A Choice-Based Approach to the Measurement of Inflation Expectations

Olga Goldfayn-Frank[1], Pascal Kieren[2], Stefan Trautmann[2]

[1] *Deutsche Bundesbank, Germany*
[2] *Heidelberg University, Germany*

Abstract: In applied research and policy analysis, economists have increasingly relied on measures of inflation expectations and uncertainty elicited via density forecasts. This method, where respondents assign probabilities to pre-specified ranges, has been subjected to criticism particularly in the recent times of high and volatile inflation. We propose a new method to elicit inflation expectations, which is rooted in decision theory and can be implemented in a standard survey. We demonstrate that it leads to well-defined expectations with central tendencies close to the corresponding point forecasts and to lower forecast uncertainty than density forecasts. The reported difficulty of the questions and survey time required are less or similar to the current standard. In contrast to the density forecasts, our method is robust to differences in the state of the economy and thus allows comparisons across time and across countries. The method is portable in the sense that it can be applied to different macroeconomic measures.

Keywords: Inflation expectations, measurement, macroeconomic beliefs, surveys
JEL: D84, E31, E37, E71

---

*Corresponding author: Stefan Trautmann, Bergheimer Straße 58, 69115 Heidelberg, Germany; trautmann@uni-hd.de. For valuable comments, we thank Michael Weber, Pär Österholm, Bernd Hayo, as well as seminar participants at the Bundesbank, De Nederlandsche Bank, the European Central Bank, the University of Heidelberg, the University of Sydney, and the U.S. Securities and Exchange Commission. We thank Manja Buchheit and Luzia Rode for providing excellent research assistance. Funding by the German Federal Ministry of Education and Research (BMBF) and the Baden-Württemberg Ministry of Science as part of Germany's Excellence Strategy is gratefully acknowledged. This paper represents the authors' personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank.

1. Introduction

Households' and firms' subjective inflation expectations play a central role in macroeconomic and intertemporal microeconomic models. Because of this prominent role, the question how to elicit such beliefs from the public is an essential issue. Inflation expectations are typically measured via large representative surveys such as the Survey of Consumer Expectations (SCE) by the Federal Reserve Bank of New York or the Consumer Expectations Survey by the European Central Bank. It is common amongst these surveys to measure expectations via probabilistic questions: respondents are shown several intervals (or *'bins'*) and are then asked to attach probabilities to all intervals that represent their beliefs about the future rate of inflation.[1]

One concern with this type of questions is that they might be cognitively demanding for many respondents, which leads to higher dropout rates that may ultimately bias inference (D'Acunto et al., 2022). Also, these questions typically have narrower bandwidths around zero than at the extreme ranges, which possibly induces survey participants to perceive that values close to zero are considered more likely by the designers of the survey (Weber et al., 2022). Additionally, recent research suggests that the elicited inflation expectations are neither robust to changes in the wording of the question (e.g., Bruine de Buin et al., 2012, Manski, 2018; Coibion et al., 2020) nor to changes in the response scale (Becker et al., 2023). These issues have no easy solution, because adjusting the size and values of bins across survey waves makes it difficult to compare survey responses over time and across countries. Recently, this issue is further exacerbated by the fact that in periods of high inflation, responses may end up being lumped in extreme bins, which effectively renders them uninformative. Thus, changes in expectations over time, often of most interest to macroeconomists, may only be poorly measured because of the framing biases when using probabilistic question formats.

We propose a new method to elicit subjective inflation expectations via surveys. We build on prior research in decision theory by Baillon (2008) to design a simple choice-based procedure to elicit the distribution of respondents' expectations. The method is based on a bisection process – i.e. the iterated partition of a state space into two equally large subevents, for which preferences are elicited through binary choices. The method starts by asking the respondent for a minimum and

[1] Manski (2018) and Curtin (2019) provide overviews of the literature on survey measures of probabilistic expectations, while Curtin (2007) provides an overview of similar surveys worldwide. Moreover, Dräger and Lamla (2024) provide an excellent overview of available survey data and issues regarding the measurement of macroeconomic expectations.

maximum level of inflation for which they think that there is "almost no chance" that actual inflation will lie outside the interval. Once this interval is established, respondents answer several series of binary choices that follow a strict algorithm until a predetermined level of precision is reached.

To validate this method in the field, we conduct a survey with UK respondents in which we elicit inflation expectations using one of four protocols. Each respondent is exposed to exactly one protocol. The first two elicitation protocols are based on two variations of our proposed method, which we will refer to as '*midpoint method*'. The two implementations differ in the precision of the elicited values, where lower precision allows for fewer steps in the elicitation, making the survey shorter. The other two protocols are based on the probabilistic question format, which is either symmetric and anchored around zero (as in the SCE) or symmetric and shifted to be anchored around respondents' point forecasts (as recently used in surveys of countries with a substantially higher level of inflation, e.g. Central Bank of Turkey; Gülsen and Kara, 2019).[2]

Our results indicate that there are notable differences in the distribution of subjective inflation expectations depending on the underlying elicitation method. For instance, we show that the implied mean of respondents' one-year ahead inflation expectations is substantially higher when elicited though a midpoint method than when elicited through the probabilistic question format of the SCE (6.2% versus 5.1%; realized inflation in the UK at the time of the survey was 6.3%), and similar to the shifted probabilistic format (6.3%). Additionally, the midpoint method leads to significantly lower forecast uncertainty (measured by the standard deviation of respondents' forecasts) relative to both methods that rely on density forecasts. Both midpoint methods and the shifted probabilistic format lead to higher disagreement (measured by the standard deviation of the respondents' mean forecasts in the sample) relative to the SCE density forecasts.

In a second step, we compare the consistency and validity of elicited inflation expectations across our four methods. We show that the midpoint method leads to a higher correlation between implied mean forecasts and respondents' point forecast relative to the probabilistic question format of the SCE (0.73 and 0.51, respectively). Additionally, when inflation expectations are elicited through a midpoint method, on average only 2.1% of the total probability is allocated to deflation scenarios. However, when inflation expectations are elicited through a probabilistic question

[2] The baseline scale was originally designed for the Federal Reserve Bank of New York's Survey of Consumer Expectations but has since been used in many consumer surveys, see Armantier et al. (2017) for an overview.

format, an average of 10.2% of the total probability mass is allocated to deflation scenarios. This stands in contrast to the fact that when respondents are explicitly asked to state the lowest possible inflation that they deem conceivable, only 4.9% even consider deflation in their answer. Given that the current rate of inflation at the time of the survey was 6.3% in the UK, it appears that the probabilistic question format induces respondents to allocate probability mass to deflation scenarios they would otherwise not consider. Next, we compare both methods with respect to perceived length and difficulty as well as in their ability to predict planned durable consumption. We show that the midpoint method is neither perceived to be more difficult nor to take more time to complete. Although our method consists of multiple screens with several questions, we show that also actual time taken is almost identical (even a bit shorter) to the time taken to complete a single density forecast.

Finally, the implied mean forecasts of our midpoint method suggest a negative relationship between planned durable consumption and inflation expectations. This negative relationship is not only consistent with prior studies (e.g., Bachmann et al., 2015; Coibion et al., 2023), but is also obtained when using participants' point forecasts. In contrast, the implied mean forecast of the SCE question format suggests a positive (albeit insignificant) relationship between planned durable consumption and inflation expectations, further reinforcing concerns regarding the validity of elicited inflation expectations in higher inflation regimes.

We contribute to an ongoing literature that aims to elicit subjective inflation expectations through surveys. Manski (2004) popularized the use of survey questions that elicit subjective probability distributions such as the probabilistic question format used in the SCE. Since then, subjective probability distributions have not only been used to infer a central tendency (such as mean or mode), but also to capture uncertainty in individuals' forecasts (Coibion et al., 2024). Although prior studies have shown that consumers are generally able to provide meaningful probability distributions which are consistent with their point predictions (Buine de Bruin et al., 2011; Zhao, 2023), the literature suggests that the probabilistic question format suffers from several problems that the method we propose is able to resolve.[3] It is robust to differences in the state of the economy and therefore allows comparison across time (e.g., in typical survey panels) with potentially very different levels of inflation, and across countries. Additionally, the method requires

[3] A related literature investigates how literate and informed consumers are about inflation and economic policy (e.g., Blinder and Krueger, 2004, for the US; Conrad et al., 2022, for Germany; and Andersson et al., 2022, for Sweden).

neither reference to the concept of probability nor a direct judgement, but only simple binary comparisons. It is thus cognitively easier and less prone to anchoring than allocating probability levels to different inflation bins.[4] Finally, the method is portable in the sense that it can be applied to different macroeconomic measures.

The rest of the paper proceeds as follows. Section 2 describes the method we propose and illustrates it using an example. Section 3 describes the survey design and provides details on data collection. Finally, Section 4 presents results and Section 5 discusses practical issues when applying the proposed method. Section 6 concludes.

2. The Midpoint Method

2.1 Algorithm Description

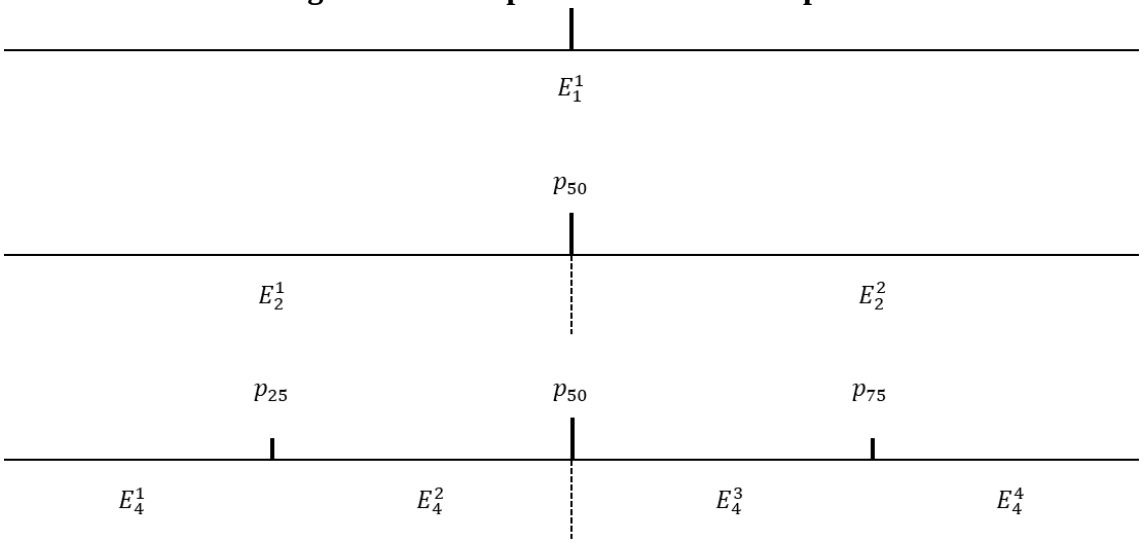
In this section, we describe the midpoint method that we use to elicit percentiles of respondents' inflation expectations. The particular method that we use in this study has originally been developed by Baillon (2008) and focuses on a choice-based implementation of methods which are based on a bisection process, i.e., the subsequent partition of the state space into two equally likely subevents. The intuition behind such bisection methods was first introduced by Ramsey (1931) and Fellner (1961). The method itself is described in Raiffa (1968) using judgements and by Spetzler and Staël von Holstein (1975) in terms of judgements and choice. Finally, Hong and Sagi (2006) formally derive the existence of probabilistic beliefs from this concept.

The basic idea of the method is that two events are exchangeable for an agent when the agent is indifferent to permutations of their outcomes. Such events are thus revealed to be equally likely. At first, the method elicits two complementary events, E_2^1 and E_2^2. The first event (E_2^1) involves outcomes that range from some minimum (or theoretically $-\infty$) to the median (p_{50}). The second event (E_2^2) involves outcomes that range from the median to some maximum (or theoretically $+\infty$). From this twofold partition of the state space, a fourfold one can be generated by splitting each of the two events into two equally likely subevents. As each partition of the state space represents an equally-likely event, the subjective probability distribution can be inferred in this manner. For

[4] Pre-defined answer options provide an anchor for answers (Hjalmarsson and Österholm, 2021), and may also induce respondents to provide an answer even though they do not have clearly defined inflation views which leads to substantial noise (see Hayo and Méon, 2023). The binary choices in our method help respondents to better discover and express their beliefs.

instance, similar to how the median can be inferred from the twofold partition of the state space, the 25-% quartile (p_{25}) and 75-% quartile (p_{75}) can be inferred from the fourfold partition of the state space. Figure 1 visualizes this process. In the following, we describe how the method is implemented to measure respondents' inflation expectations. While we elicit a fourfold partition of the state space in the experiment, in principle the process can be continued by further splitting up (some of the) elicited quartiles (see Section 5).

Figure 1: Decomposition of the State Space



Notes: The figure illustrates how the midpoint method divides the state space into equally likely subevents. The figure ends with a fourfold partition of the state-space from which the 25th-percentile, the median, and the 75th-percentile can be inferred.

The method starts by asking respondents for a minimum (b_0) and maximum (b_1) level of inflation for which they think that there is "absolutely no chance" that actual inflation will lie outside the interval. The interval has little empirical relevance and is theoretically not needed as one could work with unbounded intervals (see Baillon, 2008) but helps respondents to structure their beliefs. Once the minimum and maximum are established, the bisection process begins. In each step of the process, respondents must decide which of two possible events regarding the rate of inflation in the next year they consider to be more likely. The first event always contains ranges from the minimum (b_0) to some midpoint (m), while the second event always contains ranges from the midpoint to the maximum (b_1). The initial midpoint (m_1) which is used to dissect the state-space is half the distance between b_0 and b_1: $b_0 + \frac{b_1 - b_0}{2}$. Consequently, the first question asks subjects which of the following two events regarding the rate of inflation in 12 months they

consider more likely: $\left[b_0, b_0 + \frac{b_1 - b_0}{2}\right]$ or $\left(b_0 + \frac{b_1 - b_0}{2}, b_1\right]$. After respondents made their choice, the bisection process continues with the goal to find a midpoint for which subjects consider each of the two events to be equally likely. The calculation of the next midpoint depends on whether subjects considered the event with lower or higher rates of inflation to be more likely. If subjects think the event with lower rates of inflation is more likely to occur, the next midpoint is determined as half of the distance between the minimum (b_0) and the most recent midpoint (m_1). If subjects think the event with higher rates of inflation is more likely to occur, the next midpoint is determined as half of the distance between the most recent midpoint (m_1) and the maximum (b_1). In general, whenever subjects think that the event with lower rates of inflation is more likely to occur, the next midpoint is determined as half of the distance between the current midpoint and the next lower midpoint (or the minimum if there is no lower midpoint). Similarly, whenever subjects think that the event with higher rates of inflation is more likely to occur, the next midpoint is determined as half of the distance between the current midpoint and the next higher midpoint (or the maximum if there is no higher midpoint). The intuition of this process is that whenever subjects think that one particular event is more likely, the range of that event is reduced for the following question.

This bisection process then continues until a predetermined level of precision (between the two most recent midpoints) is reached. In the survey condition "*Midpoint Endogenous*" (see for details below), we set the precision to 1.5 percentage points. This final midpoint (m_{p50}) then represents the point for which subjects consider each of the two events equally likely. This point also represents the median of the subjective probability distribution. Afterwards, the process continues to identify both the 25th-percentile and the 75th-percentile, by splitting the final two events into two equally-likely subevents. For the 25-% percentile, the process thus starts with the two events $\left[b_0, b_0 + \frac{m_{p50} - b_0}{2}\right]$ and $\left(b_0 + \frac{b_{m50} - b_0}{2}, b_{p50}\right]$. For the 75-% percentile, the process starts with the two events $\left[m_{p50}, m_{p50} + \frac{b_1 - m_{p50}}{2}\right]$ and $\left(m_{p50} + \frac{b_1 - m_{p50}}{2}, b_1\right]$. Both processes again continue with the above-described algorithm until a precision of 1.5% is reached. At this point, we can infer the minimum, maximum, median, 25th-percentile, and 75th-percentile of subjects' subjective probability distribution of the 1-year ahead rate of inflation.

Figure A.1 in Appendix A illustrates the mapping of two exemplary latent belief distributions on elicited distributions using the midpoint method with 2 iteration steps (i.e., with a potentially low precision) vs. the SCE Bins method. These are discussed in Chapter 5.

2.2 Example

The example in Table 1 illustrates how the bisection process works. Suppose a hypothetical survey respondent believes that the lowest possible rate of inflation next year is 0% (b_0) and that the highest possible rate of inflation is 20% (b_1). Based on this interval, the first midpoint is 10%. The respondent next indicates option B as more likely (that is, she believes a higher rate of inflation is more likely). The second midpoint thus is 15% $\left(=\frac{20\%+10\%}{2}\right)$. Based on the two events that follow from the second midpoint, the respondent prefers option A, which leads to a new midpoint of 12.5% $\left(=\frac{15\%+10\%}{2}\right)$. Given that the desired level of precision is not yet reached (1.5% < 15% –12.5%), the elicitation process continues, and the respondent again prefers option A. The next midpoint is thus 11.25% $\left(=\frac{12.5\%+10\%}{2}\right)$. This midpoint satisfies the precision criterion as it is less than 1.5% different from the previous midpoint, which concludes the first bisection process. The final midpoint also serves as the median of the inferred subjective probability distribution.

Table 1: Illustration of the Bisection Process

Searched percentile	***Option A***	***Option B***	***Choice***
p_{50}	[0%; 10%]	(10%; 20%]	B
	[0%; 15%]	(15%; 20%]	A
	[0%; 12.5%]	(12.5%; 20%]	A
$\rightarrow p_{50} = 11.25\%$			
p_{25}	[0%; 5.63%]	(5.63%; 11.25%]	B
	[0%; 8.44%]	(8.44%; 11.25%]	A
$\rightarrow p_{25} = 7.03\%$			
p_{75}	[11.25%; 15.63%]	(15.63%; 20%]	A
	[11.25%; 13.44%]	(13.44%; 20%)	A
$\rightarrow p_{75} = 12.34\%$			

Notes: The table illustrates the bisection process for a hypothetical survey respondent who expects inflation in one year to not fall below 0% (b_0) and to not rise above 20% (b_1). The bisection process starts by eliciting the median (p_{50}), and then proceeds to elicit the 25-percentile (p_{25}) and 75-percentile (p_{75}).

The process continues with the elicitation of the 25th-percentile. The first midpoint of 5.63% is calculated from the minimum and the median $\left(=\frac{0\%+11.25\%}{2}\right)$. As the respondent prefers option

B, the next midpoint is 8.44% $\left(=\frac{5.63\%+11.25\%}{2}\right)$. From the next set of options, the respondent prefers A, which identifies the next midpoint as 7.03% $\left(=\frac{5.63\%+8.44\%}{2}\right)$. As this midpoint satisfies the desired level of precision, it serves as the 25th-percentile.

Finally, the process concludes with the elicitation of the 75th-percentile. The first midpoint of 15.63% is calculated from the median and the maximum $\left(=\frac{11.25\%+20\%}{2}\right)$. Since the respondent prefers option A, the next midpoint of 13.44% is calculated as half the difference between the median and the previous midpoint $\left(=\frac{11.25\%+15.63\%}{2}\right)$. Notice that each new midpoint is lower than the previous midpoint if a respondent prefers option A, while it is larger than the previous midpoint if a respondent prefers option B. From the new set of events, the participant again prefers option A, which identifies the final midpoint (and the 75th-percentile) as 12.34%, which satisfies the precision criterion.

3. Data and Survey Design

3.1 Survey Design

In our survey we elicit respondents' full distribution of one-year ahead inflation expectations (consumer price level), using four different elicitation methods in a between-subject design. First, all survey respondents are asked to provide a point forecast regarding the rate of inflation in twelve months. This question is adopted from the New York Fed Survey of Consumer Expectations (SCE). Afterwards, respondents are randomly allocated to one of the four treatments.

The first two treatments labeled "*Midpoint 2-Step*" and "*Midpoint Endogenous*" elicit inflation expectations using the above-described midpoint method. First, respondents are asked to state a minimum and maximum level of inflation for which they think that there is "absolutely no chance" that actual inflation will lie outside the interval. Afterwards, the bisection process begins by eliciting the median, the 25th-percentile, and 75th-percentile. In the "*Midpoint 2-Step*" treatment, the maximum number of steps – i.e., questions for each percentile – is set to two. That is, the process stops after two steps even if the predefined level of precision is not reached yet. As such, subjects answer two questions to elicit the median, two questions to elicit the 25th-percentile, and two questions to elicit the 75th-percentile, unless the predefined level of precision of 1.5% is reached earlier (the number of questions is thus an upper bound). In the "*Midpoint Endogenous*" treatment, the number of steps is determined endogenously and depends on the width of the interval

between the minimum and maximum level of inflation. Participants answer questions until the predefined level of precision of 1.5% is reached.

In the other two treatments labeled "*Bins SCE*" and "*Bins Shift*", respondents are asked to provide a density forecast for 12-month ahead inflation. For the "*Bins SCE*" treatment, we use the response scale of the SCE, which has been widely adopted by other surveys, such as the Bundesbank household survey. The scale has ten intervals, is centered at zero and symmetric, and has two open outer intervals. The closed intervals cover the range of –12% to 12%. The four central intervals are narrower and cover a range of 2%. The outer closed intervals are wider and cover a range of 4%. The Online Appendix provides screen shots illustrating the question format. Respondents are then asked to assign a probability mass of 100% over the ten intervals. In the "*Bins Shift*" treatment, the scale is centered at the respondents' point forecast instead of zero. All outcomes of the SCE scale are shifted by respondents' point forecast. Besides the different centering, the scale is identical to the regular scale of the SCE. The four center intervals are narrower (i.e., cover a range of 2%) than the outer closed intervals (which cover a range of 4%).

Immediately after the elicitation of the inflation expectation, the survey included two questions that assessed the participants' perceived length and difficulty or the elicitation process. Both questions are based on a 5-point Likert scale. These questions aim to control for whether the midpoint method is perceived to be more difficult or more time-intensive relative to the density forecasts, irrespective of the true survey duration, which we also measure.

Afterwards, all treatments continue with the same post-survey-questionnaire. Respondents answer questions about their age, gender, household income, and education. Additionally, we add three questions adopted from the Bundesbank household survey. First, respondents judge the likelihood that their household income over the next year will either decrease, increase, or stay the same. Second, we ask respondents how much money they spend on the following categories within the last month: major purchases (e.g., car or electrical appliances), essential goods (e.g., food and beverages or cleaning products), clothing and footwear, entertainment/recreation (e.g., restaurant visits or cultural events), mobility (e.g., fuel), services (e.g. hairdresser or childcare), travel/holidays, housing costs (e.g., rent or mortgages), and financial reserves. Third, we ask respondents to judge whether they plan to spend more, less, or roughly the same on these categories over the next year.

3.2 Procedures and Sample Demographics

The survey was conducted in September 2023 on the crowdsourcing platform Prolific, using a stratified sample of the UK population. The survey was computerized using oTree (Chen et al., 2016). The design, sample size, analyses are all pre-registered.[5] The survey received ethics-approval of the Institutional Review Board (IRB) of the Faculty of Economics and Social Sciences at Heidelberg University.[6]

In total, 811 respondents completed the survey. Respondents were paid a fixed amount of £2 for completing the survey. On average, it took respondents 6:03 minutes to finish the survey. Based on our payment, respondents earned on average an hourly wage of £19.83, which is well above the average hourly earnings on Prolific and almost double the minimum wage in the UK.

Table 2: Survey Demographics

	Mean	Standard Deviation	Median	Min	Max
Demographics					
Age	41.6	12.97	40	18	82
Female	0.503	0.50	1	0	1
Education (1-9)	4.86	2.05	6	1	9
Financials					
Household Income (in £)	3,574.91	2,598.52	2,750	500	12,000
Spending (in £):					
Major Purchases	440.99	2,774.90	0	0	50,000
Essential Goods	429.59	522.07	350	0	6,000
Clothing and Footwear	81.22	149.41	50	0	2,500
Entertainment/Recreation	118.00	193.89	75	0	3,000
Mobility	176.79	363.97	100	0	5,000
Services	80.65	450.99	20	0	12,000
Travel/Holidays	234.46	933.39	0	0	15,000
Housing Costs	861.94	1,378.10	600	0	20,000
Financial Reserves	603.76	5,323.31	20	0	100,000

Note: This table presents summary statistics. Female is a dummy that equals 1 if a person indicated to be female and zero otherwise. Household income was measured by choosing one of 13 income categories and is calculated by using the midpoint of each category. Reported are mean, standard deviation, median, minimum, and maximum.

[5] The pre-registration document is available at https://aspredicted.org/2T2_7DL.

[6] The reference number of the IRB approval is FESS-HD-2023-002.

Demographics are displayed in Table 2. Respondents are on average 42 years old and 50% are female, which results from the stratified sample collection. They report an average level of education of 4.86 (out of 9), which corresponds to education between a college degree and a Bachelor's degree. In terms of household finances, respondents report an average household disposable income of £3,575, which is higher compared to the median household disposable income of £2,692 in 2022 as reported by the UK's Office for National Statistics. The majority of monthly spending occurs for housing costs (£862), accumulation of financial reserves (£604), and purchase of essential goods (£430).

4. Results

4.1 Survey Summary Statistics

Table 3 presents summary statistics for our four methods. Overall, participants expect the rate of inflation over the next 12 months to be on average 5.51%. This is slightly lower compared to the at the time current rate of inflation of 6.3% in September 2023. Relative to the preceding year where the UK's Office for National Statistics reported a year-to-year inflation of 8.8% (September 2022), this reflects a declining trend. Histograms of the point prediction, minimum, maximum and the implied means for each elicitation methods are in Appendix B (Figure B.1 and B.2).

Respondents who provide density forecasts in the Bins SCE treatment use on average 6.01 (median 5) out of 10 possible bins, and allocate on average 21.11% probability mass to each bin they use in their forecast. In the Bins Shift treatment, it is noticeable that respondents use significantly more bins in their density forecast with an average of 6.69 (median 7) out of 10 possible bins ($p < 0.01$; $t = 2.66$, two-sided t-test). Given that they use more bins in their density forecast, the average probability mass allocated to each bin shrinks to 18.37%. The higher number of bins used in the Bins Shift treatment is not unexpected. By anchoring the response scale around the mean forecast, respondents have access to a wider range of intervals which are closer to their mean expectations. In both conditions, the number of bins used is substantially larger than in some other studies where respondents use only few bins (e.g., Delavande and Rohwedder, 2008). This probably reflects the higher uncertainty in the current inflation environment.

Next, we discuss summary statistics regarding our midpoint treatments. Both the minimum and maximum level of inflation are very similar across the two implementations. For the 2-Step

and Endogenous treatment, respondents on average expect that the lowest possible rate of inflation over the next year is 3.41% and 3.58%, respectively ($p > 0.1$; $t = 0.61$, two-sided t-test). Regarding the highest possible rate of inflation over the next year, respondents on average expect inflation to not exceed 10.18% and 10.06%, for the 2-Step and Endogenous treatment, respectively ($p > 0.1$; $t = 0.20$, two-sided t-test). As none of the differences are significant at the 10%-level, this supports a successful randomization across treatments.

Table 3: Inflation Summary Statistics

	Mean	Standard Deviation	Median	Min	Max
Point Forecast	5.51	4.43	6	-30	40
Bin Treatments:					
#Bins used (SCE)	6.01	2.65	5	1	10
Bin Size (SCE)	21.11	12.15	20	10	100
#Bins used (Shift)	6.69	2.53	7	1	10
Bin Size (Shift)	18.37	10.94	14.29	10	100
Midpoint Treatments:					
Minimum (2-Step)	3.41	2.49	4	-8	15
Maximum (2-Step)	10.18	6.23	9	1	60
Number of Steps (2-Step)	1.61	0.39	2	1	2
Minimum (Endogenous)	3.58	3.09	4	-10	17
Maximum (Endogenous)	10.06	6.15	9	1	50
Number of Steps (Endogenous)	2.11	0.86	2	1	4

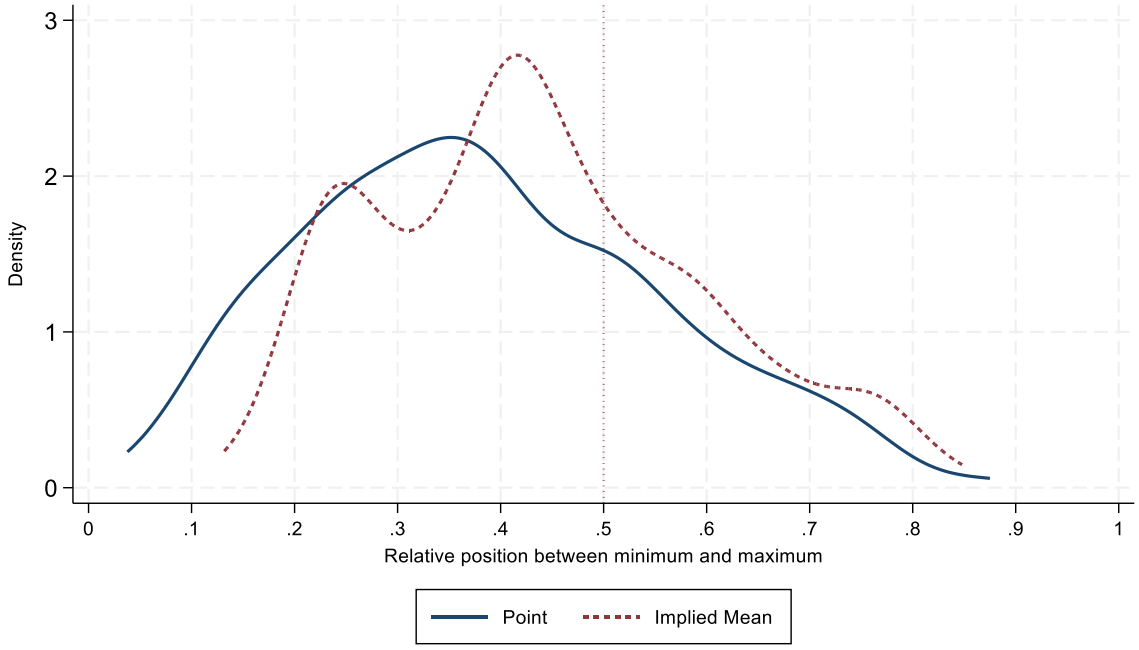
Note: This table presents summary statistics for our measurement methods. Reported are mean, standard deviation, median, minimum, and maximum.

It is noteworthy that the average expected level of inflation – as well as the median level – is closer to the minimum than to the maximum, suggesting that the distribution of inflation expectations is skewed.[7] Figure 2 visualizes this relationship. Displayed is the sample distribution of the relative position of a respondent's point forecast (and the implied mean forecast) in the range between her indicated minimum and maximum level of inflation. In this scale, 0 would imply that the point forecast is identical to the minimum level of inflation, while 1 would imply that the point forecast is identical to the maximum level of inflation. Both distributions are positively skewed

[7] The range between the minimum and maximum level of inflation expectations does not substantially correlate with demographic or financial variables, as shown in Table C.1 in Appendix C.

and contain more probability mass for values below 0.5 (which would imply that the point forecast/implied mean is in the middle between the minimum and maximum level of inflation). Note that the highly asymmetric distribution for the implied mean also implies that participants in the midpoint treatments did not simply answer randomly: on average, random answers would induce a symmetric distribution of means between the minimum b_0 and maximum b_1.

Figure 2: Sample Distribution of the Relative Position of Inflation Estimates between the Minimum and Maximum



Notes: The figure displays the sample distribution of the relative position of a respondent's implied mean or point forecast, between her minimum b_0 (0) and maximum b_1 (1). Blue line displays results for point forecasts, red line displays results for implied means. Results only include Midpoint treatments, for which data on minimum and maximum is available.

Finally, we compare the average number of steps required for each midpoint treatment. The number of steps measures how many questions respondents on average answer for each elicited percentile (i.e., median, 25th-percentile, and 75th-percentile). For the 2-Step treatment, the average number of steps is 1.61 with a median of 2. Note that the 2-Step treatment capped the number of steps at two. However, when respondents provide a relatively narrow interval (measured from the minimum to the maximum), the number of steps is sometimes lower as the level of precision is reached faster. For the endogenous treatment, the number of steps is 2.11 with a median of 2. Although the number of steps is thus slightly higher compared to the 2-Step treatment ($p < 0.01$;

$t = 7.16$, two-sided t-test), results suggest that two steps is often sufficient to approximate respondents' subjective probability distribution in the current inflation environment.

4.2 Elicited Inflation Expectations

In this section, we compare the average implied mean forecast, disagreement, as well as the average forecast uncertainty across the four elicitation methods. The implied mean forecast refers to respondents' average inflation expectations as implied by one of the four elicitation methods. Disagreement refers to the cross-sectional standard deviation of implied means and forecast uncertainty refers to the standard deviation of a respondent's forecast. To infer the statistics, we assume a mass-at-midpoint measure. For the two midpoint treatments, this means that the probability mass in each of the four elicited ranges (i.e., minimum to 25th-percentile, 25th-percentile to median, median to 75th-percentile, and 75th-percentile to maximum) is uniformly distributed for all values contained in that range. For the two treatments that rely on density forecasts, we require an additional assumption regarding the width of the two open intervals. Following prior work by Becker et al. (2023), we assume that the open intervals have twice the width of the adjacent closed interval.[8] Results are reported in Table 4.

Table 4: Inflation Expectations Across Treatments

	Treatments				Differences		
	Midpoint 2-Step (N=200)	**Midpoint Endo (N=205)**	**Bins SCE (N=200)**	**Bins Shift (N=206)**	**Midpoints vs. Bins**	**Midpoints vs. Bins SCE**	**Midpoints vs. Bins Shift**
Implied Mean Forecast	6.19	6.22	5.09	6.27	0.52**	1.12***	-0.06
					(2.07)	(3.77)	(0.19)
Disagreement	3.65	3.65	2.98	3.88	0.14	0.67***	-0.23
					(0.93)	(0.67)	(1.13)
Uncertainty	1.51	1.41	3.78	3.14	-2.46***	-2.32***	-2.02***
					(18.60)	(17.53)	(15.24)

Note: Disagreement refers to the cross-sectional standard deviation of the implied mean inflation expectations (1-year ahead). Uncertainty refers to the standard deviation of the respective question format (within-subject). Tests for the equality of moments are based on t-test with unequal variance. Tests for the equality of disagreement are based on Levene's test of homogeneity of variances. Reported are coefficients and t-statistics/f-statistics (in parentheses). ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

For the average implied mean forecast, we find that both midpoint treatments as well as the mean-shifted density forecast treatment produce similar results (6.19%, 6.22%, and 6.27%,

[8] Since both treatments only contain a small probability mass in the open interval, changing this assumption only marginally affects results.

respectively), with no significant difference between the treatments. In contrast, the average mean forecast elicited through the regular SCE question format is significantly lower relative to the other treatments with a mean of only 5.09% ($p < 0.01$ for all comparisons; $t = 3.32 / 3.41 / 3.43$, two-sided t-test). For forecast disagreement, we find a similar picture. Both midpoint treatments as well as the mean-shifted density forecast treatment lead to similar disagreement (3.65, 3.65, and 3.88, respectively) with no significant differences between the treatments. In comparison, disagreement in the regular SCE question format is significantly lower with 2.98 (all differences significant at the 1%-level, Levene's test of homogeneity of variances). It is noteworthy that the lower disagreement as measured by the SCE question format does not necessarily reflect that respondents are more aligned in their forecasts. Instead, some of the effect is mechanical in nature. Both the midpoint measures and the shifted bins allow for more personalization than the SCE bins. In the latter, all disagreement derives from different allocation of probability to the bins. In contrast, in the midpoint methods each person is asked to provide a distribution over their personalized range (defined by their elicited minimum and maximum expectations), and in the mean-shifted density forecast the center of the bin structure is also personalized (by their elicited point forecast).

Finally, we investigate differences in forecast uncertainty. As shown in Table 4, average forecast uncertainty is substantially lower for the 2-Step and Endogenous Midpoint treatments (1.51, 1.41, respectively) than for the Bins SCE and Bins Shift treatments (3.78, 3.14, respectively). All differences between the midpoint treatments and bins treatments (either individually or combined) are statistically significant at the 1%-level. Additionally, uncertainty in the Bins Shift treatments is significantly lower than in the Bins SCE treatment ($p < 0.01$; $t = 3.66$, two-sided t-test). As such, the average respondent faces the most uncertainty when asked to provide density forecasts which are not centered around their mean expectation. Overall, these results are consistent with a recent study by Pavlova (2023) who shows that asking for a minimum and maximum level of inflation as well as the mode (instead of relying on density forecasts) reduces uncertainty by 0.6 to 1 percentage points.

Overall, these findings are consistent with recent concerns that density forecasts are not invariant to the presentation format (Becker et al., 2023) and induce survey respondents to perceive that values close to zero are considered more likely by the designers of the survey (Weber et al., 2022). Although anchoring the center around the mean forecast helps respondents to some extent, our results suggest that a substantial amount of uncertainty remains. In contrast, the midpoint

method substantially reduces forecast uncertainty, potentially because it does not require respondents to exercise precise probability judgements.

4.3 Method Comparison

In the following, we will delve deeper into differences between elicited inflation expectations through the midpoint method versus density forecasts. We proceed in four steps. First, we compare the correlation between respondents' point forecasts and their implied mean forecast. Second, we compare the elicited distribution of subjective inflation expectations across treatments. Third, we assess differences in perceived and actual length as well as perceived complexity across elicitation methods. Finally, we investigate whether there are differences in how well elicited inflation expectations correlate with planned durable consumption across treatments.

Table 5: Correlation between Point Estimates and Implied Mean Forecasts

	Point Estimate	Implied Mean	Correlation	Confidence Interval
2-Step	5.56	6.19	0.79	[0.72; 0.83]
Endogenous	5.49	6.22	0.68	[0.59; 0.74]
Bins SCE	5.45	5.09	0.51	[0.40; 0.60]
Bins Shift	5.50	6.27	0.85	[0.81; 0.89]

Note: None of the differences in point predictions between treatments are significant at the 10% level. Reported are 95% confidence intervals for the Pearson's correlation coefficients. Intervals are calculated using Fisher's z transformation.

Correlation. Table 5 displays the correlation between respondents' point forecasts and their implied mean forecast split by treatment. Note that point estimates are almost identical across all four treatments, with none of the differences being significant at the 10%-level. Given that respondents provided the point estimate before they were allocated to a treatment, this further supports a successful randomization. In terms of correlation, the Bins Shift treatment achieves the highest correlation of 0.85 between point estimates and implied mean forecasts. This is not surprising given that intervals of the density forecast are anchored around the point estimate in this treatment. For the midpoint treatments, we still observe a high correlation between point estimates and implied mean forecasts. The correlation of 0.78 in the 2-Step treatment is slightly higher than the correlation of 0.68 in the Endogenous treatment. Finally, the Bins SCE treatment shows the

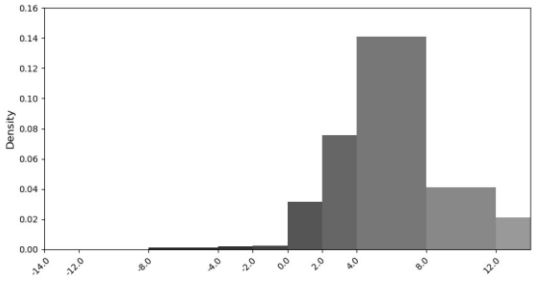
lowest correlation of only 0.51. This finding is interesting for two reasons. On the one hand, the low correlation for the Bins SCE is surprising as the implied mean is closest to the point estimate across all four treatments. On the other hand, this finding highlights one of the core problems of the SCE question format: the fixed survey scale (anchored around zero) is not robust to different inflation regimes. This problem is especially severe in periods of high inflation, where relevant responses may end up being lumped in upper extreme bins (not helping to distinguish between respondents), while still presenting respondents with irrelevant lower extreme bins. If some respondents feel they should allocate at least some probability mass to these lower extreme (deflation) bins, and this is unrelated to their actual beliefs (point forecasts), this will reduce the correlation.

Distribution of Subjective Inflation Expectations. Next, we compare the subjective distribution of inflation expectations across treatments. To allow for direct comparisons, we fit histograms of the average density assigned to the intervals of the Bins SCE format. For the Bins SCE treatment, the responses thus seamlessly translate to the histograms. For the other three treatments, we require a few distributional assumptions. For the midpoint treatments, note that all four ranges (i.e., minimum to 25th-percentile, 25th-percentile to median, median to 75th-percentile, and 75th-percentile to maximum) are equally likely. To convert the data to histograms of the average density of the SCE bins, we assume that all outcomes within a given range are uniformly distributed. For the Bins Shift treatment, we make a similar assumption. In this treatment, specific probabilities are assigned to ten different bins that all vary on the respondent-level (due to different point predictions). We assume that all outcomes within a bin are equally likely and that the open intervals of the shifted bins have twice the width of the adjacent closed interval.[9] Results are displayed in Figure 3 for the midpoint treatments and in Figure 4 for the bin treatments.

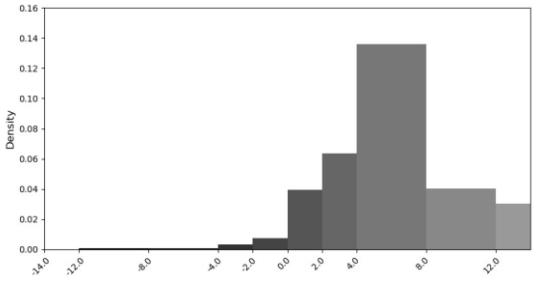
[9] As a consistency measure, we check that the assigned probability mass always adds up to 100%. This ensures that all values have been taken into account and that no double-counting occurred. All treatments meet this condition.

Figure 3: Density Plots for Midpoint Treatments

Panel A: 2-Step Treatment



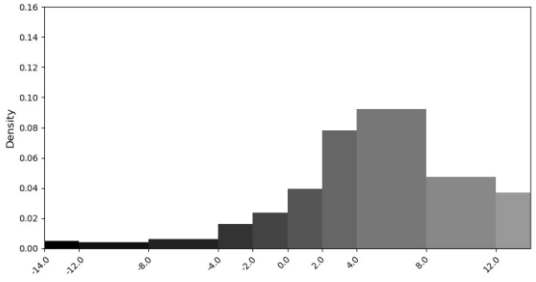
Panel B: Endogenous Treatment



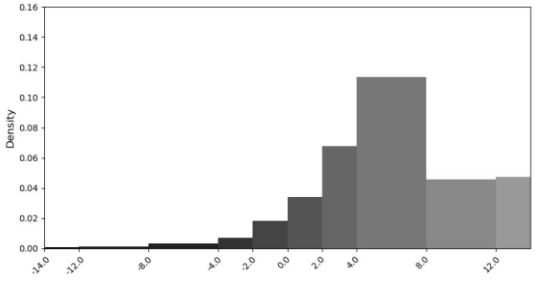
Notes: The figure illustrates histograms of average densities for 12-month ahead inflation expectations elicited through a midpoint method. Panel displays results for the 2-Step treatment and Panel B displays results for the Endogenous treatment.

Figure 4: Density Plots for Bin Treatments

Panel A: Bins SCE



Panel B: Bins Shift



Notes: The figure illustrates histograms of average densities for 12-month ahead inflation expectations elicited through a density forecasts. Panel displays results for the Bins SCE treatment and Panel B displays results for the Bins Shift treatment.

First, note that the histograms of both midpoint treatments look almost identical, which is consistent with our prior results. Most notably, both treatments assign the most probability mass to inflation ranges between 4% and 8% (56.4% and 54.3% for the 2-Step and Endogenous treatment, respectively). At the same time, almost no probability mass is assigned to scenarios that involve deflation. In particular, the average probability mass assigned to deflation is 1.4% in the 2-Step treatment, and 2.8% in the Endogenous treatment. In comparison, both bin treatments also assign most probability mass to inflation ranges between 4% and 8%. However, the total mass assigned to this bin is lower. In the Bins SCE treatment, this amounts to 36.9%, while it amounts to 45.4% in the Bins Shift treatment. The most noticeable difference between both elicitation methods is amount of probability mass that is assigned to deflation. In the Bins SCE treatment, 13.3% probability mass is assigned to deflation and in the Bins Shift treatment 7.1%, and some if it

allocated to substantial deflation rates. This higher amount of probability mass for deflation scenarios is also reflected in the relatively high forecast uncertainty discussed in Section 4.2. Ultimately, the question of whether it is desirable (or realistic) to have a certain amount of probability mass assigned to deflation is difficult to answer. However, when explicitly asked to state the lowest possible rate of inflation that respondents believe to be conceivable over the next year, only 4.9% even consider deflation in their answer. It occurs that the bin format provides an exogenous anchor for the potential variability of the measured variable.

Duration and Difficulty. In a third step, we compare both elicitation methods in terms of perceived difficulty, perceived duration, and actual time taken to complete. These are important features of a method especially when used in consumer surveys. Table 6 summarizes the results. Overall, we find no systematic differences across elicitation methods in terms of perceived difficulty and length. Respondents across all four treatments rate the survey to be rather easy to answer (ranging from 2.38 to 2.56 across treatments with maximum difficulty of 5), with no significant difference between treatments. Similarly, respondents across all treatments believe that the survey was rather fast to complete (ranging from 1.99 to 2.00 across treatments with a maximum perceived length of 5). When we contrast respondents' subjective perception with their actual time taken, we find that the average respondent is well calibrated. Although subjects in the Bins Shift treatment took on average 12 seconds more to complete the survey relative to the other treatments (122 seconds versus 108, 109, and 116 seconds in the 2-Step, Endogenous, and Bins SCE, respectively), none of the differences is statistically significant at the 10%-level. However, it is noteworthy that while both midpoint treatments consist of multiple screens with several questions, the bin treatments only consist of a single density forecast. Despite this difference, actual time taken is almost identical, suggesting that a single density forecast requires substantial time to complete. Finally, there is no relevant attrition in any of the four treatments, with a maximum drop-out rate of 3.8% during the elicitation process for the Bins SCE condition.

Table 6: Difficulty and Duration

	Perceived Difficulty (1 – 5)	Perceived Length (1 – 5)	Time Taken for Elicitation (in seconds)	Attrition
2-Step	2.45 (1.06)	2.00 (0.17)	108.13 (91.23)	5/205
Endogenous	2.38 (1.07)	1.99 (0.14)	109.62 (84.56)	4/209
Bins SCE	2.48 (1.20)	2.00 (0.20)	116.72 (94.70)	8/208
Bins Shift	2.56 (1.11)	1.99 (0.20)	122.00 (108.54)	3/206

Note: Difficulty refers to the self-reported difficulty of the survey ranging from 1 (very easy) to 5 (very difficult). Length refers to the self-reported perceived length of the survey ranging from 1 (very short) to 5 (very long). Attrition refers to respondents who drop out of survey during the elicitation of the full distribution of beliefs. Reported are means and standard deviations (in parentheses). None of the differences between treatments are significant at the 10% level.

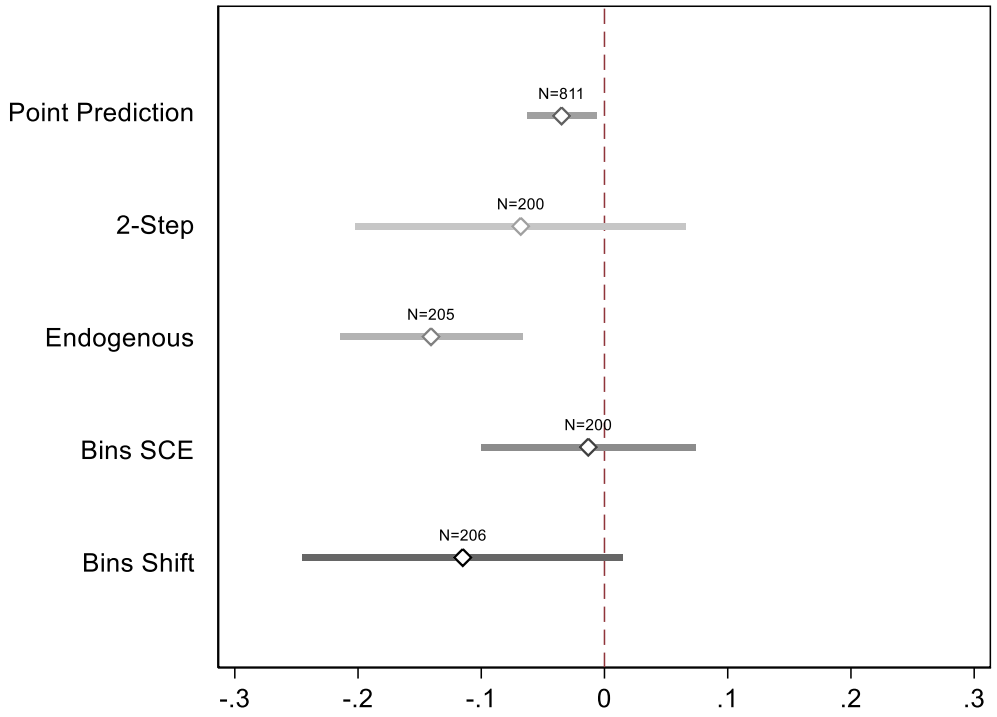
Spending on Durable Goods. Finally, we assess the external validity of the elicited measures, by linking inflation expectations to respondents' consumption decisions. In previous literature, this relationship has been studied to evaluate the extent to which consumers act on their beliefs and whether correlations are consistent with the theoretical relationship as predicted by the consumption Euler equation. As a measure of planned consumption, we use a question from our post-survey questionnaire in which respondents judge the likelihood that their household income for various categories over the next year will either decrease, increase, or stay the same (the question is taken from the Bundesbank Household Survey on Consumer Expectations). In order to capture durables consumption, we focus on the *Major Purchases* category. We regress planned spending on durable goods on inflation expectations and a set of control variables including age, gender, education, and household income using ordered logit regressions.[10] As measure of inflation expectations, we use respondents' point forecasts, as well as the implied mean forecast for each of our four elicitation methods. Coefficient estimates are displayed in Figure 5.

As can be inferred, all coefficient estimates suggest a negative relationship between planned durable consumption and inflation expectations. This relationship is largely consistent with prior studies. Bachmann et al. (2015) and Ichiue and Nishiguchi (2015) find a negative impact of households' inflation expectations on durable consumption. Similarly, Coibion et al. (2023) show that an exogenous shock on inflation expectations leads Dutch households to lower their durable

[10] Results are robust to using OLS regressions instead. Figure C.1 in Appendix C displays the results.

consumption. It is noteworthy that only participants' point forecast as well as the implied mean forecast of the Endogenous treatment are statistically different from zero, while effects for Bins Shift and the Midpoint 2-step treatments and are only directionally consistent but insignificant at the 5%-level. However, when we pool the implied mean forecasts for our two midpoint treatments (as both rely on the same question format), the combined estimate is significant at the 1%-level, suggesting that our method does fairly well in predicting planned consumption. At the same time, however, the coefficient estimate of the mean forecast elicited through the Bins SCE question format is not only weakest in absolute magnitude, but is also not statistically significant at any conventional level. Overall, this reinforces concerns regarding the external validity of elicited inflation expectations using the question format of the SCE.

Figure 5: Inflation Expectations and Durable Consumption



Notes: We regress respondents' propensity to increase (stay the same / decrease) their spending on durable consumption goods 12 months from now on their 12-month ahead inflation expectations. Displayed are coefficient estimates and 95%-confidence intervals using ordered logit regressions. The first column shows results using individuals' point forecasts while the remaining columns display results using implied means elicited in each treatment.

5. Discussion of Practical Issues

In this section, we discuss several practical issues that may become relevant in applications of the proposed method.

Wide range between minimum $\mathbf{b_0}$ and maximum $\mathbf{b_1}$. If respondents indicate a very large interval (b_0, b_1), this does not harm the precise elicitation of the different percentiles as long as the endogenous method is used with a pre-set precision. However, the assumption of uniform distribution within percentiles as made in the calculation of implied means may not be a good approximation in this case. Consider the example in Section 2.2. Possibly the respondent has indicated $b_1 = 20$ because she believed that there is a very small, but positive probability that inflation may run very high over the next year; overall though the true probability that inflation will be above, say 15% is considered basically 0. With her fourth quartile ranging from 12.34% to 20%, we would possibly make a substantial error in the calculation of the implied mean assuming uniform distribution within that bin, given that most probability mass is actually close to its lower bound of 12.34%. The problem is that the quartile is still too wide, providing too little information.

The problem of a large initial range can easily be resolved by (1) making the number of iterations endogenous (giving precise boundaries of the percentiles); and (2) by adding additional layers of subdivision as long as an elicited percentile is still considered too wide. In the above example, splitting the forth quartile by eliciting $p_{87.5}$ would reduce the problem by immediately allocating half of the probability mass of the fourth quartile to values close to 12.34%.

Random choices in later elicitation steps. The initial choices may be rather clear for respondents. However, as we approach indifference in the sequence of choices, it may become more difficult for the respondent to say which interval she considers more likely, and she may revert to answering randomly. This feature is inherent to the method and to the idea of indifference. If the respondent finds it difficult to say which interval she considers more likely, we may set the elicited percentile equal to the current midpoint, possibly adding an endogenous stopping opportunity for the respondent through an "I consider both intervals equally likely" option. If we are concerned that respondents may be too inclined to use such an option, we may instead enforce a choice (as in the experimental implementation). In this case, random choice adds a symmetric noise to the elicited value. In the example of Table 1, if the respondent was indifferent (or "cannot say") in the third step of the elicitation of the median, we would obtain a median of either 11.25% or 13.75% with equal chance, rather than the presumed true value of 12.5%. Researchers need to

decide whether they consider the chance and consequences of such random-choice noise more problematic than the potential overuse of "cannot say" options by respondents who otherwise would be able to provide more fine-grained data.

Midpoints exactly equal to elicited percentile. A related issue concerns the case where a midpoint is exactly equal (or very close) to the elicited percentile. In this case, again, the respondent should indicate indifference directly. If we do not want to offer this possibility, a forced choice approach will then initially move the midpoint away from the true value. However, in each subsequent step, the midpoints will then move back in the direction of the true value of the elicited percentile. The induced error depends on the level of precision chosen by the researcher. Moreover, if there is no systematic bias, e.g., to choose the interval with lower inflation rates in case of indifference, for the whole sample there will be no bias in the elicited values, but only additional noise. Figure A.1 in Appendix A illustrates the induced distortion if the respondent is exactly indifferent in the initial question (as is the case in symmetric distributions), and only few steps are available to mode the median back to its true value (illustration uses 2-step condition). The same effect obtains for any situation in which the respondent considers the two ranges she is asked to compare exactly equally likely.

Error propagation along the elicitation chain. Because of the chained structure of the method, any errors early in the process will propagate through the whole sequence, affecting values elicited later. For example, if the median is assessed with error, the 25%- and 75-quartiles will also be assessed with error. Error propagation cannot be prevented in the method, but it is unclear how severe the problem may be compared to errors in the elicitation using bin-based density forecasts. As discussed in the previous section on indifference, errors can be mitigated by choosing a higher precision, allowing the process to converge back to the true value after an error. Given the chained structure and possibility of error propagation, a higher precision in earlier stages of the process (notably the elicitation of the median) may be warranted to mitigate the problem in later stages.

6. Conclusion

Central banks around the world increasingly rely on measures of subjective inflation expectations collected through large-scale surveys such as the Survey of Consumer Expectations administered by the Federal Reserve Bank of New York, or the Consumer Expectations Survey by the European

Central Bank. Among these surveys, it is common practice to elicit expectations via probabilistic questions in which respondents assign probabilities to several inflation intervals. In recent years, this question format is increasingly subject to criticism.

In this study, we propose a novel method to elicit subjective inflation expectations. Our method builds on a long history of decision theoretic research and presents a potential solution to many of the existing problems. Relative to the probabilistic question format of the SCE, the elicitation process is fully driven by the respondent and thus involves no anchoring on any exogenously provided frames. A direct benefit of this aspect is that the method does not depend on the current state of the economy and works in both low and high inflation regimes, as well as in those that vary over time. It is noteworthy that our results show that anchoring the response scale of the SCE density forecast around respondents' point forecasts (instead of zero) produces similar results to our method: lower incidence of deflation expectations, higher mean forecasts, a stronger correlation with the point predictions, and better external validity. However, this procedure cannot cure one of the main problems inherent to density forecasts: the arbitrary choice of the bin structure. Using the narrow interval widths of the original SCE question in the shifted format only works in relatively low and stable inflation regimes. It may be insufficient to approximate the distribution of subjective inflation expectations in high inflation regimes (with potentially high policy uncertainty). One possible remedy is to not only link the center of the response scale, but also the interval width to participants' point forecast. However, recent research shows that elicited inflation expectations are not invariant to such compression or expansion of the response scale (Becker et al., 2023). Effectively, this severely limits the ability of any bins-based method to compare subjective inflation expectations across time and across countries with different levels of inflation.

We are not alone in proposing to substitute the current probabilistic question format with simpler questions. For instance, Pavlova (2023) compares an approach asking only for the minimum, maximum, and modal level of inflation to the SCE approach. She approximates distributions from these inputs, and shows that they have desirable properties. The simple approach also performs better than the SCE approach in terms of lower incidence of deflation expectations and lower uncertainty. In contrast to this approach of measuring only the minimum, maximum and point forecast, our method allows to directly estimate the full distribution of respondents' subjective expectations, with only few additional binary comparison questions. Altig et al. (2022) suggest an approach, where respondents are asked to select five potential realizations and subsequently assess the probabilities of these events. Similar to our method, this technique does

not force a structure or anchor on respondents' belief, which results in less biased estimates and allows for more flexibility. At the same time, this method is potentially the most cognitively demanding, as it requires not only understanding the concept of probability, but also requires constructing relevant outcome scenarios. A simpler version along these lines is used by Christelis et al. (2020), where respondents are asked to give the support of the distribution (minimum and maximum values), as well as the probability mass to the right of the midpoint. While winning in the simplicity, this approach still requires respondents to understand the concept of probability and makes strong assumptions about underlying distribution[11].

To conclude, our proposed method is based on cognitively simple direct comparisons, does not require respondents to understand numerical values of probabilities (only "more likely than" relationships), and at the same time allows to elicit full subjective probability distributions about macroeconomic outcomes. It does not impose exogenous structure or anchors on respondents' beliefs and thus leads to less bias in estimates comparing to existing benchmarks. Finally, the proposed method is portable in the sense that it can be easily applied to study beliefs about other uncertain outcomes. These include beliefs about macroeconomic events (e.g., stock market returns), personal risks (e.g., job loss, crime victimization, mortality), or future income (e.g., earnings, Social Security benefits).

[11] Another approach is to use indirect measure of inflation expectations or expected income equivalent, as proposed by Hajdini et al (2024). Contrary to that, our method is also applicable for measuring other macroeconomic beliefs.

Appendix:

A. Mapping of latent beliefs on elicited beliefs: Illustration

This section illustrates the mapping of some latent belief distributions on elicited distributions of inflation expectations. We consider the two cases in which respondents' true inflation expectations are governed by a symmetric normal distribution or a skewed normal distribution. Both distributions have a mean of $\mu = 2$ and a standard deviation of $\sigma = 4$. The skewed normal distribution has a skewness of $\alpha = 5$. These values are non-consequential for the drawn conclusions and only serve illustrative purposes.

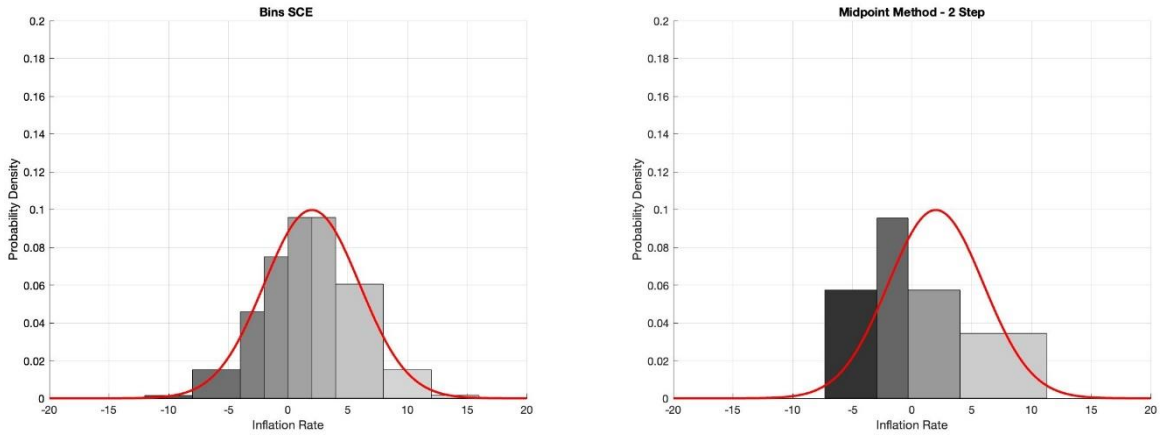
Next, we assume that respondents' inflation expectations are elicited either through a midpoint method or through the SCE density forecast. For the midpoint method, we assume only two steps. As such, the precision of elicited inflation expectations can be understood as a lower bound. We identify the maximum and the minimum as the 99th-percentile and 1st-percentile of the underlying distribution. For the SCE density forecast, we assume that respondents can perfectly match their latent beliefs to the pre-specified bins without inducing a bias. Figure A.1 Panel A illustrates results for the symmetric normal distribution and Panel B illustrates results for the skewed normal distribution.

The normal distribution (Panel A) illustrates the case discussed in Section 5 in which the initial midpoint exactly equals the elicited percentile (here: median). Since we do not permit respondents to indicate indifference directly, the method initially moves away from the true value and subsequently approaches the true value with each step. Since the direction in which the method initially moves away is random, the induced noise is symmetric around the true value (illustrated is the case in which the method initially moves to lower values). Under the assumption that respondents are perfectly capable of expressing their expectations (and that these expectations are well defined), the density forecast approximates the true distribution fairly well.

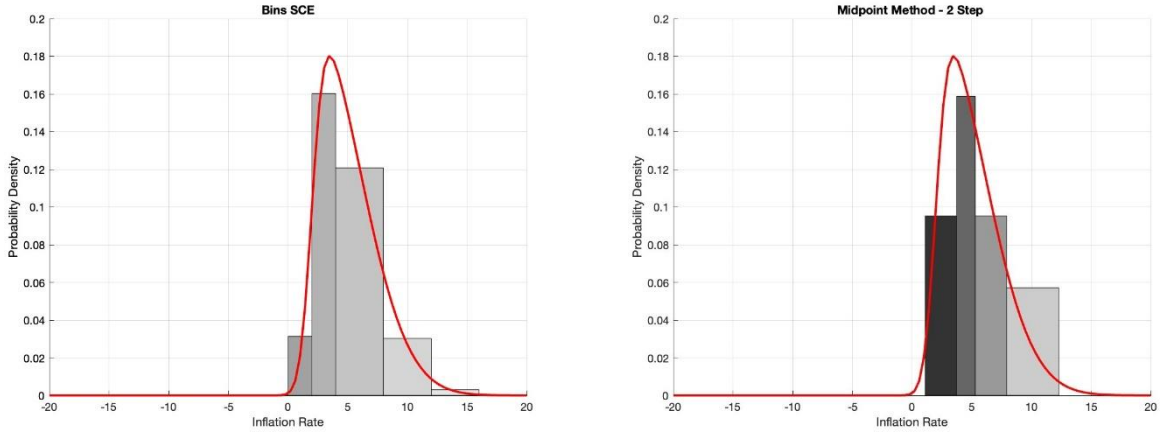
For the skewed distribution, results across both elicitation methods look pretty similar. The example illustrates that the midpoint method with only two steps approximates the true distribution already fairly well as long as the initial midpoint does not equal elicited percentile (which is always the case once we assume that the true distribution is skewed). For the density forecast, notice that respondents would only allocate a substantive amount of probability mass (>1%) to five (out of ten) bins.

Figure A.1: Mapping of latent beliefs on elicited beliefs

Panel A: Symmetric Normal Distribution



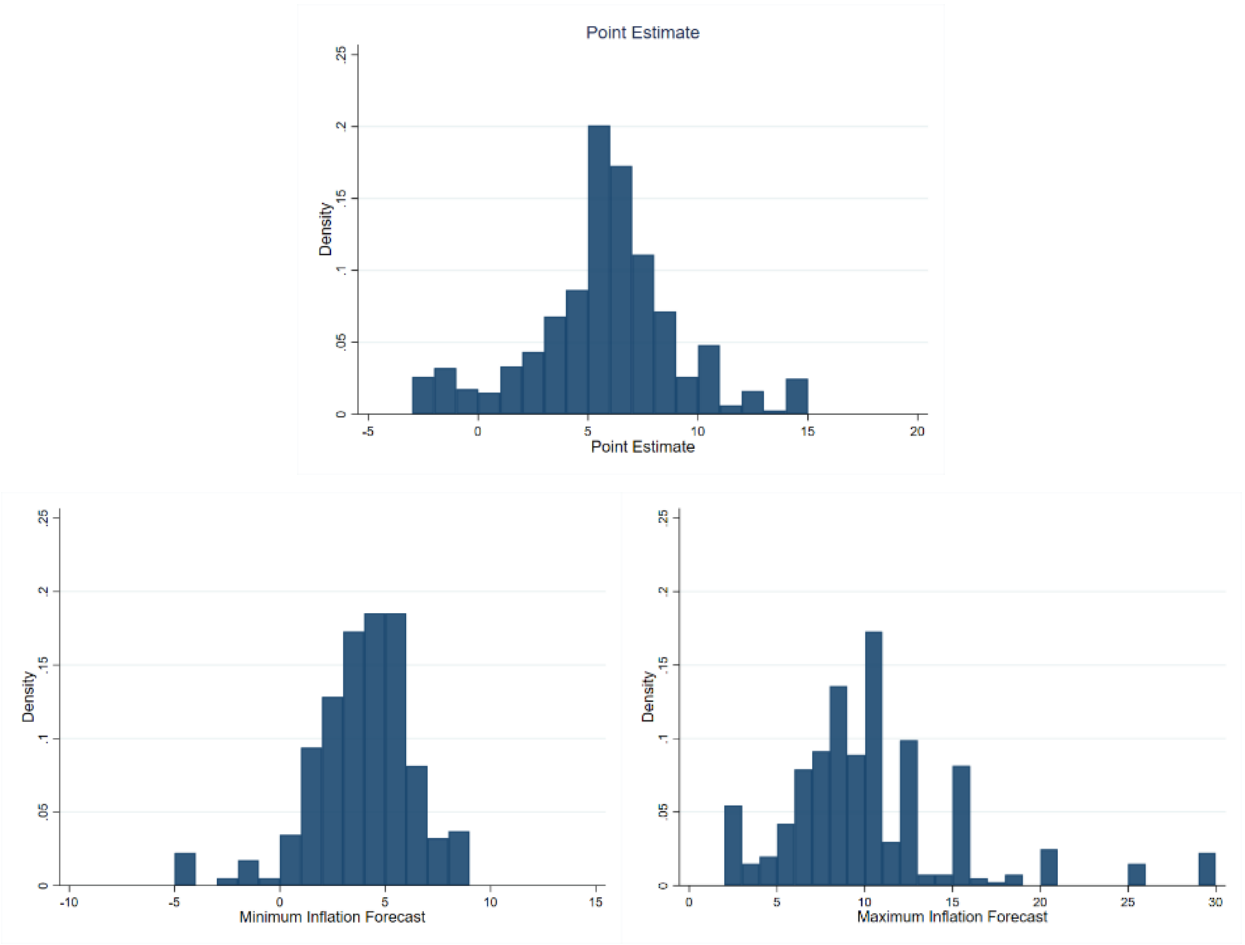
Panel B: Skewed Normal Distribution



Notes: The figure illustrates the mapping of two latent belief distributions on elicited distributions for the midpoint method (right panels) and SCE density forecast (left panels). Panel A displays results for a symmetric normal distribution (mean 2 and standard deviation of 4). Panel B displays results for a skewed normal distribution (mean 2, standard deviation 4, and skewness of 5). The red line represents the overlay of the latent distribution while the grey bars represent the results for each method. For the midpoint method, we approximate the maximum and the minimum as the 99th-percentile and 1st-percentile of the underlying distribution.

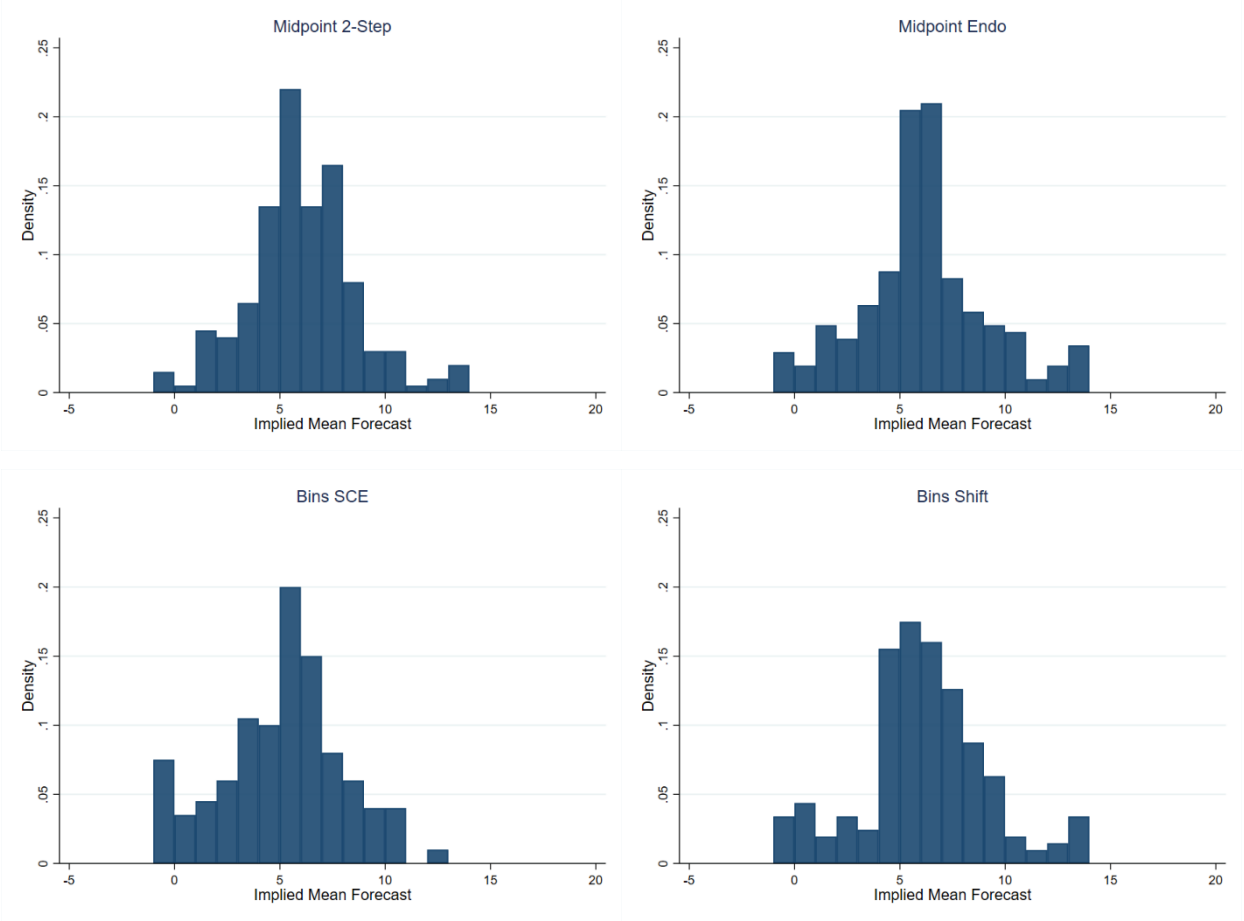
B. Distribution of Point Prediction, Minimum, Maximum and Implied Means

Figure B.1: Histograms of Point Estimate and Minimum/Maximum



Note: Histograms of the distribution of point estimate (all treatments), minimum and maximum value (midpoint treatments). All histograms use 2% and 98% winsorized data.

Figure B.2: Histograms of Implied Means by Treatment



Note: Histograms of the distribution of implied means for each treatment. All histograms use 2% and 98% winsorized data.

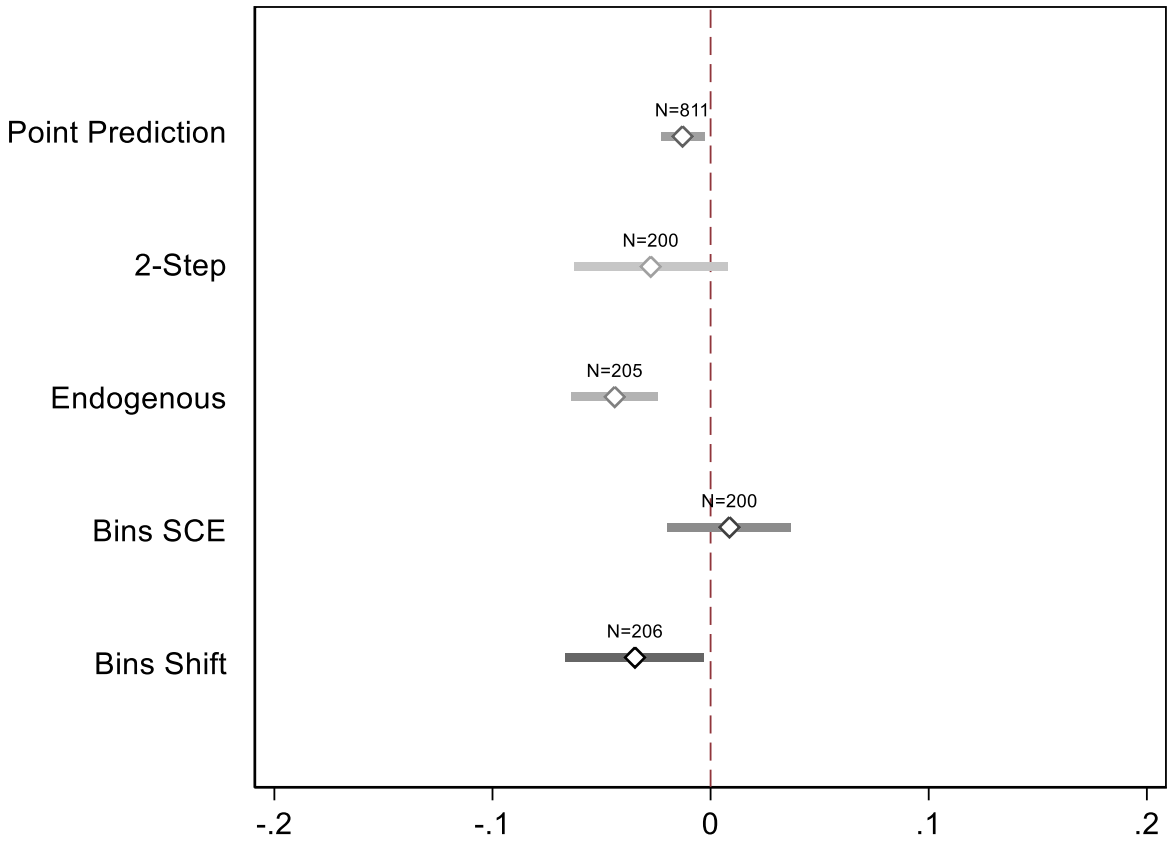
C. Additional Analyses

Table C.1: Range of Inflation Expectations

	(1)	(2)	(3)
Female	-0.639		-0.554
	(-1.12)		(-1.00)
Age	-0.0675***		-0.0672**
	(-2.88)		(-2.48)
Income (log)	-0.372		-0.248
	(-0.78)		(-0.50)
College Degree	0.195		0.183
	(1.32)		(1.22)
Major purchases		-1.017***	-0.805**
		(-2.96)	(-2.22)
Clothing and footwear		-0.297	-0.152
		(-0.56)	(-0.27)
Essential goods		0.777	0.631
		(1.14)	(0.91)
Entertainment recreation		0.248	0.376
		(0.37)	(0.56)
Mobility		-0.377	-0.195
		(-0.64)	(-0.32)
Services		0.795	0.738
		(0.97)	(0.89)
Travel holidays		-0.278	-0.523
		(-0.54)	(-0.92)
Housing costs		0.566	0.370
		(0.88)	(0.59)
Financial reserves		-0.588	-0.725
		(-1.13)	(-1.36)
Constant	11.77***	6.897***	11.40***
	(3.26)	(3.04)	(2.63)
Observations	401	405	401
R^2	0.034	0.032	0.062

Note: Dependent variable is the range of elicited inflation expectations (defined as the difference between the maximum and minimum level of inflation). Independent variables include female (=1 if female and 0 otherwise), age, income (logarithmized), and college degree (=1 if college degree or higher and 0 otherwise). Categories for spending denote whether participants plan to spend less, roughly the same, or more in the following year. Reported are coefficients and t-statistics (in parentheses) using robust standard errors. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Figure C.1: Inflation Expectations and Durable Consumption: OLS



Notes: We regress respondents' propensity to increase (stay the same / decrease) their spending on durable consumption goods 12 months from now on their 12-month ahead inflation expectations. Displayed are coefficient estimates and 95%-confidence intervals using OLS regressions. The first column shows results using individuals' point forecasts while the remaining columns display results using implied means elicited in each treatment.

References

Altig, D., Barrero, J. M., Bloom, N., Davis, S. J., Meyer, B., & Parker, N. (2022). Surveying business uncertainty. *Journal of Econometrics*, 231(1), 282-303.

Andersson, F. N. G., Hjalmarsson, E. & Österholm, P. (2022). Inflation illiteracy – a micro-data analysis. *Sveriges Riksbank Economic Review*, 2022 (2), 106-119.

Armantier, O., Topa, G., Van der Klaauw, W., & Zafar, B. (2017). An overview of the Survey of Consumer Expectations. *Economic Policy Review*, 23(2), 51-72.

Bachmann, R., Berg, T. O., & Sims, E. R. (2015). Inflation expectations and readiness to spend: Cross-sectional evidence. *American Economic Journal: Economic Policy*, 7(1), 1-35.

Baillon, A. (2008). Eliciting subjective probabilities through exchangeable events: An advantage and a limitation. *Decision Analysis*, 5(2), 76-87.

Becker, C., Duersch, P., & Eife, T. (2023). Measuring Inflation Expectations: How the Response Scale Shapes Density Forecasts. Working Paper.

Blinder, A., & Krueger, A. (2004). What Does the Public Know about Economic Policy, and How Does It Know It?. *Brookings Papers on Economic Activity*, 1, 388-393.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88-97.

Christelis, D., Georgarakos D., Japelli, T., & van Rooij, M. (2020) Trust in the Central Bank and Inflation Expectations. *International Journal of Central Banking*, 16 (6), 1-37.

Coibion, O., Gorodnichenko, Y., Kumar, S., & Pedemonte, M. (2020). Inflation expectations as a policy tool?. *Journal of International Economics*, 124, 103297.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., Kenny, G., & Weber, M. (2024). The effect of macroeconomic uncertainty on household spending. *American Economic Review*, 114(3), 645-677.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., & Van Rooij, M. (2023). How does consumption respond to news about inflation? Field evidence from a randomized control trial. *American Economic Journal: Macroeconomics*, 15(3), 109-152.

Conrad, C., Enders, Z., & Glas, A. (2022). The role of information and experience for households' inflation expectations. *European Economic Review*, 143, 104015.

Curtin, R. (2007). Consumer sentiment surveys: worldwide review and assessment. *Journal of business cycle measurement and analysis*, 2007(1), 7-42.

Curtin, R. (2019). Consumer expectations: a new paradigm. *Business Economics*, 54(4), 199-210.

De Bruin, W. B., Manski, C. F., Topa, G., & Van Der Klaauw, W. (2011). Measuring consumer uncertainty about future inflation. *Journal of Applied Econometrics*, 26(3), 454-478.

De Bruin, W. B., Van der Klaauw, W., Topa, G., Downs, J. S., Fischhoff, B., & Armantier, O. (2012). The effect of question wording on consumers' reported inflation expectations. *Journal of Economic Psychology*, 33(4), 749-757.

Delavande, A., & Rohwedder, S. (2008). Eliciting subjective probabilities in Internet surveys. *Public Opinion Quarterly*, 72(5), 866-891.

Dräger, L., & Lamla, M. J. (2024). Consumers' macroeconomic expectations. *Journal of Economic Surveys*, 38(2), 427-451.

D'Acunto, F., Hoang, D., Paloviita, M. and Weber, M., 2023. IQ, expectations, and choice. *The Review of Economic Studies*, 90(5), pp.2292-2325.

Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *The Quarterly Journal of Economics*, 75(4), 670-689.

Gülsen, E. & Kara, H. (2019). Measuring Inflation Uncertainty in Turkey. *Central Bank Review*, 19, 33-43.

Hayo, B., & Méon, P. G. (2023). Measuring household inflation perceptions and expectations: The effect of guided vs non-guided inflation questions. *Journal of Macroeconomics*, 78, 103558.

Hajdini, I., Knotek II, E. S., Leer, J., Pedemonte, M., Rich, R., & Schoenle, R. (2024). Indirect consumer inflation expectations: Theory and evidence. *Journal of Monetary Economics*, 103568.

Hjalmarsson, E. & Österholm, P. (2021). Anchoring in surveys of household expectations. *Economics Letters*, 198, 109687.

Hong, C. S., & Sagi, J. S. (2006). Event exchangeability: Probabilistic sophistication without continuity or monotonicity. *Econometrica*, 74(3), 771-786.

Ichiue, H., & Nishiguchi, S. (2015). Inflation expectations and consumer spending at the zero bound: Micro evidence. *Economic Inquiry*, 53(2), 1086-1107.

Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5), 1329-1376.

Manski, C. F. (2018). Survey measurement of probabilistic macroeconomic expectations: progress and promise. *NBER Macroeconomics Annual*, 32(1), 411-471.

Pavlova, L. (2023). Framing Effects in Consumer Expectation Surveys. Working Paper.

Raiffa, H. 1968. Decision Analysis. Addison-Wesley, London.

Ramsey, F. P. (1931). Truth and probability. R. B. Braithwaite (Ed.). The Foundations of Mathematics and Other Logical Essays. Routledge, London, 156–198.

Spetzler, C. S., & Stael von Holstein, C. A. S. (1975). Exceptional paper—Probability encoding in decision analysis. *Management Science*, 22(3), 340-358.

Weber, M., D'Acunto, F., Gorodnichenko, Y., & Coibion, O. (2022). The subjective inflation expectations of households and firms: Measurement, determinants, and implications. *Journal of Economic Perspectives*, 36(3), 157-184.

Zhao, Y. (2023). Internal consistency of household inflation expectations: Point forecasts vs. density forecasts. *International Journal of Forecasting*, 39(4), 1713-1735.

Online Appendix: Survey Instructions

Welcome to our Survey

Dear Participant,

Thank you for taking the time to complete this survey. Overall, the survey will roughly take 10 minutes to complete. Please make sure that you have the time. Your answers will be treated confidentially and no personal information will be required. All records will be assigned a subject code number for anonymity, and the anonymized data set will be published and available to other researchers. It would be a great help to us if you answer the questions as carefully as possible.

In this survey, we would like you to assess how the economic situation in your country has changed over time and what you are currently expecting for the future. You do not need any economic background to answer the questionnaire. Most importantly, there are no right or wrong answers, we are simply interested in your personal opinion.

To navigate through the survey, please only use the "Continue" button at the bottom of the screen. Please do not use the buttons in your browser or the back button on your phone or tablet. During the survey, there is no option to go back to questions previously answered. As such, it is important to read the instructions carefully. There will be questions which check whether you read the instructions.

Participation in this study is entirely voluntary and you may withdraw at any time. At the end of the survey, you will get a completion code which you need to enter on Prolific. By entering your Prolific ID and clicking the continue button below, you acknowledge that you have read and understood this consent form and agree to participate in the study.

Please enter your Prolific ID:

_____

Inflation Point Forecast [all Treatments]:

What do you expect the rate of inflation to be over the next twelve months?

Note: Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. If you expect the price level to increase, please enter a positive percentage value. If you expect the price level to decrease, please enter a negative percentage value. Please enter a value with a maximum of one decimal place.

[ ] %

Next

Minimum and Maximum Level of Inflation [only Midpoint Treatments]:

Inflation Expectations

In your opinion, what is the minimum rate of inflation over the next twelve months for which you think that there is absolutely no chance the true rate of inflation will be lower?

Note: The aim of this question is to determine the absolute lowest possible rate of inflation (or negative numbers if you expect deflation) that you would consider realistic over the next twelve months.

[ ] %

In your opinion, what is the maximum rate of inflation over the next twelve months for which you think that there is absolutely no chance the true rate of inflation will be higher?

Note: The aim of this question is to determine the absolute highest possible rate of inflation (or negative numbers if you expect deflation) that you would consider realistic over the next twelve months.

[ ] %

Next

Confirmation Screen [only Midpoint Treatments]:

Please confirm

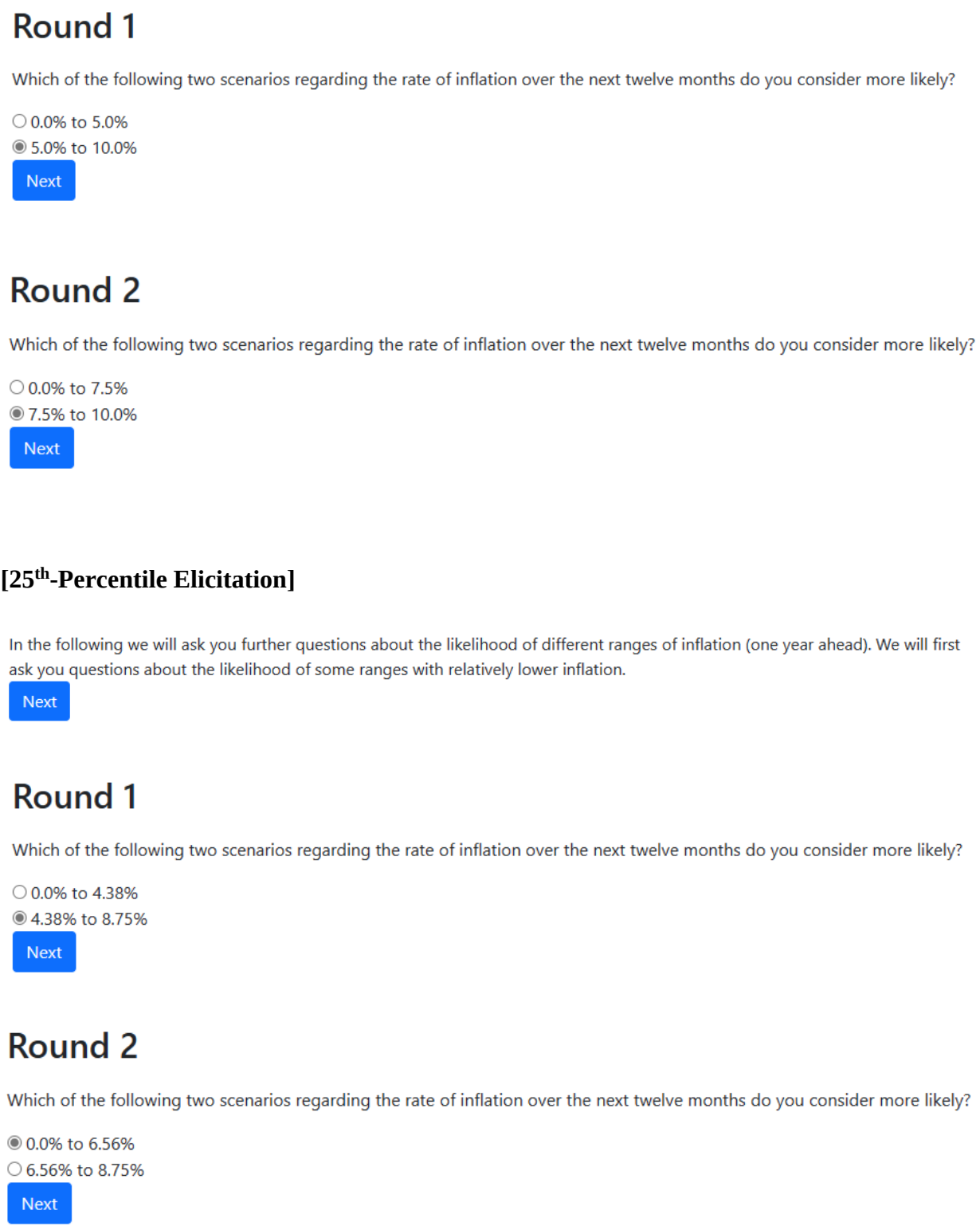
Your expectation for the minimum possible inflation is 0.0%, and your expectation for the maximum possible inflation is 10.0%. Would you like to continue? By clicking no, you will have the option to return to the prior page and edit your answers.

- ○ Yes
- ○ No

Next

Screenshots Midpoint Process [only Midpoint Treatments]:

[Median Elicitation]



[25th-Percentile Elicitation]

[75th-Percentile Elicitation]

Now we would like to ask you some questions about the likelihood of some ranges with relatively higher inflation.

Next

Round 1

Which of the following two scenarios regarding the rate of inflation over the next twelve months do you consider more likely?

- (●) 8.75% to 9.38%
- ( ) 9.38% to 10.0%

Next

Screenshots Bins SCE:

We would like you to think about the different things that may happen to inflation over the next 12 months. We realize that this question may take a little more effort.

In your view, what would you say is the percent chance that, over the next 12 months...

the rate of inflation will be 12% or higher		**percent chance**
the rate of inflation will be between 8% and 12%		**percent chance**
the rate of inflation will be between 4% and 8%		**percent chance**
the rate of inflation will be between 2% and 4%		**percent chance**
the rate of inflation will be between 0% and 2%		**percent chance**
the rate of deflation (opposite of inflation) will be between 0% and 2%		**percent chance**
the rate of deflation (opposite of inflation) will be between 2% and 4%		**percent chance**
the rate of deflation (opposite of inflation) will be between 4% and 8%		**percent chance**
the rate of deflation (opposite of inflation) will be between 8% and 12%		**percent chance**
the rate of deflation (opposite of inflation) will be 12% or higher		**percent chance**
Total	**0**	**percent chance**

Next

Screenshots Bins Shift (using a Point Forecast of 5%):

We would like you to think about the different things that may happen to inflation over the next 12 months. We realize that this question may take a little more effort.

In your view, what would you say is the percent chance that, over the next 12 months...

Note: Negative percentage values imply deflation, i.e. that you expect the price level to decrease

the rate of inflation will be 17.0% or higher		**percent chance**
the rate of inflation will be between 13.0% and 17.0%		**percent chance**
the rate of inflation will be between 9.0% and 13.0%		**percent chance**
the rate of inflation will be between 7.0% and 9.0%		**percent chance**
the rate of inflation will be between 5.0% and 7.0%		**percent chance**
the rate of inflation will be between 5.0% and 3.0%		**percent chance**
the rate of inflation will be between 3.0% and 1.0%		**percent chance**
the rate of inflation will be between 1.0% and -3.0%		**percent chance**
the rate of inflation will be between -3.0% and -7.0%		**percent chance**
the rate of inflation will be -7.0% or lower		**percent chance**
Total	**0**	**percent chance**

Next

Length and Difficulty:

Finally, we would like you to answer a few questions regarding your assessment of the survey.

How easy or hard was it to answer the questions? Please select one answer.

[Very difficult]

[Somewhat difficult]

[Partly interesting/partly uninteresting]

[Somewhat easy]

[Very easy]

How did you find the length of the survey?

[Far too long]

[Somewhat too Long]

[Just right]

[Somewhat too short]

[Far too short]

Demographics

Gender:

[Female]

[Male]

[Diverse]

[Prefer not to say]

Age:

[Number Field] Years

Income:

What is the total monthly net income of your household?

Note: This refers to the total amount, comprising wages, salary, income from self-employment and pensions, in each case after deducting taxes and Social Security contributions. In this amount, please include any income received through public aid, earnings from rents and leases, housing allowance, child benefits, and any other source of income.

Educational Attainment:

What is your highest educational attainment?

[Prefer not to answer]

[Less than high school diploma]

[High school diploma]

[Some college (no degree)]

[Associates degree occupational]

[Associates degree academic]

[Bachelor's degree]

[Master's degree]

[Professional degree]

[Doctoral degree]

Income Change:

In your opinion, how likely is it that your household's average monthly net income will change as follows over the next twelve months?

Note: The aim of this question is to determine how likely you think it is that something specific will happen in the future. You can rate the likelihood on a scale from 0 to 100, with 0 meaning that an event is completely unlikely and 100 meaning that you are absolutely certain it will happen. Use values between the two extremes to moderate the strength of your opinion. Please note that your answers to the categories have to add up to 100.

Decrease [Number Field] percent chance

Stay about the same [Number Field] percent chance

Incease [Number Field] percent chance

Household Spending:

If you think back to last month, roughly how much did you spend in GBP on the following items?

Major purchases (e.g. car, furniture, electrical appliances, etc.)

[Number Field]

Essential goods (e.g. food and beverages, non-food items such as cleaning products or similar)

[Number Field]

Clothing and footwear

[Number Field]

Entertainment/recreation (e.g. restaurant visits, cultural events, gym)

[Number Field]

Mobility (e.g. fuel, car loans and running costs, bus and train tickets)

[Number Field]

Services (e.g. hairdresser, childcare, medical costs)

[Number Field]

Travel, holidays

[Number Field]

Housing costs (e.g. rent, mortgage, ancillary costs)

[Number Field]

Financial reserves

[Number Field]

Household Spending Change:

And are you likely to spend more or less on the following items over the coming twelve months than in the last year?

Please indicate:

1: I plan to spend less.

2: I plan to spend about the same.

3: I plan to spend more.

Please select an answer in each row.

Major Purchases

Essential Goods

Clothing And Footwear

Entertainment Recreation

Mobility

Services

Travel Holidays

Housing Costs

Financial Reserves
