NBER WORKING PAPER SERIES

ON ELICITING SUBJECTIVE PROBABILITY DISTRIBUTIONS OF EXPECTATIONS

Valerie R. Boctor
Olivier Coibion
Yuriy Gorodnichenko
Michael Weber

Working Paper 32406
http://www.nber.org/papers/w32406

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2024

We thank Shannon Hazlett at Nielsen for her assistance with the survey. Results in this article are calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. Information on availability and access to the data is available at http://research.chicagobooth.edu/nielsen. Gorodnichenko thanks BB90 fund for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

On Eliciting Subjective Probability Distributions of Expectations
Valerie R. Boctor, Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber
NBER Working Paper No. 32406
May 2024
JEL No. C83,D84,E31

ABSTRACT

Using data from a large survey of American households, we compare density forecasts elicited with bins- and scenarios-based questions. We show that inflation density forecasts are sensitive to the survey question designs used to elicit them. The within-person discrepancy is smaller, but still discernible, for unemployment expectations. The discrepancy in responses is systematically related to sociodemographic characteristics of respondents. The differences shed light on the significance of priming in bins-based inflation density forecasts.

Valerie R. Boctor
University of California,
Berkeley
valboctor@berkeley.edu

Olivier Coibion
Department of Economics
University of Texas at Austin
2225 Speedway
Austin, TX 78712
and NBER
ocoibion@gmail.com

Yuriy Gorodnichenko
Department of Economics
530 Evans Hall #3880
University of California, Berkeley
Berkeley, CA 94720-3880
and IZA
and also NBER
ygorodni@econ.berkeley.edu

Michael Weber
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
michael.weber@chicagobooth.edu

"*How should we measure inflation expectations, and how should we use that information for forecasting and controlling inflation? I certainly do not have complete answers to those questions, but I believe that they are of great practical importance*." Ben Bernanke (2007).

I. Introduction

Expectations are a core element of modern macroeconomic models and policymaking. As a result, measuring expectations is central for a broad spectrum of applications ranging from understanding the Phillips curve to quantifying uncertainty to managing expectations for macroeconomic stabilization. However, eliciting subjective expectations— especially subjective probability distributions— from surveys of households and firms (i.e., the general public) is fraught with a number of measurement issues, such as limited numeric and economic literacy of respondents. Our objective is to compare subjective expectations elicited via different methods to highlight potential differences in outcomes and help researchers and practitioners choose methods appropriate for their applications. Specifically, we focus on two popular survey designs: *i*) the bins design popularized by the Survey of Consumer Expectations which is run by the Federal Reserve Bank of New York (Potter et al. 2017), and *ii*) the scenarios design proposed by Bloom et al. (2020).[1]

Using a large survey of U.S. households presented with questions based on both survey designs, we find that there are significant differences in measured subjective expectations across the designs. For example, scenario-based inflation expectations tend to convey higher levels and greater uncertainty than bins-based inflation expectations. At the same time, the cross-design differences are smaller for unemployment expectations. We also find that bins-based design may result in lumping responses at extreme bins and may prime respondents to choose unlikely outcomes. For example, few households envision deflation when expectations are elicited via scenario-based questions or via point predictions.[2] In contrast, many households assign positive probability to deflation in bin-based questions that include deflation as a possible outcome. We observe that when households are free to choose possible outcomes (especially for inflation), they tend to report scenarios outside the ranges offered in bins-based questions. This pattern reduces

[1] Because of space constraints in our survey, we did not study min-max-midpoint approach popularized by Guiso, Jappelli and Pistaferri (2002).
[2] Gorodnichenko and Sergeyev (2021) document this pattern holds for many advanced economies, including Japan.

consistency across methods for inflation expectations but the discrepancies are smaller for unemployment expectations.

Our paper contributes to several strands of research. First, Manski (2004, 2017) and others discuss the pros and cons of using different methods to elicit subjective probability distributions and measure uncertainty.[3] We provide a novel within-respondent comparison of two leading methods and thus shed new light on the strengths and weaknesses of these popular methods. Second, a large body of work studies demographic predictors of inflation expectations.[4] Becker et al. (2023) show in online surveys that the average subjective inflation expectations of households is sensitive to the location of the midpoint of the proposed probability distribution and the size of the bins. Our contribution is to document predictors of discrepancies in responses across survey designs.

II. Background and Survey Design

We utilize three methods to elicit subjective expectations. The first method is based on the New York Fed's Survey of Consumer Expectations (SCE). In this influential survey design, respondents are asked to report their subjective probabilities for 10 bins of possible inflation values. The wording of the question is

We would like to ask you some questions about the overall economy and in particular about the rate of inflation/deflation (Note: inflation is the percentage rise in overall prices in the economy, most commonly measured by the Consumer Price Index and deflation corresponds to when prices are falling).

In this question, you will be asked about the probability (PERCENT CHANCE) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, **over the next 12 months**…

	Percentage Chance
the rate of inflation will be 12% or more	______
the rate of inflation will be between 8% and 12%	______
the rate of inflation will be between 4% and 8%	______
the rate of inflation will be between 2% and 4%	______
the rate of inflation will be between 0% and 2%	______

[3] See Bruine de Bruin et al. 2023 for a survey of this literature.
[4] See D'Acunto et al. 2023, Weber et al. 2022, and D'Acunto and Weber (forthcoming) for surveys.

the rate of deflation (opposite of inflation) will be between 0% and 2%	______
the rate of deflation (opposite of inflation) will be between 2% and 4%	______
the rate of deflation (opposite of inflation) will be between 4% and 8%	______
the rate of deflation (opposite of inflation) will be between 8% and 12%	______
the rate of deflation (opposite of inflation) will be 12% or more	______
% Total	______

The main advantage of the bins-based design is that it yields probability distributions for each respondent, as opposed to point forecasts, allowing researchers to infer person-level uncertainty from the probability-weighted dispersion of point values around the implied mean. On the downside, Coibion et al. (2020) document several potential problems related to priming that may arise from offering respondents a fixed grid of possible outcomes. Here and in other designs respondents are asked to forecast a specific price index.

The second method is the scenarios-based design proposed by Bloom et al. (2020). This design asks respondents to report values for low-, medium-, and high-inflation scenarios (in some cases, respondents are asked to provide five scenarios) and assign subjective probabilities to each scenario. The wording of the question is:

Over the next 12 months, which approximate inflation rate (as measured by the Consumer Price Index) would you assign to each of the following scenarios? If you think there was inflation, please enter a positive number. If you think there will be deflation, please enter a negative number. If you think there will be neither inflation nor deflation, please enter zero.

A LOW inflation rate would be about:	______
A MEDIUM inflation rate would be about:	______
A HIGH inflation rate would be about:	______

Please distribute 100 points to the percentage changes you just entered, to indicate how likely you think it is that each inflation rate will happen. The sum of the points you allocate should total to 100.

LOW: The likelihood of realizing a “LOW” inflation rate would be	______
MEDIUM: The likelihood of realizing a “MEDIUM” inflation rate would be	______
HIGH: The likelihood of realizing a ``HIGH’’ inflation rate would be	
% Total [TOTAL ANSWERS FROM ABOVE – MUST SUM TO 100%]	______

Unlike the bins-based method, this method gives respondents more freedom to pick possible outcomes and thus priming or bunching of responses in extreme bins is less likely. On the other hand, because responses are not supervised, one may obtain a sample with outliers and bunching at multiples of 5. Appendix Figure A1 shows how we convert the three-point responses into probability distributions.

Because bins- and scenario-based questions are cognitively demanding, we also ask respondents to provide their point predictions. The wording of the question is

What do you think the inflation rate (as measured by the Consumer Price Index) is going to be over the next 12 months? Please provide an answer as a percentage change from current prices.

% [RANGE: -100-100, ONE DECIMAL] _____

If you think there will be inflation, please enter a positive number. If you think there will be deflation, please enter a negative number. If you think there will be neither inflation nor deflation, please enter zero.

Similar to the scenario-based method, this question is less likely to prime responses by offering a fixed grid of possible outcomes but is more likely to generate outlier responses. However, we note two important features of this question. First, this wording of the question prompts respondents to contemplate deflation. Second, although this question mimics the Michigan Survey of Consumers (MSC), we do not probe respondents who report high rates of inflation because we want to minimize priming.[5]

We use the same three methods to elicit expectations for unemployment rate. The wording for the bins-based question is

In THIS question, you will be asked about the probability (PERCENT CHANCE) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, **in 12 months** …

Percentage Chance

Unemployment rate will be more than 20% ______

[5] The Michigan Survey of Consumers provides this instruction to interviewers, "IF R GIVES AN ANSWER THAT IS GREATER THAN 5%, PLEASE PROBE WITH: 'Let me make sure I have that correct. You said that you expect prices to go (up/down) during the next 12 months by (X) percent. Is that correct?'"

Unemployment rate will be between 15% and 20%	______
Unemployment rate will be between 10% and 15%	______
Unemployment rate will be between 8% and 10%	______
Unemployment rate will be between 6% and 8%	______
Unemployment rate will be between 4% and 6%	______
Unemployment rate will be less than 4%	______
% Total [TOTAL ANSWERS FROM ABOVE – MUST SUM TO 100%]	______

III. Data

In our empirical analysis, we use the November 2020 wave of the survey that was launched in Coibion, Gorodnichenko and Weber (2022). This survey builds on the Nielsen Homescan Panel which is a popular platform for marketing research. The panel tracks more than 80,000 households who are broadly representative of the U.S. population (sampling weights are provided by Nielsen to correct for any imbalances). The panelists are invited to participate in occasional surveys which typically have a response rate of 20-25%. Participation is rewarded with points which panelists can cash in with Nielsen. Some information about households (e.g., household size and income) is available via Nielsen's background annual surveys and additional information (e.g., current employment status, political leanings) is collected by our survey.

The November 2020 wave targeted electoral issues and the surrounding uncertainty. Given the focus of our analysis, we utilize only questions pertinent to subjective probability distributions of macroeconomic variables: inflation and unemployment. Furthermore, because that survey wave had a randomized controlled trial about elections, we constrain the sample to respondents who were not provided with information in the information treatments (i.e., the control group). We also apply a series of filters to remove noise from survey responses. Specifically, we drop responses that include extreme responses (the point prediction for inflation is greater than 30% or less than -1%, the point prediction for unemployment is greater than 30%) and responses for which a parametric distribution fitting is infeasible, based on the method of Engelberg, Manski and Williams (2009).[6]

[6] We also drop individuals who, in the bins question, report only two, non-adjacent bins, since parametric distribution fitting is not straightforward in those cases. See Appendix B for a detailed description of the parametric approach and robustness.

To provide points of reference, the actual rates for inflation and unemployment at the time of the survey were 1.2% and 6.7%. The Survey of Professional Forecasters, which is run by the Federal Reserve Bank of Philadelphia, projected inflation and unemployment to reach 2.2% and 5.8%. In November 2021, inflation and unemployment rates realized at 6.9% and 4.1%.

IV. Basic moments

In a first pass at the data, we report basic descriptive statistics (raw as well as Huber-robust to outliers) in Table 1. Generally, the distributions of inflation and unemployment expectations have a thick right tail (columns (5)-(9) provide selected percentiles) so that average forecasts (column (1)) exceed medians. In the same spirit, Huber-robust moments (columns (3) and (4)) tend to be moderated relative to moments of the raw data. For example, the average point forecast for inflation is 6.3% in the raw data and 3.6% according to the Huber-robust method. As suggested by Reis (2021), the heavy right tail in inflation expectations in November 2020 turned out to be consistent with high realized inflation in November 2021.

The implied means of inflation expectations from bins-based questions tend to produce lower values than point predictions which in turn tend to be lower than moments implied from scenario-based questions. For instance, Huber-robust averages are 2.5%, 3.6% and 4.1%, respectively. At the same time the ranking is reversed for unemployment expectations.

Irrespective of whether we use point predictions or implied means, there is much disagreement in households' predictions for inflation and unemployment. We find that the cross-sectional standard deviation was about 3.5% for inflation and about 5% for unemployment. At the same time, the average uncertainty (measured by the standard deviation of the provided subjective probability distributions) tends to be lower than disagreement. Although prior work established this pattern for households' and firms' inflation expectations (e.g., Coibion et al. 2021), we are, to the best of our knowledge, the first to document it for unemployment expectations. We also find that the implied uncertainty is smaller in scenario-based distributions than in bins-based distributions. For example, the standard deviations implied by subjective probability distributions for inflation are 4.0% for bins and 1.4% for scenarios (Column 3 of Table 1). The difference is smaller for uncertainty in unemployment expectations, but it is still sizable.

To validate our data, we compare our moments to their counterparts in the SCE and MSC. Broadly, the results are similar across surveys but there are some differences. For instance, we tend to find more disagreement and uncertainty in the Nielsen sample than in the SCE and MSC. Some of this difference should be due to variation in survey designs (e.g., recall that MSC probes responses if reported inflation expectations exceed 5% which could compress the distribution. In our question eliciting point predictions, we ask respondents to report inflation within a value in the range of (-100, 100) *if* they report a value outside this range).

Table 2 provides additional details on the responses to questions eliciting probability distributions for inflation and unemployment. Column (1) shows that roughly 20% of respondents assign positive probability to only one bin and approximately 30% assign positively probabilities to *all* bins. The latter can be surprising given that some bins offer rather extreme scenarios such as deflation of more than 12% (this happened during the recession of 1921) or an unemployment rate greater than 20% (this happened only during the Great Depression). These patterns are consistent with priming of responses (e.g., respondents may feel the need to assign a positive probability to a scenario just because it is offered) and the cognitively demanding nature of the question (e.g., many households can have low financial/numeric literacy and cognition and thus struggle with probabilities, see Lusardi and Mitchell, 2023; D'Acunto et al, 2023).

The scenario-based question also has limitations. Ideally, respondents should have reported three distinct scenarios, each with probability $0< \Pr\left[\pi^{e}_{i,scen}\right] < 100$. However, we observe repeat scenarios relatively frequently in the data, as well as cases in which one scenario's probability is selected as 100, or the sum of two scenarios' probabilities is 100. As a consequence, we observe that 26% of respondents effectively reported a single value for expected inflation with $\Pr\left[\pi^{e}_{i,scen}\right] = 100$ (column 2). In a similar manner, 16% of individuals effectively report two scenarios for inflation so that only ~60% of respondents effectively provided three scenarios. This issue is less pronounced for unemployment expectations for which more than 80% of respondents provide three scenarios with positive probabilities. Again, these results suggest that low financial and numeric literacy may be a binding constraint and that inflation is a particularly confusing subject for households.[7] However, the potentially unusual choices for bins and scenario are

[7] Bernanke et al. (1999) observed, "Some economists have argued that the public's consistent apathy towards inflation (as evidenced by opinion polls, for example) is primarily the result of confusion about what inflation really is…

somewhat consistent: those who assign positive probability for a single bin are also more likely to assign positive probability to fewer scenarios.

V. Consistency of subjective expectations

Although different methods to elicit subjective probability distributions yield broadly similar averages and standard deviations, there is dramatic variation across methods in individual responses. Figure 1 presents typical cases for inflation and unemployment expectations. We quantify these differences using several metrics.

A. Endpoint Matching

The first measure is a check of whether the implied distribution supports match exactly. Here, consistency is measured as an indicator variable equal to 1 if the responses from each question imply the same support of possible inflation values, after adjusting scenario values to the nearest corresponding bin value. In particular, for respondent i, let LB_v^i and UB_v^i denote the extreme values from version $v = \{bins, scen\}$, respectively. Each bin (or scenario) b_n has lower bound b_n^L and upper bound b_n^R. Then the support for the reported distribution is given by

$$LB_{bins}^i = \min\{ b_n^L \mid \Pr[\pi_i^e \in b_n] > 0\}$$

$$UB_{bins}^i = \max\{b_n^R \mid \Pr[\pi_i^e \in b_n] > 0\}$$

In principle, consistency implies $LB_{bins}^i = LB_{scen}^i$ and $UB_{bins}^i = UB_{scen}^i$, but this comparison is not immediately feasible since the support of the bins- and scenario- based extrema are different. Indeed, the bins effectively run -14% to +14%, whereas scenario range from -100% to 100%. To remedy this, we adjust $(LB_{scen}^i, UB_{scen}^i)$ to the nearest encompassing endpoints from the set of possible bin-based extrema, B'. The adjusted values are denoted $\widetilde{LB}_{scen}^i$ and $\widetilde{UB}_{scen}^i$. In

Somewhat paradoxically, to a degree inflation has become perceived as a serious economic problem precisely because of the public's confusion over what inflation is and about how to make adjustments for it."

general, each scenario-based value corresponds to exactly one bin, although an exception occurs for certain responses that include only one distinct scenario.[8]

Columns (4)-(6) in Table 2 report results for this consistency check. For inflation expectations, we observe that there is low consistency in the lower bound (approximately 14%). This result obtains because respondents tend to assign positive probability to deflation in bins question but almost never envision deflation in scenario-based questions. The consistency rate is higher for the upper bound (approximately 36%) mainly because many respondents assign positively probability to the top inflation bin and thus nest high inflation scenarios. The consistency for both margins is rare (less than 5%). For unemployment expectations, the results are more similar for upper and lower bounds but the rate of consistency is very low too.

B. Distribution Support Overlap

To provide a sense of the intensive margin for consistency, we consider how often the implied distribution supports from the bins and scenarios questions overlap. For each respondent, we calculate the percent overlap by summing up the number of values that are reported as possible in both versions, then dividing by the total sum of the support ranges in each question version.[9] Specifically, we define the share of overlapping values as

$$O_i = \frac{\min\left\{UB^i_{bins}, \widetilde{UB}^i_{scen}\right\} - \max\left\{LB^i_{bins}, \widetilde{LB}^i_{scen}\right\}}{\frac{1}{2}\left[\widetilde{UB}^i_{scen} - \widetilde{LB}^i_{scen}\right] + \frac{1}{2}\left[UB^i_{bins} - LB^i_{bins}\right]}$$

where the numerator is the minimum upper bound for scenario- and bin-based responses minus the upper lower bound and the denominator is the average range for both types of questions.

We find (column 7 in Table 2) that the average percent overlap between the values reported in the bins and data is 35% for inflation expectations and 59% for unemployment expectations. We leave it to the reader to decide whether this is a half-full or half-empty glass but the method of eliciting subjective expectations is potentially important.

[8] For special cases where single-value responses lie on an endpoint from the overlapping b_ns, the transformation is 1-to-2. E.g., example, suppose $\pi_i^e = 4$. Then there are two corresponding bins that would be deemed "consistent'": either $(LB_1^i, UB_1^i) = (2,4)$ or $(LB_1^i, UB_1^i) = (4,8)$. Both of these are permitted in the consistency check.

[9] This metric assumes that intervals are continuous for both versions. For the bins-based data, this is oversimplifying in cases where respondents report positive probability for bins that are not consecutive. Thus, the overlap rates calculated with our formula can be considered as an upper bound.

C. Point forecasts vs. Implied means and uncertainty

Although we do not observe "the true subjective expectations" and thus we cannot have a clear benchmark for validating responses in the distributional questions, one can use point predictions as a benchmark because point predictions are less cognitively demanding and the question design generally has less priming. Panel A of Figure 2 presents binscatters for implied means vs point predictions of inflation expectations. There is a fairly weak relationship between point predictions and bins-based implied means: regressing implied means on point prediction yields an estimated slope of 0.11 (standard error 0.01) and a $R^2 = 0.05$. Clearly, implied means level off for high point predictions. This pattern is consistent with a cap on the maximum expected inflation that respondents can convey (recall that the top bin is inflation of 12% and above which we code as 16%). This pattern is also consistent with priming of responses in that respondents are nudged to consider more moderate outcomes for inflation based on the bins that they see. Because scenario-based questions do not have fixed bins and are less likely to suffer from priming, one should expect a stronger relationship between scenario-based implied mean and point prediction. This prediction is borne out by the data: The slope is now 0.61 (standard error 0.01), still less than 1 but substantially larger than for the subjective distribution with a $R^2$of 0.61. Furthermore, the relationship between implied means for bins-based distributions is stronger for the unemployment rate (Panel B) which is consistent with wider and a less binding set of bins for unemployment. In other words, one may expect more consistency of bins-based implied means if bins cover a wider range of possible outcomes rather than limit them to be between -16% to +16% inflation.

In a related exercise, we examine how uncertainty is related to point predictions. Intuitively, higher inflation is associated with more volatile inflation and thus one may expect a positive relationship between point predictions and uncertainty. Panels C shows that bins-based uncertainty is systematically above scenario-based uncertainty for inflation expectations. Furthermore, there is U-shaped relationship between bins-based uncertainty and point predictions when point predictions are close zero. We conjecture that the spike in uncertainty for low point predictions comes from respondents confusing the concepts of deflation and inflation in the bins-based question. We observed a similar pattern for unemployment expectations (Panel D).

D. Support and uncertainty

Because individual probability distributions are noisy, it is instructive to examine average (across respondents) CDFs for expectations (Panels G and H of Figure 2). For the bins-based CDF, the cumulative probability is set to 0 at $\pi^e = -14$, and it is set to 100 at $\pi^e = 14$.[10] For the bins-based CDFs, this censoring is required at some chosen values, since we do not observe extreme values within the highest and lowest bins. For the scenarios-based CDFs, the cumulative probability is set to 100 at the maximum support value, which we choose to be 50 in order to smooth the right side of the distribution via linear interpolation.

The CDFs corresponding to the bins- and the scenarios-based inflation expectations exhibit a familiar "S" shape with an inflection point around zero but there are important differences.[11] First, the distribution implied by the scenarios-based question lies well below the bins CDF, with roughly 40% of the probability mass in the scenarios CDF corresponding to inflation values above the cutoff midpoint value for the bins CDF, $\pi^e = 14$. In other words, the reported inflation expectations in the scenarios CDF are substantially higher than those implied by the traditional bins-based data both at the aggregate and individual levels. A similar finding applies to unemployment expectations. Second, the SCE's maximal bin, which spans [12, UB], potentially obscures a large portion of what could be the "true" aggregate distribution, given the large average probability of inflation above 12% in the scenarios-based CDF. Third, the left tail of the bins-based distribution lies well above the scenarios CDF. Specifically, the bins CDF suggests that households believe deflation will occur with a probability of up to 33.5%, in contrast with a probability of 1.3% for the same inflation range in the scenarios-based CDF. Perhaps not surprisingly, implied uncertainty is correlated across bins-based and scenario-based questions but the relationship is not linear (Panels E and F).

E. Predictors of discrepancies

What respondent characteristics predict differences across methods eliciting expectations? To answer this question, we regress the absolute value of differences across different measures on

[10] The conventionally assumed extrema for the bins-based distribution support are {-16, 16}. For the bins-based, CDFs, we plot the implied midpoints of the bins, so {-14, 14} are the effective cut-off values.

[11] The CDFs for our bins-based question and the one implied by the SCE data (for November 2020) are similar, which lends credence to the idea that our data is comparable to the SCE, and that our results apply more generally.

sociodemographic variables $\boldsymbol{X}$ such as gender, age, educational attainment, income, political leanings, and employment status:

$$\left|Expectation_i^{measure\ \#1} - Expectation_i^{measure\ \#2}\right| = \boldsymbol{X}_i\boldsymbol{\beta} + error$$

The choice of these variables is informed by earlier research documenting that these characteristics can predict cross-sectional variation in macroeconomic expectations (see D'Acunto et al. 2023 for a survey). For example, women usually have higher inflation expectations, a fact that we reproduce as well (see Appendix Table A1 for regression estimates). We find (Table 3) that some of these variables can predict discrepancies in responses, too. For example, female respondents tend to have large differences not only for forecasts (columns 1-3 and 4-7, for inflation and unemployment, respectively) but also for uncertainty in their forecasts (columns 4 and 8). Higher incomes and college+ education are associated with smaller discrepancies. Other variables can have some predictive power too, but these associations are less robust. Higher incomes and education are likely associated with stronger cognitive abilities and thus weaker inconsistencies in responses (D'Acunto et al, 2023), yet it is not clear why women would have more dissonance in their responses across different types of survey questions.

VI. Concluding remarks

Measuring macroeconomic expectations of households and firms is a difficult task. Time constraints, limited financial and numeric literacy, lack of economic knowledge, present formidable challenges. At the same time, returns to good measurement are very high for positive and normative economics. To this end, we conduct a systematic comparison of two popular methods (bins- vs scenario-based questions) to elicit subjective probabilistic distributions for inflation and unemployment expectations.

We find that elicited subjective expectations are sensitive to which method is used. There are important differences in the first and second moments as well as the support of the elicited distributions, there is limited correlation in responses across methods. Furthermore, these differences appear to vary systematically across respondent characteristics thus indicating that these differences are more than noise in the data. Although we do not have true subjective expectations to benchmark these two methods, our interpretation of the results suggests that scenario-based elicitation could be a better approach because it is less prone to priming and censoring of responses. Furthermore, because this analysis was done prior to the inflation surge 2021-22, it likely understates how large differences in question formulations may be over time, since the bins questions are generally not altered when inflation rates spike and the associated priming effects become more severe. We hope that our analysis will spur more interest and work in this arena.

References

Becker, Christoph, Peter Duersch, and Thomas Eife. 2023. "Measuring Inflation Expectations: How the Response Scale Shapes Density Forecasts." manuscript.

Bernanke, Ben. 2007. "Inflation Expectations and Inflation Forecasting." Board of Governors of the Federal Reserve System. July 10, 2007. https://www.federalreserve.gov/newsevents/speech/bernanke20070710a.htm.

Bloom, Nicholas, Steven J. Davis, Lucia Foster, Brian Lucking, Scott Ohlmacher, and Itay Saporta-Eksten. 2020. "Business-Level Expectations and Uncertainty." SSRN Scholarly Paper. Rochester, NY. https://papers.ssrn.com/abstract=3808466.Ben Bernanke (2007).

Bruine de Bruin, Wändi, Alycia Chin, Jeff Dominitz, Wilbert van der Klaauw, 2023. "Chapter 1 - Household surveys and probabilistic questions," in R. Bachmann, G. Topa, W. van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, pp. 3-31

Coibion, Olivier Yuriy Gorodnichenko, Saten Kumar, and Jane Ryngaert, 2021. "Do You Know that I Know that You Know…? Higher-Order Beliefs in Survey Data," *Quarterly Journal of Economics* 136(3): 1387-1446.

Coibion, Olivier, Yuriy Gorodnichenko, Saten Kumar, and Mathieu Pedemonte, 2020. "Inflation expectations as a policy tool?," *Journal of International Economics* 124(C): 103297

D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber, 2023. "Chapter 5 - What do the data tell us about inflation expectations?" in R. Bachmann, G. Topa, W. van der Klaauw, eds., *Handbook of Economic Expectations*, Academic Press, pp. 133-161.

D'Acunto, Francesco and Michael Weber, forthcoming. "Why Survey-Based Subjective Expectations are Meaningful and Important," *Annual Review of Economics*.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber, 2023. "IQ, Expectations, and Choice," *Review of Economic Studies*, 90(5): 2292-2325.

Engelberg, Joseph, Charles F. Manski, and Jared Williams. 2009. "Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters." *Journal of Business and Economic Statistics* 27 (1): 30–41.

Gorodnichenko, Yuriy, and Dmitriy Sergeyev, 2021. “Zero Lower Bound on Inflation Expectations,” NBER Working Paper 29496.

Guiso, Luigi, Tullio Jappelli, and Luigi Pistaferri. 2002. “An Empirical Analysis of Earnings and Employment Risk,” *Journal of Business and Economic Statistics* 20: 241–253.

Lusardi, Annamaria, and Olivia S. Mitchell. 2023. “The Importance of Financial Literacy: Opening a New Field.” *Journal of Economic Perspectives* 37 (4): 137-54.

Manski, Charles, 2004. “Measuring Expectations,” *Econometrica* 72(5): 1329-1376.

Manski, Charles, 2017. Survey Measurement of Probabilistic Macroeconomic Expectations: Progress and Promise, Charles F. Manski. in *NBER Macroeconomics Annual* 2017, volume 32, Eichenbaum and Parker, eds.

Potter, Simon, Marco Del Negro, Giorgio Topa, and Wilbert van der Klaauw. 2017. “The Advantages of Probabilistic Survey Questions.” Available at https://papers.ssrn.com/abstract=3098648.

Reis, Ricardo, 2021. “Losing the Inflation Anchors,” *Brookings Papers on Economic Activity* 52(2 (Fall)), pages 307-379.

Weber, Michael, Francesco D'Acunto, Yuriy Gorodnichenko, and Olivier Coibion. 2022. “The Subjective Inflation Expectations of Households and Firms: Measurement, Determinants, and Implications.” *Journal of Economic Perspectives* 36(3): 157-84.

Table 1. Moments of 12 months ahead inflation and unemployment expectations in November 2020

	Raw		Huber robust		Percentiles				
	Mean	St. Dev	Mean	St. Dev	P10	P25	P50	P75	P90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A. Inflation expectations.									
Nielsen									
Point Forecast	6.3	7.1	3.6	3.5	0.0	1.3	4.0	10.0	18.0
Implied means									
Bins	2.8	5.0	2.5	3.3	-2.2	0.0	2.2	5.8	10.0
Scenarios	11.6	17.1	4.1	3.8	0.0	2.1	4.8	13.5	34.0
Implied st.dev. (uncertainty)									
Bins	4.2	3.5	4.0	3.3	0.0	1.0	3.6	7.7	9.0
Scenarios	3.5	5.4	1.4	1.5	0.0	0.0	1.5	3.9	10.8
Survey of Consumer Expectations									
Point Forecast	6.6	6.6	3.5	2.1	1.9	2.0	4.0	10.0	20.0
Implied mean, bins	4.3	3.9	3.5	2.5	0.2	1.6	3.2	6.0	10.1
Implied st.dev. (uncertainty), bins	2.9	2.5	2.2	1.5	0.0	1.1	2.2	3.9	7.2
Michigan Survey of Consumers									
Point Forecast	4.7	3.3	3.7	2.1	1.0	2.0	4.0	5.0	10.0
Panel B. Unemployment expectations.									
Nielsen									
Point Forecast	11.3	7.4	9.4	5.0	4.0	6.0	10.0	15.0	24.0
Implied means									
Bins	11.2	4.9	10.5	3.8	5.2	7.6	10.5	14.1	17.7
Scenarios	14.9	15.4	8.8	5.2	3.5	6.0	9.3	17.5	35.0
Implied st.dev. (uncertainty)									
Bins	3.3	2.5	3.3	2.4	0.0	1.0	3.3	5.6	6.7
Scenarios	4.3	5.6	2.2	1.9	0.0	1.1	2.4	5.1	11.4

Notes: The sample in panel A is restricted to respondents with point predictions between -1% and 30%. The sample in panel B is restricted to respondent with point predictions between 0% and 30%. Implied uncertainty variables are based on 1% winsorized variances.

Table 2. Consistency in bin- and scenario-based questions

Number of bins with positive probability	Share of responses, %	Distribution by: Scenarios with positive probability, %			Distribution by: Share with consistent bounds, %			Average Overlap, %
		One	Two	Three	Lower	Upper	Both	
	(1)	(2)	(3)	(4)	(4)	(5)	(6)	(7)
Panel A: Inflation expectations								
1 Bin	20.1	41.7	15.4	42.9	25.8	22.9	9.2	33.8
2	10.0	20.0	20.0	59.9	32.9	28.3	13.9	51.7
3	8.3	15.0	18.5	66.5	23.7	35.9	10.8	58.8
4	8.0	18.7	14.8	66.6	17.2	34.9	3.2	53.4
5	9.7	18.0	14.2	67.8	20.5	35.5	2.4	55.5
6	3.4	19.3	13.6	67.1	0.9	31.7	0.0	45.3
7	3.1	17.6	15.1	67.2	0.1	38.2	0.1	36.8
8	3.1	19.2	15.8	65.0	0.0	44.6	0.0	29.7
9	4.8	25.5	10.7	63.7	0.2	41.2	0.2	20.0
10 Bins	29.3	26.8	15.4	57.8	0.5	47.7	0.1	11.8
All observations	100.0	25.9	15.7	58.4	14.0	36.3	4.6	34.6
Panel B: Unemployment expectations								
1 Bin	17.8	0.7	24.2	75.2	10.9	12.4	7.0	35.3
2	12.4	0.6	21.7	77.7	15.0	22.0	10.0	61.0
3	11.1	0.3	18.5	81.2	19.8	17.4	14.5	67.9
4	10.6	0.9	13.9	85.2	11.0	12.5	5.0	63.9
5	9.1	0.4	13.4	86.2	9.5	8.6	2.0	65.1
6	9.0	0.2	12.5	87.3	7.4	6.2	5.0	63.1
7 Bins	30.0	0.1	12.8	87.2	0.7	5.0	0.9	59.0
All observations	100.0	0.4	16.8	82.8	8.9	11.0	5.5	57.5

Notes: Column (1) show the share of respondents assigning positive probability to a given number of bins. Columns (2)-(4) show the share of responses with a given number of distinct scenarios. For each row, columns (2)-(4) sum up to 100. Columns (4)-(6) show the share of respondents who report consistent bounds as described in Section V.A. Column (7) report the percent overlap in supports of the reported subjective bins-based and scenario-based distributions. See Section V.B for the method to compute the overlap.

Table 3. Predictors of differences in expectations.

	Inflation				Unemployment			
	Level			Std.	Level			Std.
Measure 1	Point	Point	Bins	Bins	Point	Point	Bins	Bins
Measure 2	Bins	Scenarios	Scenarios	Scenario	Bins	Scenarios	Scenarios	Scenarios
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	1.06***	0.70***	1.29***	0.72***	0.51***	0.86***	0.97***	0.50***
	(0.07)	(0.06)	(0.10)	(0.06)	(0.05)	(0.08)	(0.08)	(0.05)
Age	0.07***	0.03**	0.05**	-0.05***	-0.03**	-0.02	-0.00	-0.01
	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
$Age^2/100$	-0.05***	-0.03**	-0.05**	0.03**	0.02**	0.02	-0.01	0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)
Log Income	-0.70***	-0.32***	-1.00***	-1.19***	-0.86***	-0.82***	-1.34***	-0.88***
	(0.10)	(0.08)	(0.14)	(0.08)	(0.07)	(0.11)	(0.12)	(0.06)
Republican	0.08	0.06	0.06	0.08	0.02	0.06	0.10	0.04
	(0.08)	(0.07)	(0.11)	(0.07)	(0.06)	(0.09)	(0.09)	(0.05)
Green Party	0.39	-0.44	0.26	-1.04**	0.03	-0.57	0.79	0.13
	(0.47)	(0.34)	(0.75)	(0.40)	(0.42)	(0.62)	(0.63)	(0.35)
Libertarian Party	0.05	0.10	0.01	-0.16	-0.12	-0.78***	-0.50*	-0.29**
	(0.26)	(0.21)	(0.33)	(0.19)	(0.18)	(0.23)	(0.26)	(0.15)
Other Party	0.08	0.02	0.21*	0.21***	0.12*	0.06	0.09	0.01
	(0.09)	(0.08)	(0.12)	(0.08)	(0.07)	(0.10)	(0.10)	(0.06)
Party not reported	0.36***	0.08	0.83***	0.87***	0.98***	0.64***	1.21***	0.48***
	(0.13)	(0.11)	(0.18)	(0.11)	(0.10)	(0.16)	(0.17)	(0.09)
Some high school	-0.81***	0.16	0.42	0.79***	0.11	-0.09	0.77*	0.87***
	(0.28)	(0.28)	(0.46)	(0.29)	(0.31)	(0.41)	(0.44)	(0.21)
Graduated high school	0.03	-0.18**	0.09	0.30***	0.12	-0.31***	-0.16	0.09
	(0.10)	(0.09)	(0.14)	(0.09)	(0.07)	(0.12)	(0.12)	(0.07)
Some college	0.00	-0.02	0.17	0.09	-0.04	-0.09	0.07	-0.03
	(0.09)	(0.07)	(0.12)	(0.07)	(0.06)	(0.10)	(0.10)	(0.05)
Post college graduate	-0.41***	-0.39***	-0.68***	-0.42***	-0.11	-0.39***	-0.30***	-0.16***
	(0.10)	(0.08)	(0.12)	(0.08)	(0.07)	(0.11)	(0.11)	(0.06)
Under 30 hours of work	-0.20*	0.26**	-0.06	0.06	0.07	-0.10	0.02	0.13*
	(0.12)	(0.11)	(0.16)	(0.10)	(0.09)	(0.13)	(0.14)	(0.08)
30-34 hours of work	0.11	0.27**	0.22	0.41***	-0.03	0.50**	0.01	0.18*
	(0.16)	(0.13)	(0.23)	(0.14)	(0.12)	(0.21)	(0.19)	(0.11)
Not employed for pay	-0.15*	0.12	-0.09	0.16**	-0.17***	-0.15	-0.27***	-0.08
	(0.09)	(0.08)	(0.12)	(0.08)	(0.06)	(0.10)	(0.10)	(0.06)
Observations	7,361	6,651	6,891	7,620	6,867	6,502	6,511	6,875
R-squared	0.04	0.03	0.05	0.13	0.08	0.04	0.08	0.09

Notes: The table reports estimates for regressions of absolute value of difference in expectations between two methods (measure #1 and Measure #2) on sociodemographic characteristics of responses. All specifications are estimated using Huber robust regression. ***,**,* denote statistical significance at 1, 5, and 10 percent levels.

Figure 1. Distribution of subjection probabilistic distributions for selected respondents.

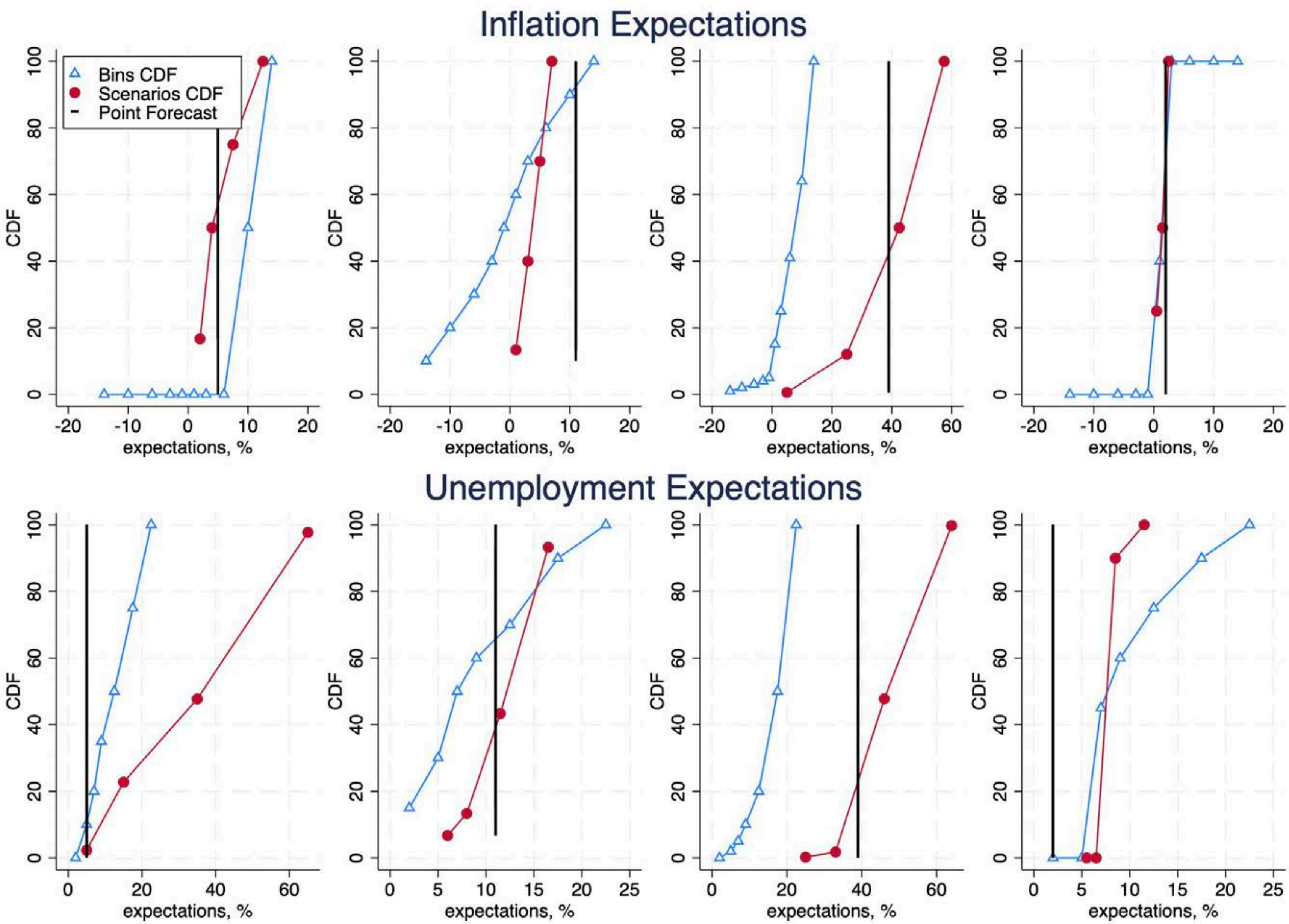
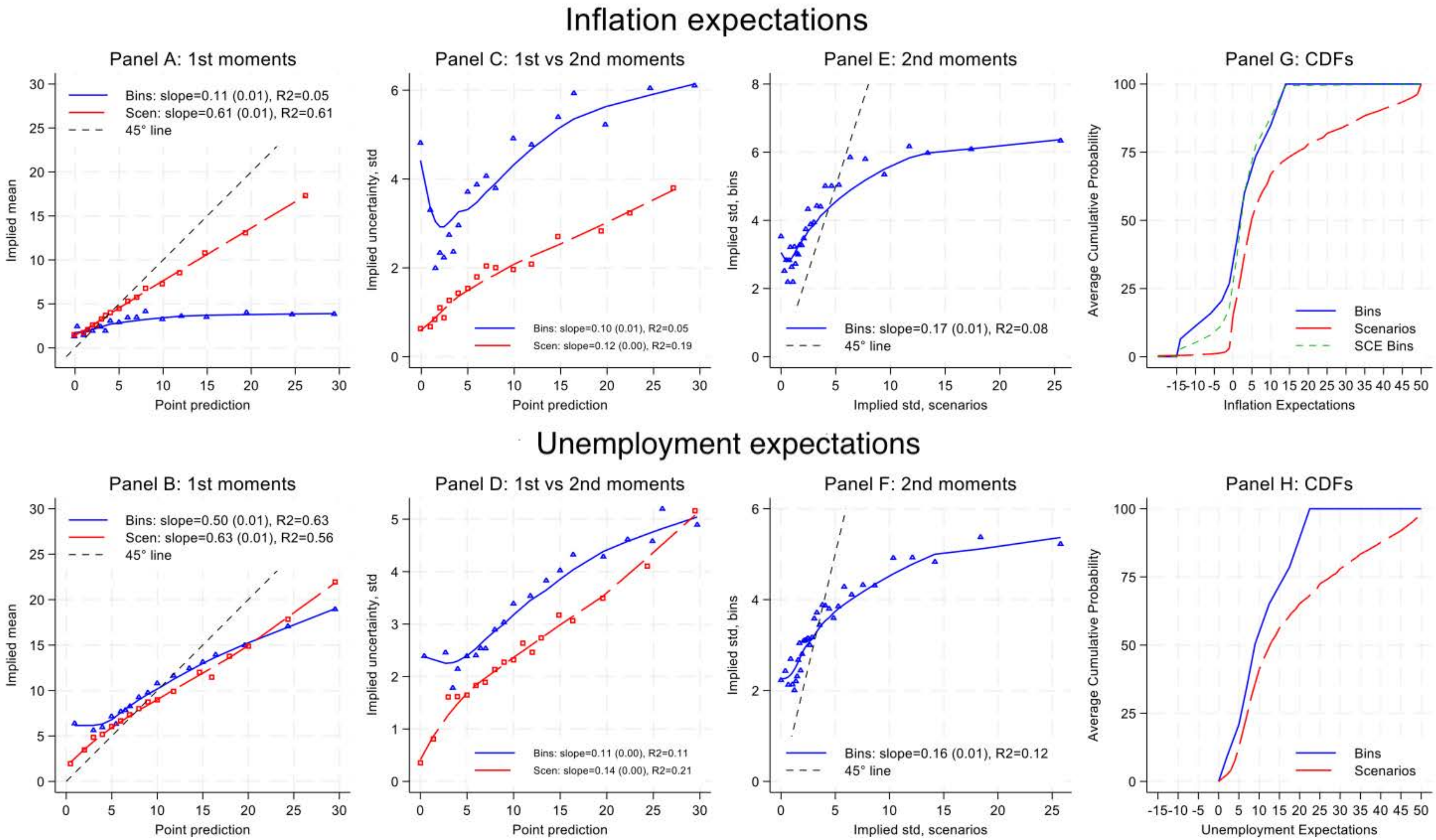


Figure 2. Comparison of moments across methods.

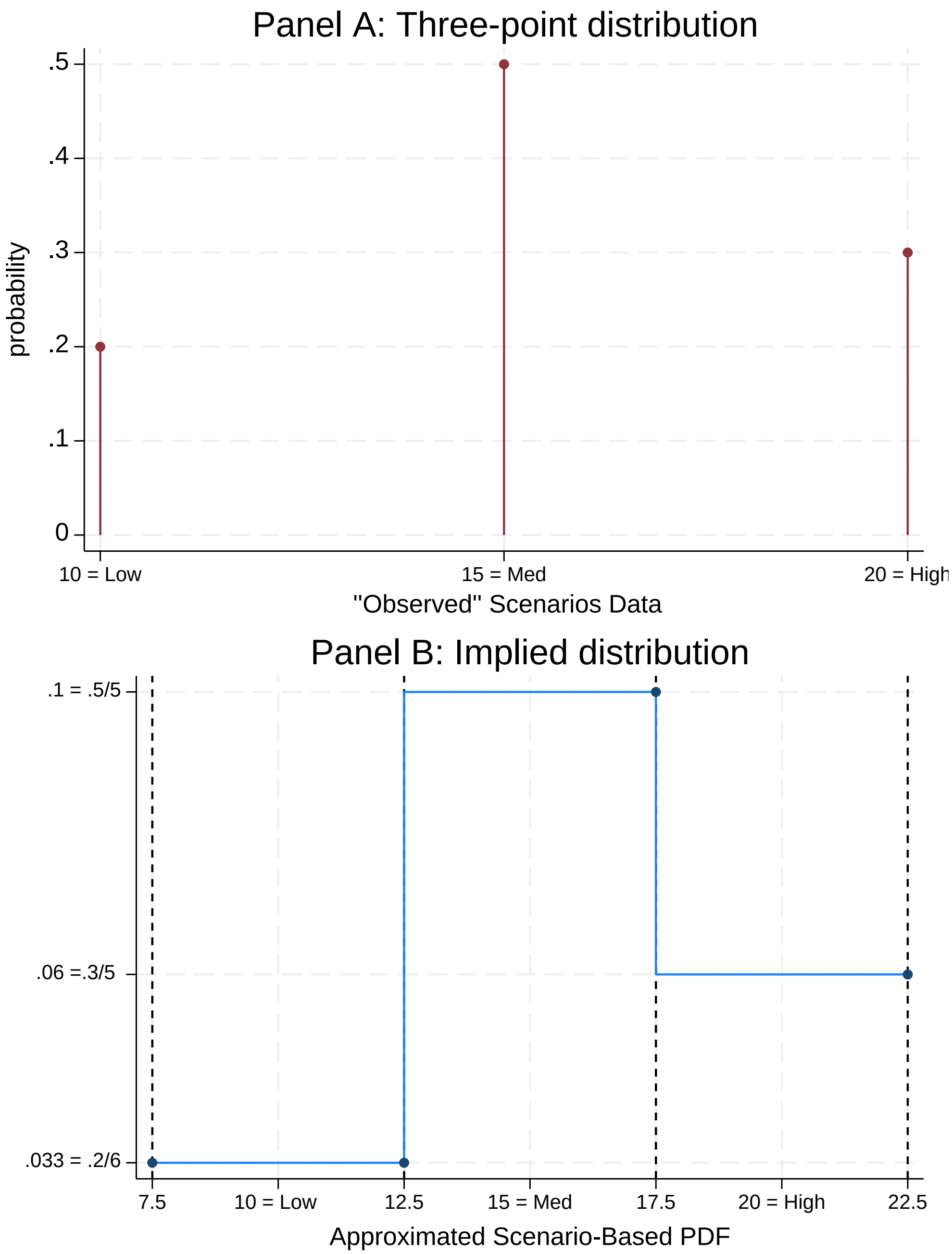


Notes: Panels A-F present binscatter plots. Panels G and H report average (across respondents) cumulative distribution functions (CDFs). Blue solid line is from bins-based elicitation. Red, long-dash line is from scenario-based elicitation. Green, dashed line (Panel G) is from bins-based elicitation in the Survey of Consumer Expectations. Black, short-dash line is the 45° line.

Online Appendix A

Additional Figures and Tables

Appendix Figure A1. Stylized Example of PDF Interpretation for Scenarios Data.



The left panel shows a stylized example of reported expected inflation values and corresponding probabilities for low- medium- and high-inflation scenarios. The right panel demonstrates the method we use to interpret the data as a PDF. In principle, we set each scenario to be the midpoint of a uniformly distributed range of values around that point. The probability within each range is given by the corresponding scenario probability, divided by the number of discrete points in the range of values. As shown above, if an individual reports three scenarios, this results in a 4- point mapping required to pin down the approximated PDF (and CDF). The method is analogous for individuals who report two scenarios, i.e., the approximated distribution has a 3-point mapping. For individuals that report a single scenario with 100% probability, we interpret the CDF as-is.

Appendix Table A1. Predictors of expectations.

	Inflation					Unemployment				
	Level			Uncertainty		Level			Uncertainty	
	Point	Bins	Scenarios	Bins	Scenarios	Point	Bins	Scenarios	Bins	Scenarios
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.77***	0.64***	0.93***	0.83***	0.28***	1.87***	1.26***	1.52***	0.55***	0.58***
	(0.09)	(0.08)	(0.10)	(0.08)	(0.04)	(0.12)	(0.10)	(0.13)	(0.07)	(0.05)
Age	0.08***	0.08***	0.07***	-0.09***	0.02***	0.02	-0.04**	0.04	-0.07***	-0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.01)	(0.01)
$Age^2/100$	-0.04**	-0.05***	-0.03*	0.05***	-0.01	-0.02	0.01	-0.02	0.04***	0.01
	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)
Log Income	0.10	0.19**	-0.43***	-1.45***	-0.02	-0.18	-0.60***	-0.33*	-0.99***	0.05
	(0.10)	(0.09)	(0.13)	(0.09)	(0.05)	(0.17)	(0.11)	(0.19)	(0.07)	(0.07)
Republican	0.40***	0.38***	0.27**	-0.27***	0.02	-0.35**	-0.31***	-0.09	-0.28***	0.07
	(0.10)	(0.09)	(0.11)	(0.09)	(0.04)	(0.14)	(0.11)	(0.15)	(0.07)	(0.06)
Green Party	0.56	0.85	0.60	0.23	0.57*	-1.23	0.06	-0.75	0.27	0.60
	(0.56)	(0.67)	(0.62)	(0.59)	(0.31)	(1.14)	(0.74)	(1.14)	(0.55)	(0.40)
Libertarian Party	0.71**	1.43***	0.01	-0.71***	0.04	-0.42	-0.49	-0.79*	-0.82***	-0.27*
	(0.32)	(0.31)	(0.29)	(0.25)	(0.14)	(0.44)	(0.35)	(0.41)	(0.22)	(0.16)
Other Party	0.06	0.13	0.08	0.06	-0.03	0.36**	0.23*	0.13	-0.02	0.08
	(0.11)	(0.10)	(0.12)	(0.10)	(0.05)	(0.16)	(0.12)	(0.17)	(0.08)	(0.07)
Party not reported	-0.44***	-0.27**	-0.21	0.90***	-0.21***	-0.42*	0.86***	-0.80***	0.44***	-0.31***
	(0.15)	(0.13)	(0.17)	(0.14)	(0.07)	(0.23)	(0.16)	(0.25)	(0.11)	(0.09)
Some high school	-1.28***	-1.65***	-0.50	2.30***	-0.13	-2.94***	-0.08	-2.09***	1.08***	-0.39
	(0.33)	(0.27)	(0.45)	(0.29)	(0.19)	(0.55)	(0.43)	(0.79)	(0.28)	(0.26)
Graduated high school	-0.32***	-0.67***	-0.72***	0.78***	-0.35***	-0.82***	-0.02	-1.06***	0.40***	-0.50***
	(0.12)	(0.11)	(0.14)	(0.11)	(0.05)	(0.18)	(0.13)	(0.19)	(0.09)	(0.07)
Some college	-0.03	0.11	0.02	0.13	-0.05	0.15	0.33***	-0.16	0.01	-0.18***
	(0.10)	(0.10)	(0.12)	(0.09)	(0.05)	(0.15)	(0.12)	(0.16)	(0.07)	(0.06)
Post college graduate	-0.33***	-0.12	-0.64***	-0.53***	-0.10*	-0.31*	-0.25*	-0.29*	-0.06	-0.17**
	(0.11)	(0.11)	(0.12)	(0.11)	(0.05)	(0.17)	(0.14)	(0.17)	(0.09)	(0.07)
Under 30 hours of work	-0.45***	-0.04	-0.14	0.01	0.04	0.22	-0.02	-0.08	0.06	0.06
	(0.13)	(0.13)	(0.16)	(0.12)	(0.06)	(0.20)	(0.16)	(0.22)	(0.10)	(0.09)
30-34 hours of work	-0.33*	-0.01	-0.10	0.33**	0.12	0.80***	0.10	0.66**	0.24*	0.09
	(0.18)	(0.18)	(0.21)	(0.17)	(0.09)	(0.30)	(0.21)	(0.32)	(0.13)	(0.11)
Not employed for pay	-0.11	-0.11	-0.35***	-0.00	-0.10**	0.04	-0.13	0.32*	0.05	0.00
	(0.11)	(0.09)	(0.12)	(0.10)	(0.05)	(0.16)	(0.12)	(0.17)	(0.08)	(0.06)
Observations	7,386	7,677	6,790	7,845	7,103	7,124	7,237	6,618	7,289	6,701
R-squared	0.04	0.04	0.03	0.17	0.02	0.04	0.06	0.03	0.12	0.03

Notes: The table reports estimates for regressions of expectations on sociodemographic characteristics of responses. All specifications are estimated using Huber robust regression. ***,**,* denote statistical significance at 1, 5, and 10 percent levels.

Appendix B

Construction of CDFs

Naturally, one could wonder how sensitive our results are to the methods we use to interpret data as PDFs. Specifically, in our handling of the scenarios data, we assume a particular method for interpreting individual PDFs (see Appendix A Figure 1), choose the support of the aggregate distribution as [-20, 50], and use linear interpolation to smooth the right side of the aggregate distribution. At the same time, we take a conventional approach when analyzing the bins-based data, using the NY Fed's implicit range of [-16, 16] as the distribution support, and assuming the probability within each bin is distributed uniformly. To ensure that the key features of the bins- and scenarios- based aggregate CDFs are robust to minor differences in data treatment, we abstract from these empirical choices, to the extent possible, and instead use the parametric approach of Engelberg, Manski, and Williams (2009) (hereafter referred to as EMW). After fitting the individual PDFs to continuous parametric distributions, we show that the implied aggregate CDFs are strikingly similar to those in our main results (see Figure B1). Below, we describe the EMW method, including some minor adaptations to fit the scenarios data, and discuss the aggregate CDF results.

In the EMW method, individual bins-based probability data are fitted to parametric distributions using a small set of assumptions on the parameters and estimation with non-linear least squares, as needed. The target parametric distribution for each response is determined by the number of bins (or scenarios) used. One- and two-bin (scenario) responses are fitted to the uniform and isosceles triangular distributions, respectively.[12] Responses including three or more bins (or three scenarios) are fitted to the generalized beta distribution.

For any distribution, we denote the set of parameters as $\theta = \{\eta, l\ s\}$. l is the location parameter, equivalent to the left endpoint of the support, and $s = r - l$ is a scale parameter, equal to the distance between the right and left endpoints, and η is a set of shape parameters, whose elements depend on the specific distribution. For the beta distribution, $\eta = \{\alpha, \beta\}$. For the isosceles triangular distribution, $\eta \equiv .5$ by definition of isosceles. Finally, since the uniform distribution does not take a shape parameter, $\eta = \{\cdot\}$. We describe each of the parametric distributions, using the notation stated above.[13]

[12] Responses including only two bins which are not adjacent are not fitted to any parametric distribution. Due to this, these observations are omitted from the sample throughout the analysis in this paper.
[13] Our PDF notation differs from Armantier (2017) and EMW (2009) and reflects our use of the SciPy library in Python to estimate the default "location" and "scale" parameters (as opposed to estimating left and right endpoints).

Generalized Beta Distribution

$$\eta = \{\alpha, \beta\}; \theta = \{\{\alpha, \beta\}, l, s\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} 0, & x < l \\ \dfrac{(x-l)^{\alpha-1}(l+s-x)^{\beta-1}}{B(\alpha,\beta)s^{\alpha+\beta-1}}, & l \leq x \leq l+s \\ 0, & x > l+s, \end{cases}$$

where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

As stated in EMW, one of the main advantages of using the generalized beta distribution is that it is highly flexible, by dint of the two shape parameters. By the same token, the low number of data points (between 3 to 10) per response relative to the number of parameters (up to four) implies that estimating $\widehat{\theta_i}$ precisely is a challenge, and estimates are notably sensitive to the initial guess, θ_0, as a result. Furthermore, in the bins data, it is relatively rare to observe $\widehat{r_i} \gg r_0$ or $\widehat{l_i} \ll l_0$. We attribute the lack of substantial variation in the endpoint estimates to compression of the underlying bins data, related to priming.

Isosceles Triangular Distribution

$$\eta = \{.5\}; \theta = \{\{.5\}, l, s\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} \dfrac{4(x-l)}{s^2}, & l \leq x \leq \dfrac{s+2l}{2} \\ \dfrac{4}{s^2}(s+l-x), & \dfrac{s+2l}{2} < x \leq s+l \\ 0, & otherwise \end{cases}$$

Uniform Distribution

$$\eta = \{\cdot\}; \theta = \left\{\{\cdot\}, l, s\right\}$$

The probability distribution function is given by:

$$f(x;\theta) = \begin{cases} \frac{1}{s}, & l \leq x \leq s + l \\ 0, & otherwise \end{cases}$$

Whenever the uniform distribution is used, we assume $l_i = \overline{l_i}$ and $r_i = \bar{r}_i$, which together imply $s_i = \overline{s_i}$. Since all parameters in θ are known, no estimation is required.

After classifying responses by parametric distribution type, we split them into cases depending on which, if either, of the distribution's support endpoints are known, given the assumptions in the EMW method. For any given distribution, the general endpoint cases are defined as follows. (See Table B2 for additional information on how we assign support ranges according to distribution type and endpoint case.)

- <u>Case 0: l, r unknown; $\theta = \{\eta, l, s\}$</u>

  This case describes PDFs for which neither endpoint of the support is known. Both l, s are estimated using non-linear least squares.

  Responses are classified as Case I if:

  <u>Bins:</u> the respondent uses 3+ bins, including the largest and smallest bins.

  <u>Scenarios:</u> 3 distinct scenarios, each with positive probability are used.

- <u>Case 1: l unknown, $r = \bar{r}$; $\theta = \{\eta, l\}$</u>

  This occurs when the max value of the support is pinned down by the data, based on the method of EMW. In this case estimating l is sufficient for s, since $s = \bar{r} - l$.

  Responses are classified as Case II if:

  <u>Bins:</u> the respondent uses at least 3 bins, including the lowest (left-censored) bin, but excluding the highest (right-censored) bin; or, the respondent uses exactly 2 bins, the higher bin used is not censored, and has higher probability than the lower bin.

Scenarios: the respondent uses 2 distinct scenarios with positive probabilities. In addition, the higher scenario value is assigned a higher probability.

- Case 2: r unknown, $l = \bar{l}$; $\theta = \{\eta, s\}$

  This occurs when the min value of the support is pinned down by the data, based on the method of EMW. In this case estimating s is sufficient to recover r , since $r = \bar{l} + s$. Responses are classified as Case III if:

  Bins: the respondent uses at least 3 bins, excluding the lowest (left-censored) bin, but including the highest (right-censored) bin; or, the respondent uses exactly 2 bins, the lower bin used is not censored, and has higher probability than the higher bin.

  Scenarios: the respondent uses 2 distinct scenarios with positive probabilities. In addition, the lower scenario value is assigned a higher probability.

- Case 3: $l = \bar{l}, r = \bar{r}; \theta = \{\eta\}$

  This case describes situations in which both endpoints are pinned down by the observed data, based on the assumptions in EMW. We estimate only η, if needed. Recall that $\eta \equiv .5$ for the isosceles triangular distribution, and $\eta = \{.\}$ for the uniform distribution, so estimation is required only if the target distribution is the generalized beta distribution.

  Responses are classified as Case IV if:

  Bins: the respondent uses exactly 1 bin.

  Scenarios: the respondent uses only 1 distinct scenario with 100% probability, or the respondent uses two bins, each having 50% probability.

After fitting each response to the corresponding PDF type, we use the resulting parameter estimates to obtain micro-level CDF values across the grid ranging [-20, 40]. In estimation, the location (left endpoint) parameter is bounded below at -100, and the scale parameter is unconstrained. If applicable, the shape parameters are also unconstrained. We find that the estimation results are relatively invariant to alternative assumptions on the bounds. Using the parametric approach delineated above, we show that the implied aggregate CDFs for both the bins and scenarios subsamples track closely with the original CDFs used in our main results. One noteworthy difference is that the tails of parametric scenarios CDF are flatter than those of the corresponding non-parametric curve, which implies the extreme values of the support could lie beyond [-20,40] This finding bolsters our view that the conventional cut-off values of ±16 for the bins-based distributions are unrealistic. Overall, the similarity between the aggregate parametric and non-parametric curves provide evidence that our results to obtain even when using an alternative set of assumptions over the underlying data. This reinforces the validity of our original approach and demonstrates the robustness of our main results.

Figure B1: Comparison of Parametric and Non-Parametric Inflation CDFs

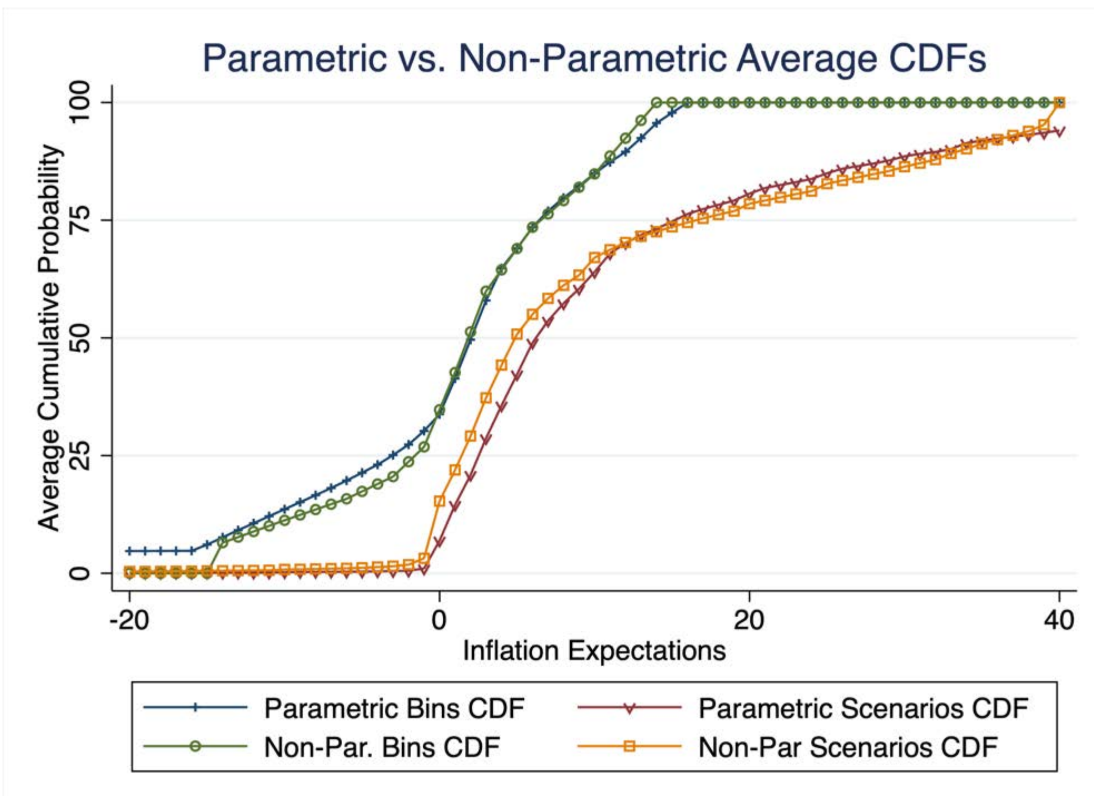


Table B1: Count of Responses by Parametric Distribution Type and Endpoint Case

# Bins/ Scenarios	Case 0	Case 1	Case 2	Case 3
1 (Uniform)				
Bins	0	0	0	1840
Scenarios	0	0	0	2347
2 (Triangular)				
Bins	0	244	237	384
Scenarios	0	409	682	322
3+ (Beta)				
Bins	3129	1595	307	1119
Scenarios	0	0	0	5095
N = 8855				

Table B2: Guide to Assigning Support by Parametric Distribution Type and Endpoint Case

Distr. Type	Case I	Case II	Case III	Case IV
Uniform	N/A	N/A	N/A	Bins: If $\bar{l}_i$ ($\bar{r}_i$) censored, use -16 (16). Scen: Support is $[\pi_i^e - \delta, \pi_i^e + \delta]$, $\delta \to 0$. No estimation required.
Triangular	N/A	$\bar{r}_i = UB^i_{bins;\,scen}$ $\hat{s}_i \equiv \bar{r}_i - \hat{l}_i$ Estimate l_i.	$\bar{l}_i = LB^i_{bins}$ $\hat{r}_i \equiv \hat{s}_i - \bar{l}_i$ Estimate s_i.	Support is $[LB^i_{bins;\,scen}, UB^i_{bins;scen}]$ No estimation required.
Beta	Endpoints unknown. Estimate $\eta_i\ l_I, s_i$.	$\bar{r}_i = UB^i_{bins;\,scen}$ $\hat{s}_i \equiv \bar{r}_i - \hat{l}_i$ Estimate η_i, l_i.	$\bar{l}_i = LB^i_{bins;\,scen}$ $\hat{r}_i \equiv \hat{s}_i - \bar{l}_i$ Estimate η_i, s_i.	Support is $[LB^i_{bins;scen}, UB^i_{bins;scen}]$ Estimate only η_i.