve months (10^3 dollars)	0.896 (0.036)	0.172 (0.032)
Labor force participation by respondent and spouse (2 if both, 1 if either)	-1.088 (1.465)	-0.106 (1.339)
Unemployment indicator (1 if respondent or spouse is unemployed)	-4.439 (2.069)	10.193 (6.220)
Gender of respondent (1 if female)	-1.661 (1.013)	-1.083 (1.244)
Marital Status of respondent (1 if married or cohabiting)	-0.999 (1.132)	0.266 (1.149)
Age of respondent (years)	-0.014 (0.047)	-0.138 (0.050)
Respondent has postsecondary schooling but no Bachelor's degree (1 if yes)	-0.915 (1.172)	-1.105 (1.712)
Respondent has Bachelor's degree (1 if yes)	-1.077 (1.263)	0.517 (1.621)
Constant	7.333 (3.710)	8.316 (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

Table 6: Descriptive Statistics for the Variables in Table 5

Variable	Mean	Std. Dev.	Min	Max
subjective median	47.7	32.8	0.0	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	0.67	0	2
unemployment indicator	0.04	0.19	0	1
respondent gender	0.51	0.50	0	1
respondent marital status	0.62	0.49	0	1
respondent age	43.6	14.6	18	85
postsecondary schooling	0.44	0.50	0	1
Bachelor's degree	0.35	0.48	0	1

4.1. 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 23.9 thousand dollars, the average absolute deviation between m and its BLP is just 10.2 thousand dollars. Thus, the BLP "explains" more than half the empirical variation in m. We consider the predictive power of the BLP of m to be remarkably good, especially when it is remembered that m

is derived by fitting a log-normal distribution to the raw expectations data obtained from each respondent.

Realized household income, henceforth denoted y, is the dominant predictor variable. The estimated BLP of m increases 896 dollars with every one thousand dollar increase in y. Other respondent attributes have modest or negligible effects on the BLP. The predicted value of m in a household where the respondent or spouse is unemployed is 4.4 thousand dollars lower than in a household where neither is unemployed, ceteris paribus. The predicted value of m varies little, if at all, 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 $\text{Med}(m|y)$, a kernel-smoothed estimate of the population median of m conditional on y. A bootstrapped ninety-percent confidence interval is also shown.

The graph shows $\text{Med}(m|y)$ to be close to a linear function of y. The confidence interval is quite tight at all but the highest income levels. Observe that $\text{Med}(m|y) > y$ at low values of y and and $\text{Med}(m|y) < y$ 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.

4.2. Predicting the IQRs of Subjective Income Distributions

The findings on prediction of respondents' fitted subjective IQRs r may be less dramatic but still are interesting. The average absolute deviation between r and its empirical median is 10.7 thousand dollars, while the average absolute deviation between r and its BLP is 9.5 thousand dollars. Thus, the BLP "explains" about 11 percent of the empirical variation in m.

Realized household income is an important predictor variable but is not the only important one. The estimated BLP of r increases 172 dollars with every one thousand dollar increase in y. The BLP decreases 138 dollars with every year increase in the age of the respondent. The predicted value of r in a household where someone is unemployed is fully 10.2 thousand dollars 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 percent 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 non-decreasing 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, however, too
wide for us to draw firm conclusions.

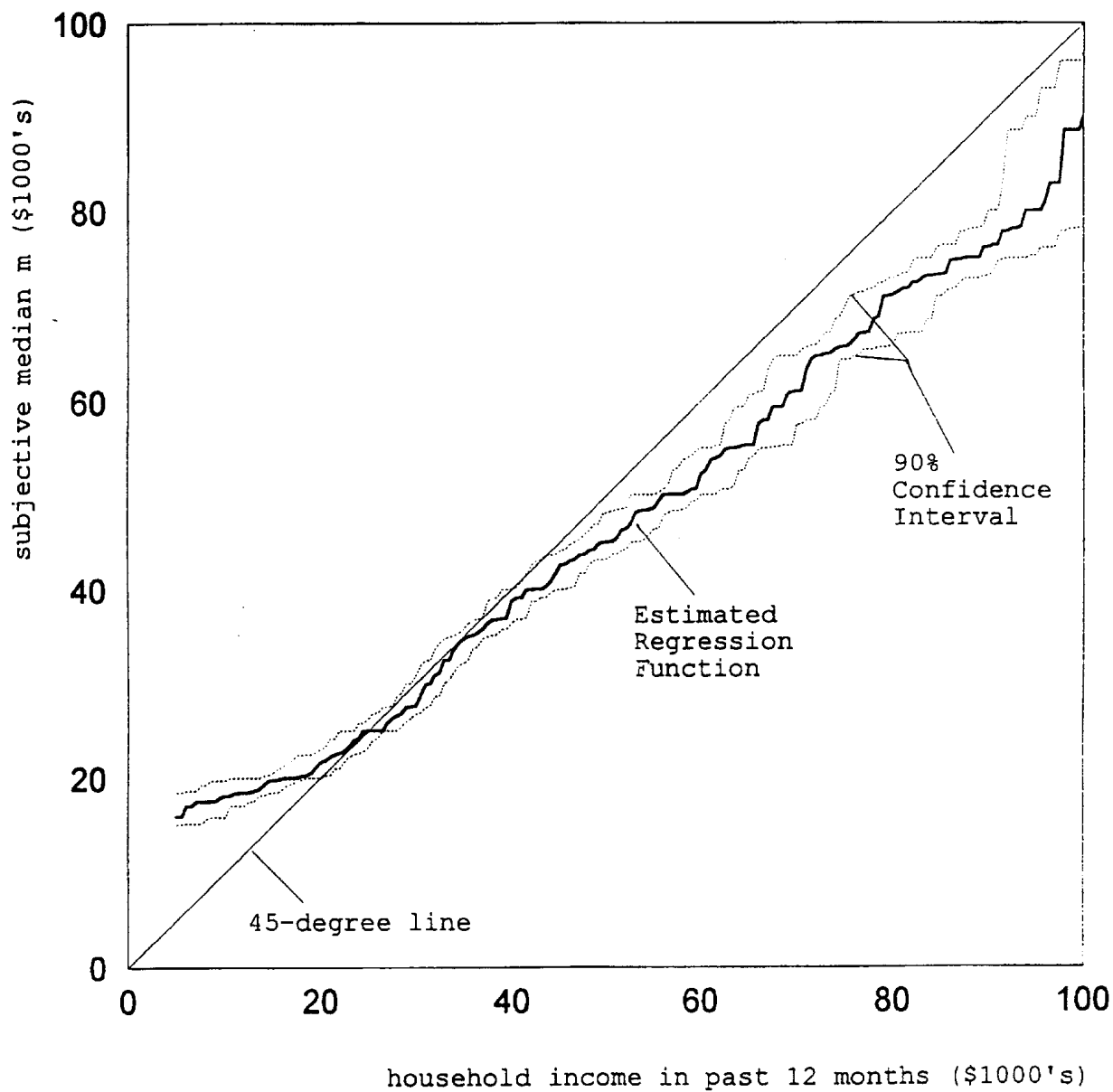Figure 1: Best Prediction of Expectations Conditional on Realizations
Med(m|y)

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 percent 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

5. Conclusion

We began this work with some concern about the feasibility of eliciting meaningful income expectations in a telephone survey. We conclude with a clear sense that elicitation is not only feasible but that the specific way we pose questions and fit subjective distributions, described in Section 2, works quite well. Figure 1, which shows the close association between realized income and fitted subjective median income, provides especially striking evidence of this.

From a substantive perspective, the most interesting part of this study may be our findings on subjective income uncertainty, reported in Section 3. Lacking expectations data, economists have only been able to speculate about the uncertainty that persons perceive concerning their future incomes. In studies inferring expectations from realizations, it has been common to assume that the spread and central tendency of income expectations are proportional to one another. We find that the subjective IQR of future income does tend to rise with the subjective median, but more slowly than proportionately. We also find that, conditioning on any specified value of the subjective median, the subjective IQR varies substantially across respondents.

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. Section 4 shows how the SEE data may be used to estimate best empirical predictors of expectations conditional on realizations data. These best predictors improve on the conventional approach of inferring expectations from realizations data alone.

Appendix: Questions Eliciting Expected and Realized Income

A.1. 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 houshold 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.2. 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.

And, just 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.)

References

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 R. Curtin, ed., Surveys of Consumers: 1974-75, 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 C. Manski (1994), "Eliciting Student Expectations of the Returns to Schooling," Department of Economics, University of Wisconsin-Madison.

Dynan, K. (1993), "How Prudent Are Consumers?" Journal of Political Economy, 101: 1104-1113.

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.

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

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

Härdle, W. (1990), Applied Nonparametric Regression, Cambridge, UK: Cambridge University Press.

Hurd, M. and K. McGarry (1993), "Evaluation of Subjective Probability Distributions," Department of Economics, State University of New York at Stoney Brook.

Jamieson, L. and F. Bass (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: Princeton University Press.

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

Juster, T. and R. Suzman (1993), "The Health and Retirement Study: An Overview," Core Paper No. 94-1001, Health and Retirement Study Working Paper Series, Institute for Social Research, University of Michigan.

Katona, G. (1957), "Federal Reserve Board Committee Reports on Consumer Expectations and Savings Statistics," <u>Review of Economics and Statistics</u>, 39: 40-46.

MaCurdy, T. (1982), "The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis," <u>Journal of Econometrics</u>, 18: 83-114.

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

Manski, C. (1993), "Adolescent Econometricians: How Do Youth Infer the Returns to Schooling?" in C. Clotfelter and M. Rothschild, eds., <u>Studies of Supply and Demand in Higher Education</u>, Chicago: University of Chicago Press.

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

Morrison, D. (1979), "Purchase Intentions and Purchase Behavior," <u>Journal of Marketing</u>, 43, 65-74.

National Bureau of Economic Research (1960), <u>The Quality and Economic Significance of Anticipations Data</u>, Special Conference Series, Princeton: Princeton University Press.

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

Skinner, J. (1988), "Risky Income, Life Cycle Consumption and Precautionary Savings," <u>Journal of Monetary Economics</u>, 22: 237-255.

Ullah, A. and H. Vinod (1993), "General Nonparametric Regression Estimation," in C. R. Rao, G. S. Maddala, and H. Vinod (editors), <u>Handbook of Statistics, Vol. 11: Econometrics</u>, Amsterdam: North-Holland.

Urban, G. and J. Hauser(1980), <u>Design and Marketing of New Products</u>, Englewood Cliffs: Prentice-Hall.

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

Zeldes, S. (1989), "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence," <u>Quarterly Journal of Economics</u>, 104: 275-298.