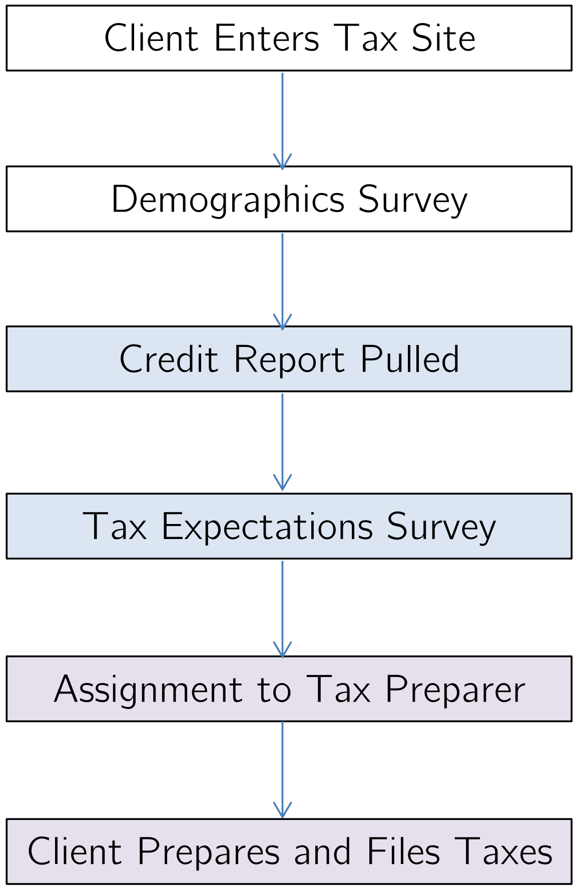
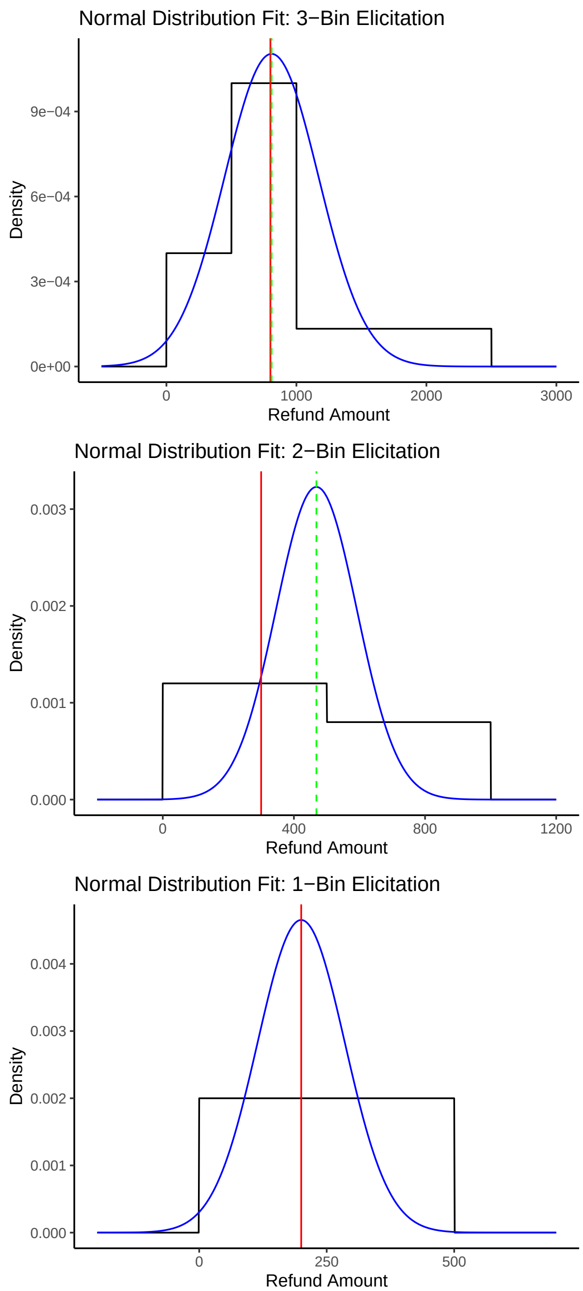
A For Online Publication: Appendix Tables and Figures

Figure A1: Tax Site Client Flow



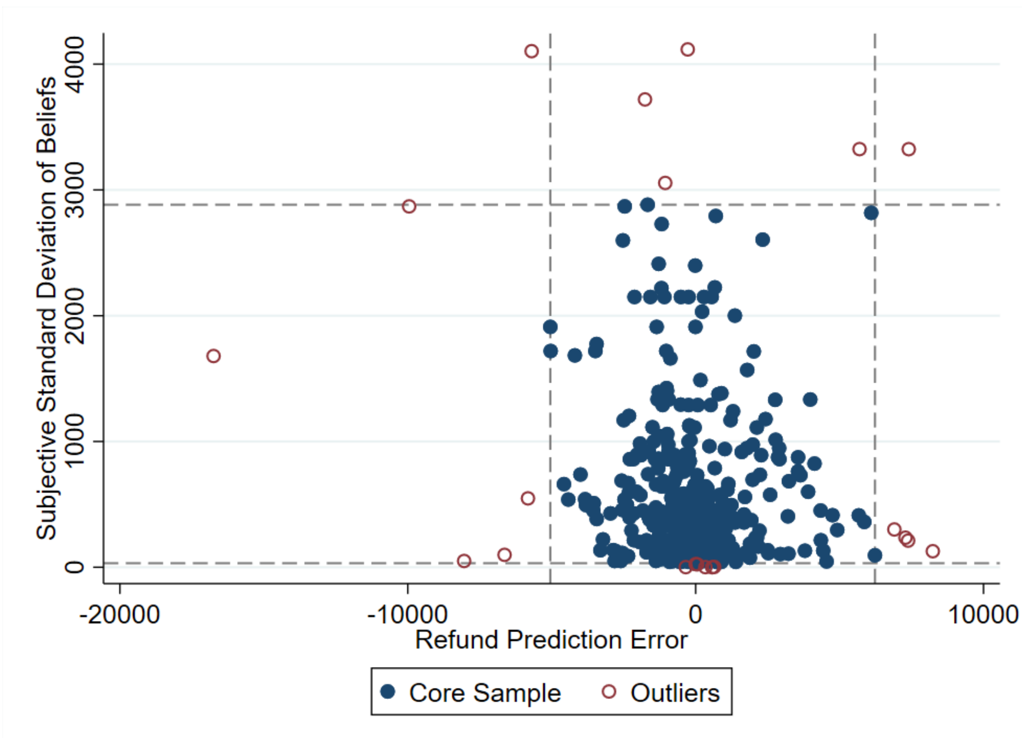
Note: This figure shows the steps a tax filer would go through upon arriving at the center. The steps in white occur before a filer has met with a financial guide or tax preparer. The steps in blue are completed in collaboration with one of the site's financial guides. Filers provided consent for their tax, credit, and survey information to be used for research purposes immediately prior to the tax expectations survey. The steps in purple are completed with the help of a volunteer tax preparer.

Figure A2: Fitting Beliefs to Normal Distributions



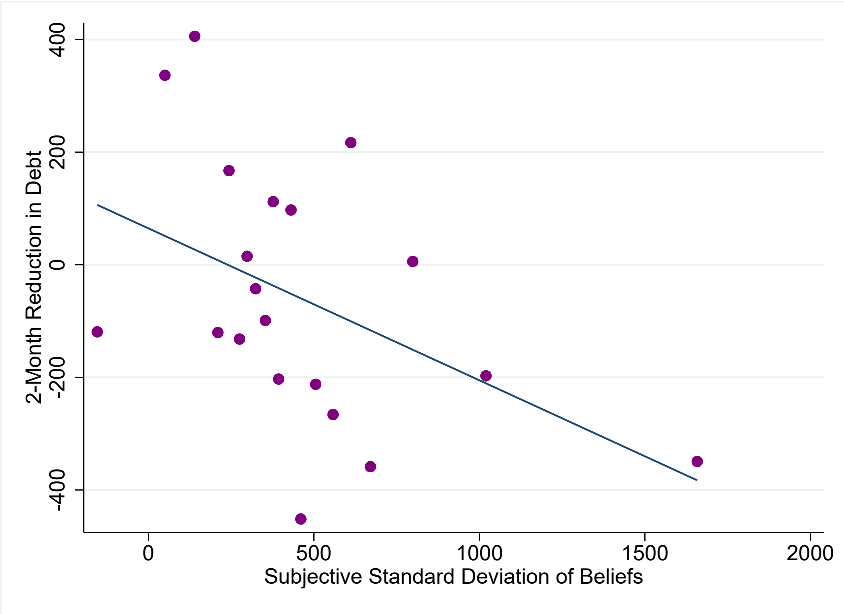
Note: This figure shows how we fit probabilistic beliefs to normal distributions if the individual places positive mass in 3 or more bins (top); in 2 bins (middle); or in 1 bin (bottom). Each solid black line represents the distribution of beliefs implied by a uniform density within each bin. Each red line denotes the filer's point estimate. Each blue curve is the density of the fitted normal distribution, with the corresponding dashed green line showing the mean of this distribution. More information on how we fit beliefs to normal distributions is provided in Section 3. Graphs describing how we fit beliefs to beta distributions are provided in Figure A6. Table 2 presents descriptive statistics on the fitted beliefs.

Figure A3: Expectation Outliers and Core Sample



Note: This figure plots the fitted standard deviation of subjective beliefs about refund size against realized refund prediction errors. Dotted lines denote the thresholds at which the top and bottom 1% of refund prediction errors and the top and bottom 1% of subjective standard deviations are excluded as outliers. Solid circles represent the core sample excluding outliers, and hollow circles represent the outliers. See Tables 1, A1, and A2 for summary statistics on these two groups.

Figure A4: Refund Uncertainty and Borrowing Behavior



Note: This figure shows a binned scatterplot of 2-month repayment of non-installment balances against subjective uncertainty corresponding to the regression specification in equation 2. These data are plotted after partialling out the demographic and tax filer characteristics included in column 4 of Table 3.

Table A1: Descriptive Statistics

	Core Sample	All Filers			
	Tax Data & Expectations Data	Tax Data & Expectations Data	Tax Data, Expectations Data, & Demographics	Current and Prior Tax Data & Expectations Data	Tax Data, Expectations Data, & Credit Data
	(1)	(2)	(3)	(4)	(5)
Demographic Characteristics					
Female	0.62	0.62	0.62	0.65	0.68
	(0.49)	(0.49)	(0.49)	(0.48)	(0.47)
Age	40.21	40.46	40.29	42.82	41.79
	(15.92)	(15.90)	(15.78)	(15.76)	(15.96)
High School or Above	0.82	0.82	0.82	0.84	0.85
	(0.38)	(0.39)	(0.39)	(0.37)	(0.36)
BA Degree	0.15	0.15	0.15	0.17	0.20
	(0.36)	(0.36)	(0.36)	(0.38)	(0.40)
Economic and Tax Characteristics					
Adjusted Gross Income ($)	20,636.93	20,997.95	21,040.55	23,844.29	24,310.66
	(15930.39)	(15941.50)	(15776.71)	(16125.67)	(16189.67)
Has Dependents	0.32	0.32	0.32	0.36	0.35
	(0.47)	(0.47)	(0.47)	(0.48)	(0.48)
Married	0.08	0.08	0.07	0.07	0.08
	(0.27)	(0.28)	(0.26)	(0.25)	(0.28)
Single Head of Household	0.27	0.27	0.28	0.31	0.30
	(0.44)	(0.45)	(0.45)	(0.46)	(0.46)
Filed Schedule C	0.08	0.07	0.07	0.06	0.07
	(0.27)	(0.26)	(0.26)	(0.24)	(0.26)
Lost Job	0.08	0.08	0.07	0.07	0.06
	(0.27)	(0.26)	(0.26)	(0.25)	(0.23)
Tax Refund					
Refund Amount ($)	1,542.33	1,584.63	1,604.54	1,866.76	1,745.39
	(2207.11)	(2372.20)	(2383.22)	(2511.12)	(2507.85)
Received EITC	0.35	0.35	0.35	0.35	0.31
	(0.48)	(0.48)	(0.48)	(0.48)	(0.46)
EITC Credit (If >0)	1,654.16	1,730.11	1,722.83	2,007.79	1,957.06
	(1661.35)	(1703.41)	(1716.82)	(1796.08)	(1745.59)
EITC share	0.50	0.50	0.49	0.53	0.46
	(0.43)	(0.42)	(0.37)	(0.41)	(0.38)
Chose Direct Deposit	0.59	0.58	0.58	0.63	0.65
	(0.49)	(0.49)	(0.49)	(0.48)	(0.48)
Observations	618	692	616	375	400
with Demographics	548	616	616	339	357

Note: This table provides descriptive statistics on our sample of low-income filers. The first column describes our core sample, as shown previously in Table 1. The remaining columns show samples analogous to those shown in Table 1, here including outlier observations. These are individuals with subjective uncertainty in the top or bottom 1% of expectations survey respondents, or tax refund prediction errors in the top or bottom 1%, as well as individuals with adjusted gross incomes below 0. Additional descriptive statistics are provided in Table A2.

Table A2: Descriptive Statistics

	Core Sample	All Filers			
	Tax Data & Expectations Data	Tax Data & Expectations Data	Tax Data, Expectations Data, & Demographics	Current and Prior Tax Data & Expectations Data	Tax Data, Expectations Data, & Credit Data
	(1)	(2)	(3)	(4)	(5)
Savings and Credit					
Estimated Savings Balance	523.36	522.05	522.05	542.86	627.04
	(576.15)	(572.40)	(572.40)	(579.86)	(602.61)
FICO Score	666	664	663	672	682
	(87)	(86)	(86)	(87)	(80)
Credit Card Balances ($)	1,686	1,680	1,749	1,954	2,638
	(4,985)	(4,836)	(5,029)	(5,698)	(5,850)
Non-Mortgage Installment Balances	9,612	9,359	9,632	11,394	12,348
	(23,488)	(22,694)	(23,438)	(25,964)	(26,223)
Has Mortgage	0.04	0.05	0.05	0.06	0.06
	(0.21)	(0.21)	(0.22)	(0.24)	(0.23)
Filing Characterisstics					
Absolute Change in AGI	6.15	6.27	6.17	6.24	6.21
	(8.79)	(9.01)	(8.79)	(9.01)	(9.59)
Change in Filling Status	0.09	0.10	0.10	0.10	0.07
	(0.29)	(0.30)	(0.30)	(0.30)	(0.26)
Change in Number of Dependents	-0.04	-0.04	-0.04	-0.04	-0.04
	(0.55)	(0.57)	(0.59)	(0.57)	(0.54)
Any Change in Number of Dependents	0.13	0.13	0.14	0.13	0.10
	(0.33)	(0.34)	(0.35)	(0.34)	(0.30)
Observations	618	692	616	375	400
with Demographics	548	616	616	339	357

Note: This table provides descriptive statistics on our sample of low-income filers. The first column describes our core sample, as shown previously in Table 1. The remaining columns show samples analogous to those shown in Table 1, here including outlier observations. These are individuals with subjective uncertainty in the top or bottom 1% of expectations survey respondents, or tax refund prediction errors in the top or bottom 1%, as well as individuals with adjusted gross incomes below 0. Additional descriptive statistics are provided in Table A1.

Table A3: Elicited Beliefs by Filer Group

		Has Dependents		Marital Status		Any College		Relative to 2x Federal Poverty Line	
	Core Sample	Yes	No	Married	Not Married	Yes	No	Below	Above
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Number of Bins with Positive Probability									
1 Bin	22.2%	24.1%	21.3%	22.4%	22.1%	20.6%	24.4%	25.2%	20.5%
2 Bins	38.7%	39.0%	38.5%	36.7%	38.8%	37.3%	39.4%	41.3%	37.3%
3 Bins	20.7%	16.4%	22.7%	14.3%	21.3%	19.4%	20.1%	18.8%	21.8%
4 Bins	11.0%	11.3%	10.9%	12.2%	10.9%	13.5%	9.7%	8.3%	12.5%
5 Bins	5.0%	7.2%	4.0%	8.2%	4.7%	6.3%	3.9%	5.0%	5.0%
6 Bins	2.4%	2.1%	2.6%	6.1%	2.1%	2.8%	2.5%	1.4%	3.0%
Qualitative Uncertainty									
Very Certain	34.0%	30.3%	35.7%	44.9%	33.0%	32.5%	37.3%	38.5%	31.5%
Somewhat Certain	41.7%	48.2%	38.8%	36.7%	42.2%	38.9%	42.7%	39.9%	42.8%
Not Certain At All	23.5%	21.0%	24.6%	18.4%	23.9%	27.0%	19.7%	21.1%	24.8%
Observations	618	195	423	49	569	252	279	218	400

Notes: This table describes qualitative feature of tax filers' responses to the beliefs survey. All statistics are means within each group. Any College refers to any college experience regardless of degree attainment. Sample sizes for columns 6 and 7 reflect response rates to the Demographics survey (see Table 1), minus 17 respondents who did not provide education information.

Table A4: Parametric Belief Distributions

	Normal Distribution				Beta Distribution			
	Baseline	Exclude 50/50	Exclude Single Bins	All	Baseline	Exclude 50/50	Exclude Single Bins	All
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mean	1,605	1,641	1,322	1,678	1,837	1,905	1,435	1,932
	(2000)	(2061)	(1407)	(2187)	(2584)	(2698)	(1705)	(2796)
Median	1,605	1,641	1,322	1,678	1,943	2,026	1,582	2,068
	(2000)	(2061)	(1407)	(2187)	(3138)	(3299)	(2626)	(3407)
Std. Dev.	426	457	385	454	690	739	578	733
	(510)	(535)	(456)	(599)	(895)	(941)	(725)	(1005)
Observations	618	541	584	647	618	541	584	647

Notes: This table reports features of parametric belief distributions under alternative assumptions. Statistics are aggregated across all tax filers in the main analysis sample. Columns 1-4 present statistics based on the normal distributions fit to the probabilistic survey question described in Section 3. Columns 5-8 contain statistics based on beta and triangle distributions. We describe how we fit these distributions in Section 3 and Appendix E. Columns 1 and 5 present our baseline sample. Columns 2 and 6 exclude individuals who put 50/50 probabilities on two bins. Columns 3 and 7 exclude individuals who put 100% probability on a single bin. Columns 4 and 8 include all tax filers who filled out the expectations survey; the sample size in these two columns differs from that in Table A1, column 2, because parametric belief distributions are only available for individuals in the full sample who reported internally consistent beliefs, as described in footnote 10 of the text.

Table A5: Characteristics of Beliefs

	Subjective Standard Deviation		Absolute Prediction Error		Absolute Change in MTR
	(1)	(2)	(3)	(4)	(5)
Absolute Change in MTR	1.141**	0.0828	0.980**	-0.107	
	(0.450)	(0.564)	(0.460)	(0.485)	
Absolute Change in AGI		0.174***		0.0155	0.154***
		(0.0651)		(0.0439)	(0.0562)
Has Dependents		0.470***		0.444***	0.170**
		(0.0757)		(0.0808)	(0.0811)
Change in No. Dependents		-0.0454		0.216***	0.102*
		(0.0567)		(0.0799)	(0.0615)
Married		0.102		-0.0636	-0.102
		(0.0722)		(0.0651)	(0.0659)
Change in Filing Status		-0.0186		-0.0980	0.0491
		(0.0493)		(0.0629)	(0.0629)
Received UI during Past Year		0.0190		-0.0341	0.0412
		(0.0631)		(0.0549)	(0.0656)
Absolute Change in AGI * Has Dependents		-0.0286		-0.0189	0.0815
		(0.0698)		(0.0430)	(0.0888)
Demographics	X	X	X	X	X
Observations	337	337	337	337	337

Note: Estimates from ordinary least squares regressions. The dependent variable in each specification is indicated in the column header. Both the dependent and independent variables are standardized to have a variance of 1. Absolute Forecast Error is the absolute difference between each filer's refund amount and their mean elicited belief. All specifications include the listed covariates, plus controls for whether each demographic variable was missing. Tables A1 and A2 present additional descriptive statistics. Robust standard errors are in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Table A6: Belief Updating Rates over Prior Year

	Number of Filers	Fraction with ratio of expected and realized change			Mean Ratio (%)
		< 0	[0,100]	> 100	
	(1)	(2)	(3)	(3)	(4)
All	324	24%	67%	9%	47
Male	96	19%	76%	6%	43
Female	178	21%	70%	9%	58
Below Median Age	151	21%	72%	8%	42
Above Median Age	173	27%	62%	11%	52
Has Kids	119	25%	66%	9%	43
No Kids	205	22%	68%	10%	54
HS or Less	138	22%	70%	8%	35
More than HS	142	21%	70%	9%	62
Received EITC	118	30%	59%	11%	38
No EITC	206	18%	74%	8%	55

Note: Numbers based on the statistic $\frac{m_{1,i}-r_{0,i}}{r_{1,i}-r_{0,i}}$, for tax filers who also filed their taxes at the tax site in the previous year. As described in Section C, this is the difference between an individual's expectation of this year's refund and their prior year refund, scaled by the change in realized refunds from last year to this year. The three middle columns show the fraction of filers, weighted by the size of refund, for whom the ratio is negative, between 0% and 100%, or over 100%. Filers for whom the ratio is negative have expectations that moved in the opposite direction (relative to their prior year refund) than their realized refund. Filers for whom the ratio is between 0% and 100% updated in the "correct" direction, but less than 100%. Filers for whom the ratio is over 100% updated in the "correct" direction, but thought their refund would change more than it did. Mean Ratio (%) is the mean of this statistic across tax filers in each subgroup. The sample size in the first row differs from that in Column 3 of Table 1 because a few individuals had zero change in refund $(r_{1,i} - r_{0,i})$ from the prior year.

Table A7: Updating Rate Heterogeneity

Dependent Variable: Difference between Mean Expectation and Last Year's Refund				
	No Heterogeneity	Full Heterogeneity	Tax Determinants Only	Demographics Only
	(1)	(2)	(3)	(4)
Change in Refund Amount over Last Year	0.597***	0.233	0.264**	0.580***
	(0.0722)	(0.149)	(0.122)	(0.140)
Interacted with Change in Refund Amount				
Absolute Change in AGI		0.0118**	0.0113**	
		(0.00491)	(0.00486)	
Absolute Change in MTR		0.457*	0.476*	
		(0.246)	(0.250)	
Has Dependents		-0.180	-0.0277	
		(0.160)	(0.144)	
Any Change in No. Dependents		0.0127	0.0416	
		(0.144)	(0.140)	
Married		0.0305	-0.0174	
		(0.153)	(0.173)	
Change in Filing Status		0.184	0.194	
		(0.146)	(0.153)	
Received UI during Past Year		-0.316	-0.221	
		(0.225)	(0.222)	
Age 25 or Younger		-0.403		-0.514**
		(0.285)		(0.250)
Above Age 50		0.0207		0.0133
		(0.115)		(0.177)
Any College		0.112		0.0231
		(0.131)		(0.152)
Female		0.205*		0.0469
		(0.119)		(0.158)
Observations	337	337	337	337
R-squared	0.336	0.411	0.395	0.348
Updating model p-values				
No Updating	0.00	0.00	0.00	0.00
No Heterogeneity in Updating Rates	0.00	0.00	0.00	0.31
Full Updating	0.00	0.00	0.00	0.00

Notes: Estimated coefficients from equation 7 in Appendix C. Each control is interacted with the tax filer's change in refund amount. The sample includes all filers for whom tax refund information is available from the prior tax year. Specifications with demographic and economic controls (columns 2-4) also control for missing value indicators for each variable; these coefficients are omitted for brevity. The last three rows present p-values from F-tests of the hypotheses of no updating ($\beta = 0$); no updating rate heterogeneity by filer characteristics; and complete updating ($X_i'\beta = 1\,\forall i$). Robust standard errors are in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table A8: Additional Robustness Checks for Borrowing Results

		Alternate Belief Distribution: Beta Distribution					
	Baseline	Full Sample	No Direct Deposit	No Savings	Can't Change Income	No Dependents	LIML
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Expected Refund Amount	40.38	54.92	10.04	68.22	33.69	116.5	208.5
	(38.07)	(44.14)	(48.67)	(93.92)	(49.81)	(78.39)	(155.2)
Subjective Standard Deviation	-259.3**	-154.0	-48.57	-329.0*	-224.6*	-510.1**	-1300.1
	(131.5)	(120.6)	(116.0)	(193.8)	(135.7)	(206.1)	(924.9)
Controls							
Demographics	X	X	X	X	X	X	X
Tax Determinants	X	X	X	X	X	X	X
Observations	359	359	234	91	211	237	359
R-squared	0.096	0.096	0.103	0.273	0.130	0.107	---

Note: This table investigates the robustness of the borrowing results in Table 3. The regressions include all of the core sample tax filers for whom we have expectations data and credit report data. Column 1 repeats the main specification in Column 4 of Table 3. Columns 2-6 present present results analogous to those in columns 1-4 in Table 4, but assuming beliefs follow beta or triangle distributions rather than normal distributions. The samples in Columns 2-6 are described in Table 4. Column 7 presents LIML estimates for a regression analogous to that in column 1, where we have instrumented for the subjective standard deviation with indicators for our two qualitative measures of uncertainty. The demographic controls include controls for whether a filer is female, over 50, a college graduate, married, or has dependents. The tax determinants include controls for the (absolute value of the) change in AGI, a dummy for any change in the number of dependents, a dummy for a change in filing status, and a dummy for whether the filer received UI this year. Robust standard errors are in parentheses. * $p < .1$ ** $p < 0.05$ *** $p < 0.01$

B For Online Publication: Data and Empirical Setting

This appendix provides more information on the tax filer surveys, as well as information on the context in which we conducted these surveys.

B.1 Expectations Survey

The expectations survey consisted of four questions, printed on the next page. The survey was administered by the financial guides at the tax site.

The first question produces a point estimate of individuals' beliefs. The second question measures individuals' qualitative uncertainty: whether they are "not sure at all," "somewhat sure," or "very sure" that their refund would fall within a $1,000-interval around the number they reported in the first question. The third question was used to measure labor income flexibility.

The fourth question elicits probabilistic beliefs. The number of bins was chosen in coordination with the VITA partner in order to balance the need to run the survey quickly with the desire to obtain richer information on individuals' beliefs. The boundaries of the bins were chosen using data on the distribution of refunds for filers at the site in the previous year, so that roughly an equal number of actual refunds would fall in each bin, with a smaller number in the two tail bins. In our core sample, the middle four of these six bins ultimately covered 24%, 19%, 24%, and 13% of tax filers' actual refunds, while the two tail bins covered 20% of tax filers' actual refunds.

1) If you get a tax refund this year, how much do you think it will be? Please choose an amount:

$___________________

(Financial Guide volunteer: please write $500 above this number, and $500 below this number, in the two blank lines in the question below)

2) How sure are you that your refund will be between $________ and $________? Please circle one:

NOT SURE AT ALL SOMEWHAT SURE VERY SURE

3) Suppose you want to make some extra money by working more hours next week. Do you think you could you get your manager/supervisor to schedule you for more hours?

YES

NO

I AM NOT WORKING RIGHT NOW

I AM NOT PAID HOURLY

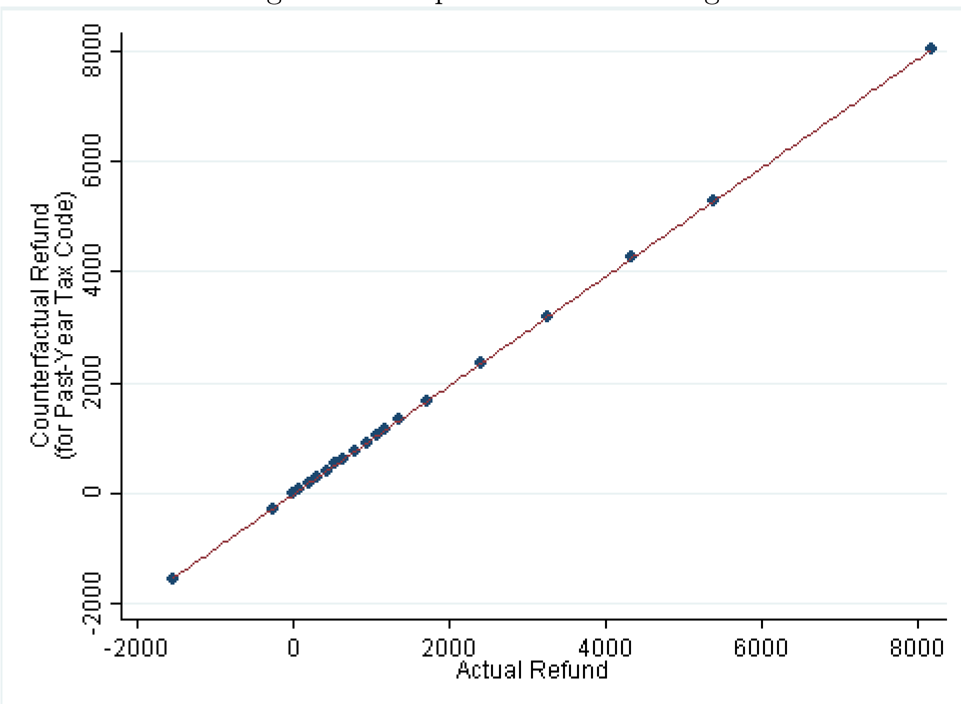
4) We have one final question about your tax refund. Below we show six possible amounts that your refund could be (for example, "between $1000 and $2500"). For each of the six possibilities, please say what is the "percent chance" that you think your refund could be that amount:

Could my refund be…	**(Please Enter % Chance for Each)**
Over $5000	%
Between $2500 and $5000	%
Between $1000 and $2500	%
Between $500 and $1000	%
Between $0 and $500	%
Negative: I will owe taxes	%

B.2 Tax Environment

We conducted our survey in spring 2016, when filers were filing their 2015 tax year returns. Figure A5 shows that there were no major changes in either the federal or state tax schedule that would have resulted in large refund changes between tax years 2014 and 2015.

Figure A5: Imputed Refund Changes



Note: This figure plots a binned scatterplot of the refund an individual would have received under the 2014 tax rules (y-axis), relative to what they received under the 2015 schedule. The 2014 refunds were calculated using NBER TAXSIM (Feenberg and Coutts, 1993).

This is not surprising, because both the federal and state income schedules remained fairly stable between 2014 and 2015. The EITC and CTC also saw no major changes.

C Updating Model

Suppose filers' prior beliefs ($m_{0,i}$) are normally distributed and centered at their prior year refund ($r_{0,i}$) with precision $h_0(X_i)$, and that filers receive noisy signals of the change in their refund, $\Delta r_i + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, 1/h_\epsilon(X_i))$. Filers' Bayesian posterior beliefs ($m_{1,i}$) and "updates" ($m_{1,i} - r_{0,i}$) are then given by:

$$m_{1,i} = r_{0,i} + \underbrace{\frac{h_\epsilon(X_i)}{h_0(X_i) + h_\epsilon(X_i)}}_{\equiv I(X_i)}(\Delta r_i + \epsilon) \tag{6}$$

$$\underbrace{m_{1,i} - r_{0,i}}_{\text{update}} = (r_{1,i} - r_{0,i}) \times \underbrace{I(X_i)}_{X_i'\beta} + \epsilon \times I(X_i) \tag{7}$$

The amount that filers update depends on the relative precision of their prior and signal. In our regressions we parameterize the updating rate $I(X_i) = X_i'\beta$. The primary restriction is that conditional on observables, households update towards their signal at the same rate relative to their prior – in other words, they have the same ratio of their signal and prior precisions. In practice, we view our estimates as capturing an average updating rate among filers in a particular group, averaging over any possible unobserved heterogeneity in updating rates.

Heterogeneity in updating suggests that uncertainty may reflect access to tax-relevant information. We examine heterogeneity in updating by regressing the the filer's "update"– on the realized difference ($r_{1,i} - r_{0,i}$) between this year's refund and the prior year's refund, interacted with economic and demographic characteristics X_i:

$$m_{1,i} - r_{0,i} = (r_{1,i} - r_{0,i})\, X_i'\beta + \eta_i. \tag{8}$$

The updating rates $X_i'\beta$ describe how much tax filers with different characteristics update their beliefs toward the actual refund they receive. Updating rates of 1 would reflect full updating on average, in which average expectations fully reflect how refunds have changed from the prior year. Rates less than 1 reflect partial updating on average, whereas rates greater than 1 would reflect over-reaction to filers' most recent changes in their tax circumstances. Across all specifications we can reject the nulls of no updating (rates of 0) or complete updating (rates of 1).

We reject no heterogeneity in updating rates in column 3, where we focus on heterogeneity by tax determinants, but find no detectable heterogeneity across demographic groups in column 4. In particular, filers who saw larger changes in income (AGI) and larger changes in marginal tax rates (MTRs) update more.[25] This is consistent with filers exerting more effort to reduce uncertainty when the stakes are higher, consistent with models of rational inattention (Coibion et al., 2018). Of course these patterns could also be consistent with other changes (e.g., filing status) just being more difficult for tax filers to understand.

D Computing Compensating Variation

In order to calculate compensating variation for each individual, we have to make assumptions about the interest rate, discount rate, take-home pay, distribution of refund amounts, and form of the utility function.

- **Take-home Pay**: Take-home pay in each period, $c_{0,i}, c_{1,i}$, is equal to the individual's quarterly take home pay (adjusted gross income, minus withholding)/4.
- **Distribution of** y: We use the elicited belief distribution as our measure of $F(y)$.
- **Credit Constraints**: The borrowing limit is $c_{i,1} + E[y]$. A few households choose negative debt (positive savings) given expectations of a negative refund. Given the high levels of baseline non-installment debt in this population, we interpret savings as a marginal repayment of non-installment (e.g., credit card) debt.
- **Consumption Commitments**: Individuals must consume at least $100 each period.
- **Interest Rate**: Individuals can borrow or save at a quarterly interest rate of $R = 1.05$. This is a realistic credit card interest rate for this population.
- **Discount Rate**: Individuals discount the future using $\beta = .98$.

Algorithm

We calculate the compensating variation for each individual. For each functional form for utility, we calculate CV as follows:

- For each s in $1, ..., B$
    1. Draw realizations of the refund amount y_{is} $\{s = 1, ..., S\}$ using the elicited belief distribution $N(\mu_i, \sigma_i^2)$.
    2. Calculate CV_i^{nu}
    3. Calculate CV_i^d assuming $y = E[y]$.
- Save the average realization of CV_i^{nu} and CV_i^d for each individual.

We average over individuals to report the mean CV^{nu} and CV^d for a given utility function and set of preference parameters. These results are presented in Table 5 and Figure 4.

[25] Column 2 of Table A7 indicates that an additional $1,000 change in AGI predicts a 1.2 percentage point increase in $X_i'\beta$. This relationship is statistically and economically significant: the mean absolute AGI change is $6,120, with a standard deviation of $8,780. A similar relationship holds for marginal tax rates.

E For Online Publication: Belief Distributions

E.1 Normal Distributions

Our baseline estimates use beliefs fitted to normal distributions. Our procedure for fitting these beliefs is provided in Section 3 in the main text. As described in that section, our procedure fits reported beliefs to:

$$\min_{\mu,\sigma} \sum_{x \in \mathcal{X}_i} \left[p_{x,i} - \Phi\left(\frac{x-\mu}{\sigma}\right)\right]^2 + \left(\max\{0, 1+\Phi\left(\frac{\underline{x}-\mu}{\sigma}\right) - \Phi\left(\frac{\overline{x}-\mu}{\sigma}\right) - \alpha\}\right)^2 \tag{9}$$

Example For example, suppose a filer reports a "best guess" of $400 and says that there is a 60% chance they will receive between $0 and $500 and a 40% chance they will receive between $500 and $1000. This corresponds to $\mathcal{X} = (\$400, \$500)$, $p = (0.5, 0.6)$, and $(\underline{x}, \overline{x}) = (\$0, \$1000)$. The middle plot in Figure A2 shows the normal distribution which best fits this elicitation. The first and third plots present analogous figures for filers who placed positive probability on three and one bins, respectively. In the single-bin case, equation 1 does not pin down σ, so we restrict the mass outside the bin to equal exactly α.

E.2 Beta Distributions

Fitting beliefs to normal distributions has the advantage of being consistent with the updating model we describe in Appendix Section C. However, normal distributions are also restrictive. For this reason, much of the literature on subjective expectations has fit probabilistic beliefs to beta distributions. Beta distributions have finite support and need not be symmetric or single-peaked. Since probabilistic survey questions can imply a wide range of distributional shapes, beta distributions may do a better job of fitting elicited beliefs.

In order to probe the robustness of our empirical results, we compare our baseline measures of uncertainty to those we would obtain if we fit beliefs to beta distributions.

E.2.1 Fitting Beliefs

As for normal distributions, our procedure for fitting beliefs depends on the number of bins on which the respondent placed positive probability. Using only the elicited probabilities, we can identify the parameters of a beta distribution if there is positive mass on at least three bins.[26] For respondents who placed positive mass on one or two bins, we assume beliefs

[26] A beta distribution has four parameters including its support endpoints, and an elicitation with K bins provides $K+1$ quantiles.

follow triangle distributions. Single-bin reports are fit with a scalene triangle; the support is the full bin, and the mode is the point estimate. The two-bin reports are fit with an isosceles triangle with the widest possible support that is consistent with the probabilities for each bin. These sets of assumptions uniquely pin down a distribution for one- and two-bin responses.

For elicitations with positive mass on three or more bins, we follow Engelberg et al. (2009) in fitting a beta distribution to the reported quantiles. However, we depart from their approach by not restricting the estimated beta densities to be single-peaked. Regardless of the number of bins, we must also make an assumption about the lowest and highest possible refund amounts. The maximum refund amount was a little below \$20,000, and the lowest refund amount was approximately -\$500 (the tax filer had \$500 due). We take these two values as the endpoints of the support of the highest (over \$5,000) and lowest (negative) bins when they contain positive probability.

The triangle distributions are exactly identified and fit using analytical formulas. To fit the beta distributions, we follow Engelberg et al. (2009) and minimize the sum of squared differences between the reported cumulative probabilities at each point in the distribution's support and those of a beta distribution with the same support. Let $\mathcal{X}$ denote the support points of the response to the probabilistic survey question. Let Z denote a beta-distributed random variable governed by parameters (α, β) and normalized to have support on $\mathcal{X}$. Finally, let p_x denote the reported cumulative probability at each point $x \in \mathcal{X}$. We find the $(\hat{\alpha}_i, \hat{\beta}_i)$ for the elicited distribution from each individual i which solves

$$\min_{\alpha,\beta} \sum_{x \in \mathcal{X}_i} [p_{x,i} - P(Z \leq x \mid \alpha, \beta)]^2$$

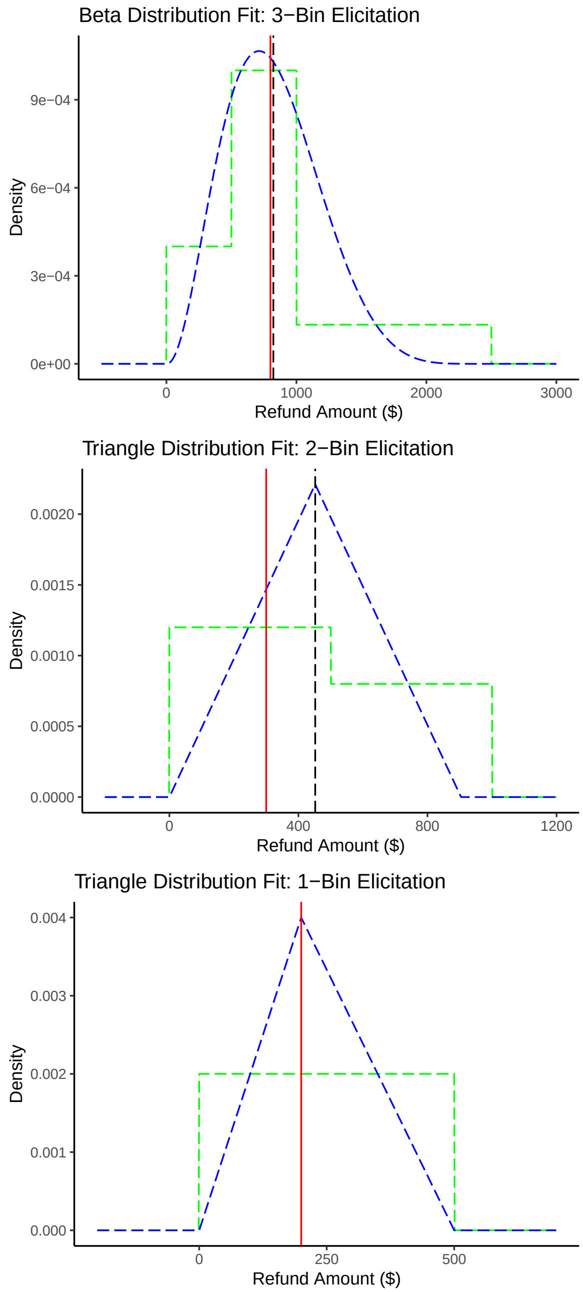
E.2.2 Comparison with Normal Beliefs

Figure A7 compares the means and standard deviations from the normal and beta fitted belief distributions. The first panel shows that the mean beliefs track each other closely; the slope of the fitted regression line lies nearly on top of the 45-degree line. The second panel shows that the standard deviations of uncertainty also track each other closely. However, we obtain larger standard deviations when using the beta distribution. This is especially true for individuals with high absolute levels of uncertainty. This is because the more flexible beta distribution allows us to capture uncertainty that is not "symmetric." By contrast, the normal distribution smooths out uncertainty that leads to skewness in the distribution.

Table A4 presents descriptive statistics on the means and standard deviations of different groups of tax filers under different parametric assumptions. Dropping individuals that put

50/50 probability on two bins does not affect the mean or standard deviation meaningfully (Fischhoff and Bruine De Bruin, 1999). Dropping individuals who placed a hundred percent probability on a single bin reduces the standard deviation somewhat, especially when we use the beta distribution. Our estimates using the normal distribution are less sensitive to the choice of sample.

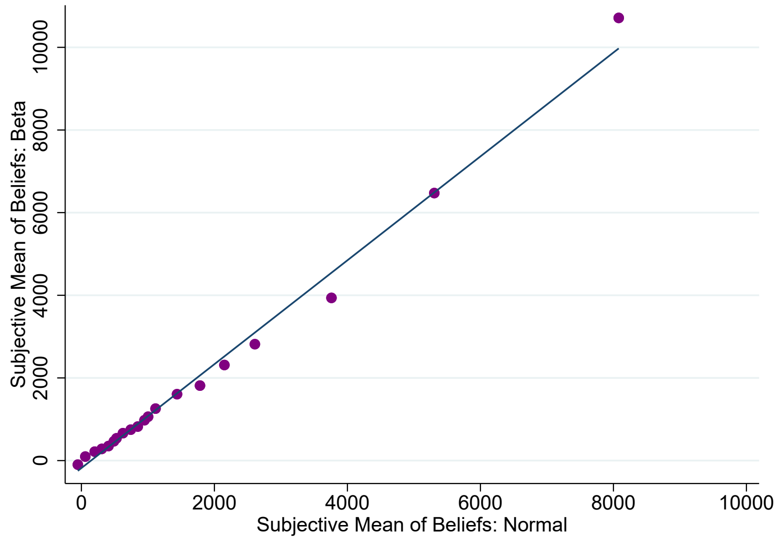
Figure A6: Fitting Beliefs to Beta Distributions



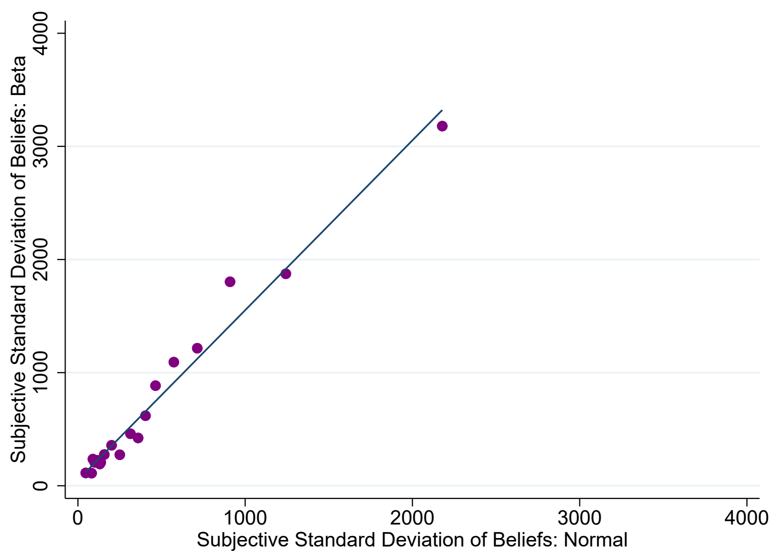
Note: This figure shows how we fit probabilistic beliefs to beta distributions if the individual places positive mass in 3 or more bins (top), in 2 bins (middle) or 1 bin (bottom). Solid lines denote data; dashed lines denote fitted distributions. The green dashed lines report the distribution of beliefs, assigning a uniform density over the density in each bin. The red line denotes the point expectation. The dashed blue curves show the density of the fitted distribution and the dashed black line shows the mean of this distribution. More information on how we fit beliefs to beta distributions is provided in Appendix Section E. Graphs describing how we fit beliefs to normal distributions are provided in Figure A2. Table 2 presents descriptive statistics on the fitted beliefs.

Figure A7: Distributional Assumptions for Beliefs

A. Mean Expectation



B. Subjective Standard Deviation



Note: This figure plots the fitted mean beliefs (Panel A) and fitted standard deviations of beliefs (Panel B) from a normal distribution against those from a beta distribution. Section 3 describes how we fit normal distributions; Appendix Section E describes how we fit beta distributions.