Discussion Paper

Deutsche Bundesbank
No 25/2025

A choice-based approach to the measurement of inflation expectations

Olga Goldfayn-Frank
(Deutsche Bundesbank)

Pascal Kieren
(Heidelberg University)

Stefan Trautmann
(Heidelberg University)

Discussion Papers represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or the Eurosystem.

Editorial Board:	Daniel Foos
	Stephan Jank
	Thomas Kick
	Martin Kliem
	Malte Knüppel
	Christoph Memmel
	Hannah Paule-Paludkiewicz

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

Reproduction permitted only if source is stated.

ISBN 978–3–98848–046–0
ISSN 2941–7503

Deutsche Bundesbank Discussion Paper No 25/2025

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*

30 March 2025

Abstract: Economists widely rely 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 the full distribution of inflation expectations, which is rooted in decision theory and can be implemented in standard surveys. In two large surveys and one laboratory experiment, we demonstrate that it leads to well-defined expectations that fulfil both subjective and objective quality criteria. The method is neither perceived as more difficult nor does it take more time to complete compared to the current standard. In contrast to density forecasts, the method is robust to differences in the state of the economy and thus allows comparisons across time and across countries. The method is portable and can be applied to elicit different macroeconomic expectations.

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 the editors Boragan Aruoba and Luminita Stevens, an anonymous referee, as well as Michael Weber, Pär Österholm, Bernd Hayo, Willibert van der Klaaw, Giorgio Toppa, and seminar participants at the Bundesbank, Czech National Bank, De Nederlandsche Bank, the European Central Bank, the Central Bank of Canada, the University of Heidelberg, Indiana University, the University of Sydney, and the U.S. Securities and Exchange Commission. We thank Christoph Becker, 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. A supplementary online appendix is available at https://www.bundesbank.de/dp-en/2025-25-appendix.

1. Introduction

Households' and firms' subjective inflation expectations play a central role in macroeconomic models and are increasingly important for monetary policy authorities as both indicator and a policy tool (Coibion et al., 2022). The question how to elicit such beliefs from the public is therefore an essential issue. In his pioneering work, Manski (2004) highlighted the importance of eliciting full subjective probability distributions. Since then, subjective probability distributions have been used to go beyond measures of central tendency (Engelberg et al., 2009; Potter et al., 2017; Manski, 2018; Curtin, 2019), and to capture the economic consequences of uncertainty in individuals' forecasts (Coibion et al., 2024). In particular, a probabilistic question format has gained popularity in which respondents are shown several intervals (or *'bins'*) and are then asked to attach probabilities to the intervals according to their beliefs about the future rate of inflation. The question format was developed and implemented in the Survey of Consumer Expectations (SCE) by the Federal Reserve Bank of New York and has since been implemented by several other major central banks.[1]

The recent surge in inflation has made evident not only the importance of the data on subjective probabilistic beliefs to gauge individual-level uncertainty, but also exposed the major challenge of the existing measurement. The primary concern revolves around the bin structure: the narrower bandwidths around zero than at the extreme ranges may induce survey participants to perceive that values close to zero are considered more likely by the designers of the survey, which may bias responses towards zero (Weber et al., 2022). Another concern is that density forecasts are cognitively demanding for many respondents, which leads to higher dropout rates that may ultimately bias inference (Guiso et al., 2002; D'Acunto et al., 2020). Additionally, recent research

[1] Dräger and Lamla (2024) provide an overview of available survey data and issues regarding the measurement of macroeconomic expectations.

suggests that the elicited inflation expectations are neither robust to changes in the wording of the question (e.g., Bruine de Bruin et al., 2012, Manski, 2018; Coibion et al., 2020) nor to changes in the response scale (Becker et al., 2023). During periods of high and volatile inflation these shortcomings are even more pronounced, as responses may end up being lumped in extreme (open) bins, which effectively renders them uninformative. These issues have no easy solution, because adjusting the size and value of bins across survey waves makes it difficult to compare survey responses over time and across countries.

One approach to these problems has been to develop and fit flexible subjective distributions to reported bin probabilities. This aims to preserve the available time-series of data on long standing panels in the presence of framing issues (e.g., Engelberg et al., 2009; Armantier et al., 2017; Ryngaert 2023). However, for new surveys and ad hoc research, a method that avoids these biases and that is less sensitive to variation in inflation rates across time and across countries may be desirable. We propose such a method to elicit individual's subjective probability distribution over expected inflation via surveys, avoiding several of the problems associated with density forecasts. We build on prior research in decision theory by Baillon (2008) to design a simple binary choice-based procedure to elicit the distribution of respondents' expectations. The method is based on a bisection process to partition the state space into equally likely subevents. The method starts by asking the respondent for a minimum and maximum level of inflation for which they think that there is "almost no chance" that actual inflation will lie outside the interval. This way the method avoids inducing external anchors. Once this interval is established, respondents answer a series of binary choices that follow a strict algorithm from which relevant percentiles of respondents' subjective probability can be inferred.

To validate this method in the field, we conduct two surveys, one in the U.K. and one in the U.S. In the U.K. survey, we focus on testing the feasibility of the method by eliciting expectations using one of four protocols. The first two elicitation protocols are based on two variations of our proposed method, which we refer to as '*Midpoint method*' and differ in the precision of the elicited values. The other two protocols are based on the probabilistic question format, which is either anchored around zero (as in the SCE) or shifted to be anchored around respondents' point forecasts (as recently used in surveys of countries with high and volatile inflation, e.g. Central Bank of Turkey; Gülsen and Kara, 2019). In the U.S. survey, we consider two elicitation formats (the Midpoint method and the probabilistic question format of the SCE) and include a randomized controlled trial (RCT) in which we randomly allocate survey respondents into five groups with different information treatments to generate exogenous variation in respondents' inflation expectations. We use this variation to study the causal effects of the resulting change in expectations, as measured by the two different elicitation formats, on spending.

Our results show that there are notable differences in the distribution of subjective inflation expectations depending on the elicitation method. The Midpoint method elicits inflation expectations with substantially higher implied means and higher disagreement than the SCE probabilistic question format, but similar implied means and disagreement as the shifted probabilistic format. The Midpoint method also delivers significantly lower forecast uncertainty relative to both methods that rely on density forecasts.

In a second step, we assess the consistency and validity of elicited inflation expectations. First, 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. Second, when inflation expectations are elicited through a midpoint method, only 2% to 4% of the total probability mass (depending on the survey) is allocated to deflation scenarios. In contrast, when inflation expectations are elicited through a probabilistic question format, 10% to 17% of the total probability mass is allocated to deflation scenarios. Given that the current rate of inflation at the time of both surveys was above target, it appears that the probabilistic question format induces respondents to allocate probability mass to deflation scenarios they would otherwise not consider.

Since it is difficult to judge whether these differences indicate a better performance of the Bins or the Midpoint method, we added two additional quality checks in the U.S. survey. First, we confront respondents with statistics obtained from their own answers (without telling them where the statistics come from). We calculate in real-time the implied median and the implied deflation probability from respondents' answers and then let them judge whether they broadly agree with the number, or whether they believe the number should higher or lower. For both statistics, we find that that a substantially larger share of respondents in the Midpoint treatment agrees with statements derived from their own previous answers. In contrast, the Bins SCE method seems to elicit too large probabilities of deflation, which then biases the implied median downward and leads to overall lower agreement by the respondents.

Second, we assess the causal effects of how well elicited inflation expectations map to respondents' consumption choices by making use of the RCT design. We measure durable consumption intentions using three different items that have been used by prior studies and in which a positive relation to subjective inflation expectations is expected (Duca-Radu et al., 2021; D'Acunto et al., 2022). When post-treatment inflation expectations are elicited using the Midpoint method, we find a positive and statistically significant effect on all three durable consumption categories. In contrast, when inflation expectations are elicited using the SCE density forecasts we find effects that are often insignificant and of different sign across categories.

In a final test, we benchmark the methods against an objective underlying distribution. The main concern is that the true beliefs of respondents are unobserved, and therefore it is difficult to draw conclusions on how well they have been measured. To circumvent this, we designed an experiment in which participants can repeatedly sample (positive and negative) numbers from one of two underlying distributions and afterwards report their beliefs regarding the sample using either the SCE bins or the Midpoint method. We find that the Midpoint method elicits beliefs that better approximate both the mean and the standard deviation of the underlying distribution. In terms of the probability of a negative number (akin to deflation probability), we find that the Midpoint method underestimates this probability, while the Bins SCE method strongly overestimates it, driven by the inappropriate use of "negative" bins. This finding is consistent with our survey evidence on the use of deflation bins. The severity of the biases in terms of deflation probability in surveys thus depends on the underlying distribution. If a country faces little to no deflation, the bias in the Midpoint method will be rather small while the bias of the bin method can be sizable, as documented in our two surveys.

We contribute to an ongoing literature that aims to elicit subjective inflation expectations through surveys. Although prior studies have shown that consumers are generally able to provide meaningful probability distributions which are consistent with their point predictions (Bruine 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: It is robust to differences in the state of the economy and therefore allows comparison across time with potentially very different levels of inflation (e.g., in typical survey panels), but also across countries. Additionally, the method requires neither reference to the concept of probability nor a direct judgement, but only simple binary comparisons.[2] It is thus cognitively easier and less prone to anchoring than allocating probability levels to different inflation bins.[3] Finally, the method is portable in the sense that it can

[2] A related literature investigates how literate and informed consumers actually 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). Although it only requires simpler binary comparisons, the midpoint method still implicitly assumes that consumers are informed enough to make these judgments.

[3] Pre-defined answer options provide an anchor for answers (Hjalmarsson and Österholm, 2021), and may also induce respondents to provide an answer although 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.

be applied to elicit different macroeconomic beliefs.

We are not alone in proposing potential alternatives to the current probabilistic question format. For instance, Pavlova (2025) 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. Her approach also elicits lower incidence of deflation expectations and lower uncertainty than the SCE bins. In contrast to this approach of measuring only the minimum, maximum and mode, 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 Guiso et al. (2002) and more recently by Christelis et al. (2020). Their 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. In contrast to the Midpoint method proposed in the current paper, this approach still requires respondents to understand the concept of probability and makes strong assumptions about the underlying distribution.[4]

[4] Another approach is to use an 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.

2. The Midpoint Method

2.1 Algorithm Description

The midpoint method we use in this study is based on Baillon (2008) and provides a choice-based implementation of a bisection process to partition the state space into 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. Chew and Sagi (2006) formally derive the existence of probabilistic beliefs from this concept.

The basic idea is that two events are exchangeable for an agent when she is indifferent to permutations of their outcomes. Such events are thus revealed to be equally likely. At first, the method elicits two complementary events, $\underline{E}$ and $\overline{E}$. Event $\underline{E}$ involves outcomes that range from some minimum (or theoretically $-\infty$) to the median (p_{50}). Event $\overline{E}$ 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 again 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 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. 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 our studies, in principle the process can be continued by further splitting up (some of the) elicited quartiles.

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 decide which of two 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) is half the distance between b_0 and b_1: $b_0 + \frac{b_1 - b_0}{2}$. Consequently, the first question asks subjects whether they consider the event $\left[b_0, b_0 + \frac{b_1 - b_0}{2}\right]$ or the event $\left(b_0 + \frac{b_1 - b_0}{2}, b_1\right]$ more likely. 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, 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, 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 inflation rates is more likely to occur, the next midpoint is determined as half of the distance between the current midpoint and the next lower previous midpoint (or the minimum if there is no lower midpoint); whenever subjects think that the event with higher inflation rates is more likely to occur, the next midpoint is determined as half of the distance between the current midpoint and the next higher previous midpoint (or the maximum if there is no higher midpoint). Intuitively, whenever an event is considered 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 our surveys in the U.K. and the U.S., 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 (approximately) equally likely, thus representing the median of the subjective probability distribution. 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.

2.2 Example

Table 1 gives an illustration of the bisection process. A hypothetical survey respondent indicates that she expects the lowest possible rate of inflation next year to be 0% (b_0) and the highest possible rate to be 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), implying the second midpoint of $\frac{20\%+10\%}{2} = 15\%$. Based on the two events that follow from the second midpoint, the respondent prefers option A, which leads to the next midpoint of $\frac{15\%+10\%}{2} = 12.5\%$. 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 equals $\frac{12.5\%+10\%}{2} = 11.25\%$. It satisfies the precision criterion as it is less than 1.5% different from the previous midpoint. This concludes the first bisection process, and the final midpoint of 11.25% serves as the median of the inferred subjective probability distribution.

Table 1: Illustration of the Bisection Process

Searched percentile	*Option A*	*Option B*	*Choice*	*Implication*
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 process continues with the elicitation of the 25th-percentile. The first midpoint of $\frac{0\%+11.25\%}{2} = 5.63\%$ is calculated from the minimum and the median. As the respondent prefers option B, the next midpoint is $\frac{5.63\%+11.25\%}{2} = 8.44$. From the next set of options, the respondent prefers A, implying the next midpoint of $\frac{5.63\%+8.44\%}{2} = 7.03\%$, which already satisfies the desired level of precision. It serves as the 25th-percentile. Finally, the process concludes with the elicitation of the 75th-percentile in the same way.

Online Appendix A discusses several practical issues when implementing the method and how to handle them: a wide range between the minimum and maximum; random choices in later steps of the process; a midpoint that is exactly equal to the percentile that is elicited; and the problem of error propagation. It also provides further illustrations by showing the mapping of two exemplary latent belief distributions on elicited distributions using the Midpoint method vs. the SCE Bins method.

3. Survey Design and Data Collection

We conducted two surveys in which we elicit respondents' full distribution of one-year ahead CPI inflation expectations using different elicitation methods. The U.K. survey aimed to test the feasibility of the Midpoint method, while the U.S. added further tests to gauge its performance relative to the Bins SCE approach.

U.K. Survey. The first survey was conducted in September 2023 in the U.K., when the current rate of inflation was 6.3% (a decline from 8.8% in September 2022). First, all survey respondents were 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). Then respondents were randomly allocated to one of the four treatments, using four different elicitation methods in a between-subject design.

The first two treatments labeled "*Midpoint Endogenous*" and "*Midpoint 2-Step*" elicit inflation expectations using the Midpoint method as described in Section 2. 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 level of precision of 1.5% is reached. In the "*Midpoint 2-Step*" treatment, the maximum number of steps for each percentile is set to two. That is, the process stops after two steps, even if the predefined level of precision of 1.5% is not reached yet.

In the other two treatments labeled "*Bins SCE*" and "*Bins Shift*", respondents were asked to provide a density forecast for 12-month ahead inflation. The "*Bins SCE*" treatment uses the standard response scale of the SCE. The symmetric scale has ten intervals, is centered at zero and has two open outer intervals. The closed intervals cover the range from –12% to 12%. Respondents are asked to assign a probability mass of exactly 100% over the ten intervals. In the "*Bins Shift*" treatment, the scale is individually centered at a respondent's previously elicited point forecast instead of zero. All outcomes of the SCE scale are thus shifted by the respondent's point forecast. Besides the different centering, the scale is identical to the regular scale of the SCE. Online Appendix B provides screenshots of the two formats.

Immediately after the elicitation of the inflation expectation, the survey included two questions that assessed the respondents' perceived length and difficulty of the elicitation process, measured on a 5-point Likert scale. These questions test 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 was also recorded. The survey concludes with a questionnaire on demographics and financial variables.

U.S. Survey. The second survey was conducted in November 2024 in the U.S. The latest available CPI inflation release for October 2024 was 2.6% as reported by Bureau of Labor Statistics. The survey follows a similar structure as the U.K. survey, with a few relevant differences discussed next. The survey focuses on two elicitation formats, the *"Midpoint Endogenous"* and the *"Bins SCE"* treatment, but includes an RCT randomly allocating survey respondents into five groups with different information to make causal inferences about the effect of measured inflation expectations on consumption decisions. Before the RCT, the inflation point forecast is collected, while after the RCT, inflation expectations are measured using one of the two elicitations methods. This allows us to evaluate for each elicitation method the effect of the information treatments on expectations and on the subsequent transmission to consumption choices. The information treatments are similar to the ones used in existing studies to exogenously move households' inflation expectations and make causal inferences (see Kumar et al., 2023; Coibion et al., 2024; Georgarakos et al., 2024).

The first group serves as a control group and did not receive any information. The second group was informed about the average professional forecast for inflation in the U.S. (Mean

Treatment). The third group was informed about the average professional forecast and the difference between the lowest and highest predictions among professional forecasters (Mean + Uncertainty Treatment). The fourth group was informed about professional forecasters' perceived lower bound of inflation over the next year (Low Treatment), and the fifth group was informed about the professional forecasters' perceived upper bound (High Treatment). Information provided was based on the most recent data from the Survey of Professional Forecasters from the Federal Reserve Bank of Philadelphia. The exact wording can be found in Online Appendix B. Since we randomly assign respondents to one of five information treatments (including the control group) and to one of two elicitation methods, we have a total of 10 different conditions.

For the U.S. survey, we developed two quality checks to assess how well the elicitation methods identify respondents' 'true' subjective expectations. These questions follow after the elicitation of the full distribution of respondents' inflation expectations, from which we calculate in real-time the implied median and the implied deflation probability for each respondent. We then ask participants the following two questions:

Median: *Would you agree that it is approximately equally likely that inflation in 12 months is higher or lower than [implied median from own answers to Midpoint or Bin method]%?*

Deflation: *Would you agree that that the probability of there being deflation over the next 12 months is about [implied deflation probability from own answers to Midpoint or Bin method]%?*

For each question, participants could either agree that the statement is 'about right' or state that inflation (resp. deflation probability) should be 'rather higher' or 'rather lower'. Note that while these questions directly confront participants with statistics obtained from their own answers, we do not tell them where the numbers come from in order to avoid priming or demand effects. The survey then proceeds with a questionnaire on socio-demographic and financial variables.

The surveys were distributed via the crowdsourcing platform Prolific, which allows to collect stratified samples of the U.K. and U.S. population. 811 respondents completed the U.K. survey (4 conditions), and 4,018 respondents completed the U.S. survey (10 conditions). Respondents were paid a fixed amount of £2 for completing the survey. On average, it took respondents 6:03 and 6:46 minutes to finish the U.K. and U.S. surveys, respectively. This implies an average hourly wage of £19.83 and £17.73, which is well above the average hourly earnings on Prolific and double (triple) the minimum wage in the U.K. (U.S.).[5]

4. Results

4.1. Survey Summary Statistics

Table 2 presents summary statistics for the four elicitation treatments used in the U.K. survey (left panel) and for the two elicitation treatments used in the U.S. survey (right panel).

U.K. Survey. Participants expect the rate of inflation over the next 12 months to be on average 5.51%, slightly lower than the current rate of 6.3% at that time. Histograms of the point prediction, minimum, maximum and implied means for each elicitation method are in Online Appendix C.

Respondents who provide density forecasts in the Bins SCE treatment use on average 6.01 out of 10 possible bins, significantly less than the 6.69 in the Bins Shift treatment ($p < 0.01$; $t = 2.66$, two-sided t-test). The higher number of bins used in the Bins Shift treatment is not unexpected. By anchoring the response scale around their point forecast, respondents have access to a wider range of "sensible" intervals closer to their expectations. The number of bins used is substantially larger than in some other studies (e.g., Delavande and Rohwedder, 2008), probably reflecting the higher

[5] The surveys were computerized using oTree (Chen et al., 2016). The design, sample size, analyses are all pre-registered (https://aspredicted.org/2T2_7DL (U.K. survey) and https://aspredicted.org/fnpb-5cb2.pdf (U.S. survey)). The study received ethics-approval of the Institutional Review Board of the Faculty of Economics and Social Sciences at Heidelberg University (reference number FESS-HD-2023-002).

uncertainty in the inflation environment at the time.

Table 2: Inflation Measurement Summary Statistics

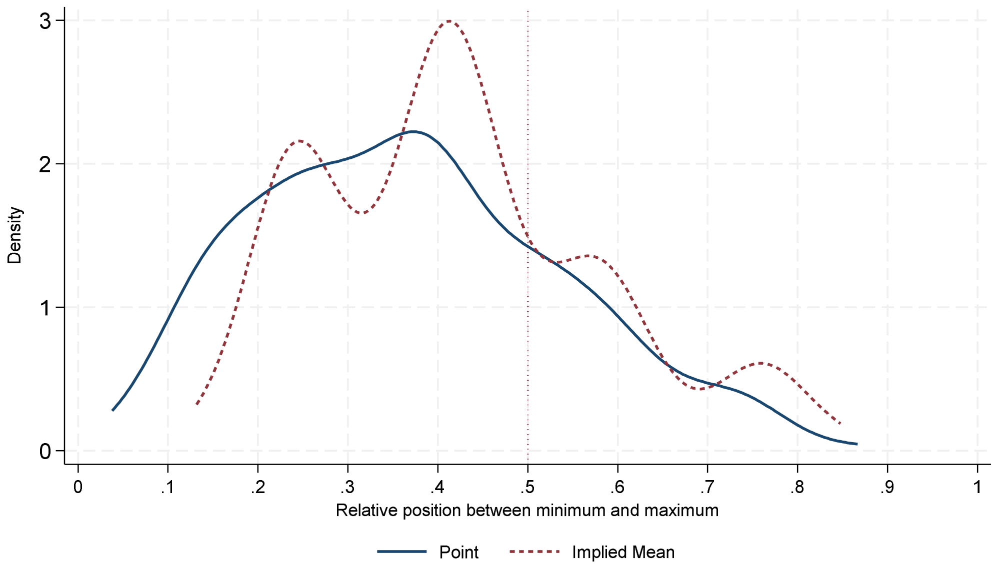
	U.K. Sample		U.S. Sample	
	Mean	**P10 / P50 / P90**	**Mean**	**P10 / P50 / P90**
Point Forecast	5.51	1 / 5.5 / 9.2	4.33	0 / 3 / 9
Bin Treatments:				
#Bins (SCE)	6.01	3 / 5 / 10	5.96	3 / 5 / 10
#Bins (Shift)	6.69	3 / 7 / 10	-	-
Midpoint Treatments:				
Minimum (2-Step)	3.41	1 / 4 / 6	-	-
Maximum (2-Step)	10.18	5 / 9 / 15	-	-
Number of Steps (2-Step)	1.61	1.3 / 2 / 2	-	-
Minimum (Endogenous)	3.58	1 / 4 / 6.7	1.79	-1 / 2 / 3
Maximum (Endogenous)	10.06	4.5 / 9 / 15	6.97	3 / 5 / 10
Number of Steps (Endogenous)	2.11	0.7 / 2 / 3.3	0.99	0 / 1 / 2

Note: Reported are the mean, 10^{th} (P10), 50^{th} (P50), and 90^{th} (P90) percentile of each variable.

In the midpoint treatments, respondents on average expect that the lowest possible rate of inflation over the next year is 3.41% (2-step) and 3.58% (endogenous), and that the highest possible rate of inflation over the next year is 10.18% (2-step) and 10.06% (endogenous). There are no significant differences between the two implementations, supporting consistency across the treatments. The number of steps shown in the table measures how many questions respondents on average answer for each elicited percentile. For the 2-Step treatment (number of steps capped at 2), the average number of steps is 1.61. That is, for narrow intervals (b_0, b_1), the number of steps is sometimes lower than 2 because the level of precision has already been reached. For the endogenous treatment, the number of steps is 2.11. 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), it turns out that two steps is often sufficient to approximate respondents' subjective probability distribution in an environment with a current rate of about 6%.

Figure 1 visualizes the sample distribution of the relative position of the point forecast and the implied means from the Midpoint method between the minimum and maximum expectation (U.K. survey, for U.S survey see Online Appendix D). Both distributions are highly skewed, with point forecasts and implied means closer to the minimum than the maximum. Note that the highly asymmetric distribution for the implied mean implies that participants in the midpoint treatments did not simply answer randomly: on average, random answers in the elicitation process would induce a symmetric distribution of implied means between the minimum b_0 and maximum b_1.

Figure 1: Sample Distribution of the Relative Position of Inflation Estimates between the Minimum and Maximum (U.K. Survey)



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 and at least a single choice was answered.

Feasibility is a precondition for the application of the Midpoint method in large population surveys. In Appendix A.3 we show that there are no differences between the methods in terms of perceived difficulty, perceived duration, actual time taken and attrition, despite the larger number of questions and screens for the Midpoint methods.

U.S. Survey. Participants expect the rate of inflation over the next 12 months to be on average 4.43%, which is higher than the current rate of 2.6% at that time. Despite the lower inflation environment, the number of bins used are basically identical to the U.K. sample. Consistent with the different inflation environment in the U.S. in 2024 compared to the U.K. in 2023, for the Midpoint treatment we observe clear downward shift in both the minimum. maximum, and the number of steps used. In terms of feasibility, we find that the actual time taken is lower for the Midpoint method than for the Bins SCE method, and again no differences in perceived difficulty, perceived duration and attrition.

4.2 Elicited Inflation Expectations

In this section, we compare the average implied mean forecast, disagreement, and average forecast uncertainty across the elicitation methods. The implied mean forecast refers to respondents' average inflation expectations as implied by their subjective belief distribution. 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 these statistics, we assume a mass-at-midpoint measure. For the midpoint treatments, this means that the probability mass in each of the four elicited ranges is uniformly distributed for all values contained in that range. For the density forecasts, it means that the probability mass in each bin is uniformly distributed. For the density forecasts, we also follow work by Becker et al. (2023) and assume that the open intervals have twice the width of the adjacent closed interval.[6] We pool results for the two Midpoint methods in the U.K. survey as results are very similar and do not differ for any of the statistics.

U.K. Survey. Results for the U.K. Survey are in the left panel of Table 3. The midpoint treatments and the mean-shifted density forecast treatment produce similar average implied mean

[6] Results are almost identical if we follow the Engelberg et al. (2009) procedure for the bin treatments. We fit a generalized beta distribution to the SCE bin treatment when the respondent assigns positive probabilities to three or more intervals and a triangular distribution when the respondent uses one or two intervals, from which we infer statistics such as mean, uncertainty, and disagreement. Results are in Online Appendix E.

forecast (6.21% and 6.27%, 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; two-sided t-tests). For forecast disagreement, we find a similar picture. The Midpoint and the mean-shifted density forecast treatments lead to similar disagreement (3.65 and 3.88, respectively) with no significant differences between the treatments. Disagreement in the regular SCE question format is significantly lower with 2.98 (differences significant at the 1%-level, Levene's test of homogeneity of variances). Note that the lower disagreement in the SCE format does not necessarily imply 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 are more personalized than the SCE bins. In the SCE format, 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, and in the mean-shifted density forecast the center of the bin structure is also personalized (by their elicited point forecast).

Table 3: Inflation Expectations Across Treatments

	U.K. Survey (Sep. 2023)			**U.S. Survey (Nov. 2024)**	
	Midpoint (N=405)	**Bins SCE (N=200)**	**Bins Shift (N=206)**	**Midpoint (N=2,015)**	**Bins SCE (N=2,003)**
Implied Mean	6.21	5.09***	6.27	3.99 (5.06)	2.75 (3.24)***
Disagreement	3.65	2.98***	3.88	6.01 (8.08)	2.75 (3.07)***
Uncertainty	1.46	3.78***	3.14***	1.30 (1.75)	3.36 (3.78)***

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). Values for the control group for the U.S. survey are reported in parentheses. 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. All tests are against the midpoint treatment, with ***, **, and * denoting significance at the 1%, 5%, and 10% level, respectively.

Average forecast uncertainty is substantially lower for the midpoint treatments (1.46) than for the Bins SCE and Bins Shift treatments (3.78, 3.14, respectively). All differences between the midpoint treatments and bins treatments 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. These results are consistent with Pavlova (2025), who finds that asking for a minimum, maximum and modal inflation instead of relying on density forecasts reduces uncertainty by 0.6 to 1 percentage points.

U.S. Survey. Results for the U.S. Survey are in the right panel of Table 3. Because the elicited distributions are affected by the randomized information treatments, we provide both statistics based on the whole sample and those based on the control group only (in parentheses). We observe that all patterns found in the U.K. survey replicate, for both the full sample and the control group.

Our finding of higher uncertainty for the SCE bins compared to all personalized methods suggests that the provided framing may conflict with respondents' own views, increasing uncertainty. The Midpoint methods may also elicit lower uncertainty because they do not require respondents to exercise precise probability judgements, and cannot induce them to allocate positive probability mass to extreme bins they would not otherwise have considered. This is especially true for the deflation bins, which may then also downward-bias the implied means, explaining the observed differences between the SCE bins and the other conditions. Section 4.3. will investigate these issues in more detail.

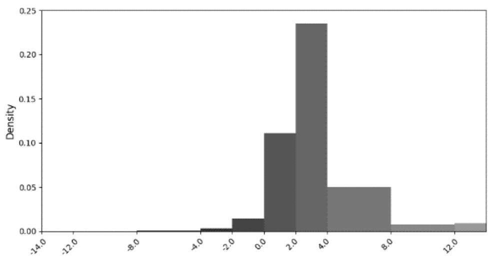
4.3 Method Comparison

We will now assess the quality of the subjective expectations the Midpoint and density forecast methods elicit. We first compare the elicited distributions and how well they reflect respondents' subjective beliefs. We then consider the consequences for the correlation between implied mean and point forecasts. Finally, we consider how well expectations elicited with each of the methods map on respondents' consumption decisions.

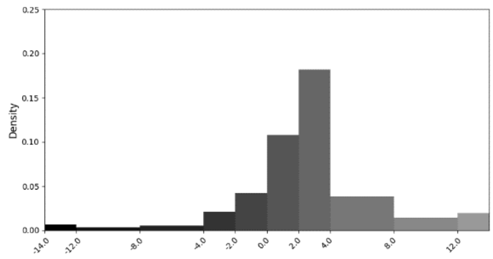
Distribution of Inflation Expectations and Subjective Quality Checks. To compare the subjective distribution of inflation expectations across treatments, we fit histograms of the average density assigned to the intervals of the Bins SCE format for all methods. To this end, we assume that outcomes within a given quartile in the midpoint treatments, or a given bin in the Bins Shift treatment, are uniformly distributed. In the latter, we also assume that the open intervals of the shifted bins have twice the width of the adjacent closed interval.

Figure 2: Density Plots

Panel A: Midpoint (U.S.)



Panel B: Bins SCE (U.S.)



Notes: Bins SCE format histograms of average densities for 12-month ahead inflation expectations elicited through midpoint method or density forecast, U.S. survey.

Results are displayed in Figure 2 for the two conditions ran in the U.S. survey. Two insights emerge from the figure. First, the probability mass allocated to the modal bin (2% to 4%) is more pronounced in the to the Midpoint method (47.08%) than in the Bins SCE method (36.31%). Second, the Bins SCE treatment elicits 17.25% probability mass for deflation, with some of it allocated to substantial deflation rates. In contrast, the Midpoint treatment elicits just 4% probability mass for deflation. In the U.S. survey, when explicitly asked for the lowest possible rate of inflation, 11.4% of the respondents consider any deflation in their answer.

Histograms for the U.K. survey are in Online Appendix F and replicate these patterns. Interestingly, in the U.K. survey the Bins Shift treatment elicits a 7.1% probability mass for deflation, which falls in between the midpoint treatments (1.4% and 2.8%) and the Bins SCE treatment (13.3%). In the U.K. survey, when explicitly asked for the lowest possible rate of inflation, 4.9% of the respondents consider deflation in their answer.

In both surveys, the SCE bins format elicits significantly lower implied means and substantially larger deflation probabilities than the Midpoint method. To understand whether these differences indicate a better performance of the Bins or the Midpoint method, we confront respondents with the characteristics of the distributions based on their answers as described in Section 3. Results are shown in Table 4. When presented with a statement that inflation is equally likely to be above and below the median value resulting from their own answers, 60.40% of respondents agree for the Midpoint method, and 57.56% agree for the Bins SCE method (p = 0.068). Importantly, those who are not happy with the statement are equally split between rather higher and rather lower inflation for the Midpoint method, but are significantly more likely to say that it should be higher for the Bins SCE method. Thus, the Bins method seems to elicit too low values for the median.

Table 4: Quality Checks for Median and Deflation Probability

	Perception Median		*Probability of Deflation*	
	Midpoint	**Bins SCE**	**Midpoint**	**Bins SCE**
Agree	60.40	57.56*	56.18	46.28***
Rather higher	19.75	25.76	26.20	11.68
Rather lower	19.85	16.67$^{\#\#\#}$	17.62$^{\#\#\#}$	42.04$^{\#\#\#}$

Note: Entries are % or respondents. t-tests for the difference in agreement are denoted using ***, **, * which correspond to significance at the 1%, 5%, and 10% level, respectively. t-tests for the difference between higher/lower (within method) are denoted using $^{\#\#\#}$, $^{\#\#}$, $^{\#}$ which correspond to significance at the 1%, 5%, and 10% level, respectively.

Consistent with this finding, when asked about the probability of deflation derived from their own previous answers, 42.04% of Bins SCE respondents say that this number should rather be lower. This is significantly larger than the 11.68% who say that it should be larger, and close to the share of 46.28% who agree with the value. In the Midpoint method, a significantly larger share of 56.18% agree with the probability of deflation derived from their answers. 26.20% think it should be larger, significantly more than the 17.62% who think it should be lower.

The quality check on the basis of respondents' own implied medians and probabilities of deflation thus reveals that a substantially larger share of respondents in the Midpoint treatment agrees with statements derived from their own previous answers. The Bins SCE method seems to elicit too large probabilities of deflation, which then biases the implied median downward. The Midpoint method seems to elicit too small probabilities of deflation, but this effect is much less pronounced than the opposite effect for the Bins method.

Correlation. Armantier et al. (2013) suggest that the correlation between respondents' point forecasts and the implied means of their elicited distributions provides a measure of reliability of the probabilistic questions. Table 5 displays these correlations. In the U.K. survey, 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 pooled midpoint treatments, we still observe a high correlation between point estimates and implied mean forecasts of 0.73. The Bins SCE treatment shows the lowest correlation of only 0.51. In the U.S. survey, the correlation is 0.75 for the Midpoint treatment, and a very low 0.2 for the Bins SCE treatment. A closer look at the low correlation for the Bins SCE in the U.S. sample reveals that it is driven by respondents who fill in positive probability mass in the most extreme deflation bin. Excluding these respondents yields a correlation of 0.6 for the U.S. sample, a much higher consistency between density and point forecasts. The same effect obtains in the U.K. sample, but less pronounced: excluding those with positive probability mass in the most extreme deflation bin increases the correlation to 0.54.

Overall, in both surveys the respondents are more consistent in their assessments when answering the Midpoint method compared to the Bins SCE method. The Bins SCE structure seems to induce some respondents to allocate some probability mass to the lower extreme (deflation) bins, unrelated to their actual beliefs. This reduces the correlation between point and density forecasts.

Table 5: Correlation between Point Estimates and Implied Mean Forecasts

	U.K. Survey (Sep. 2023)		U.S. Survey (Nov. 2024)	
	Correlation	**Confidence Interval**	**Correlation**	**Confidence Interval**
Midpoint	0.73^{***}	[0.68, 0.77]	0.75^{***}	[0.71, 0.79]
Bins SCE	0.51^{***}	[0.40, 0.61]	0.20^{***}	[0.10, 0.29]
Bins Shift	0.85^{***}	[0.81, 0.89]	–	–

Note: ***, **, and * denote significance at the 1%, 5%, and 10% level. Reported are 95% confidence intervals for the Pearson's correlation coefficients. Intervals are calculated using Fisher's z transformation.

Spending on Durable Goods: U.S. Survey. Finally, we assess the external validity of the elicited measures, by linking inflation expectations to respondents' consumption decisions. In particular, as described in Section 3, we make use of randomized information treatments included in the U.S. survey to identify causal effects of inflation expectations on consumption intentions. We measure consumption intentions by the following three items: *Is this a good time or a bad time to buy durable goods like electronics, mobile phones, home appliances? Is this a good time or a bad time to invest in real estate in your city? Is this a good time or a bad time to buy a car?* In contrast to some consumption measures that are distorted by other constraints the respondents face, these items have the advantage that they should show a clearly positive relation to a person's inflation expectations (see Duca-Radu et al., 2021; D'Acunto et al., 2022). We follow the approach by Coibion et al. (2024) and Georgarakos et al. (2024) and estimate the following 2-stage model:

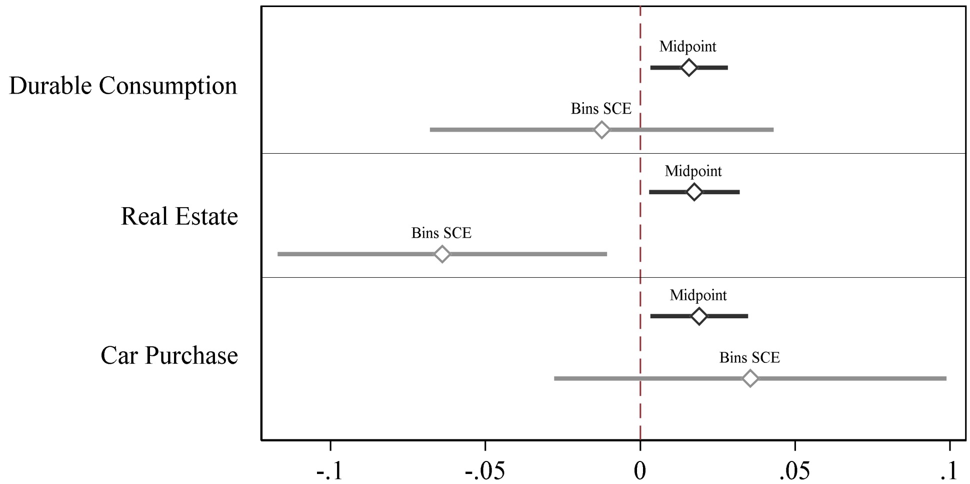
$$1st\ Stage: Post_i^{mean} = a_0 + \sum_{j=1}^{4} a_j \times I\{i \in Treat_j\} + \sum_{j=1}^{4} b_j \times I\{i \in Treat_j\} \times Prior_i^{mean} + Controls_i + error_i \quad (1)$$

$$2nd\ Stage: Spending_i = \alpha_0 + \alpha_1 \widehat{Post}_i^{mean} + Controls_i + error_i \quad (2)$$

where $Post_i^{mean}$ is the implied mean elicited after the information treatments using either the Midpoint or bins method, $Prior_i^{mean}$ is the point forecast before the information treatments, and

$I\{i \in Treat_j\}$ is an indicator for each of our four information treatments (relative to the control group). Finally, we include all available household-level controls including age, gender, education, and household income. The first stage specification essentially consists of regressing posteriors on priors interacted with treatment group indicators, while the second stage regresses spending behavior on the treated posterior beliefs. Intuitively, we use the exogenous variation in inflation expectations to isolate the effects of inflation expectations (measured using two different methods) on spending behavior.

Figure 3: Inflation Expectations and Durable Consumption



Notes: This figure displays coefficient estimates for Equation 2 for each elicitation method and spending category. Displayed are 95%-confidence intervals using heteroscedasticity robust standard errors.

We estimate the model for each elicitation method (Bins SCE and Midpoint) and spending category (durable consumption, real estate, car).[7] Figure 3 shows the coefficient estimates of $Post_i^{mean}$ in the second stage specification for the six models. We find a small but positive effect of the expectations measured by the Midpoint method for all three spending categories. In contrast, for Bins SCE expectations we find a significantly negative coefficient for the real estate question, an insignificantly negative effect for durable consumption and an insignificantly positive effect for

[7] The first-stage results (how households revise their beliefs) are displayed in Online Appendix G.

car purchases. That is, while we find a consistent effect of expectations on consumption intentions for the Midpoint method, we find insignificant and contradictory effects for the Bins SCE method. As shown above, the expectations measured through the Bins SCE method appear to not reflect respondents' subjective expectations well. The implied means used to predict consumption intentions may thus not reflect true differences in expectations between respondents, leading to these inconsistencies.

5. A Further Quality Test: Predicting Known Distributions

In this section we report on a further test of the Bins versus Midpoint approach, benchmarking the methods to an objective underlying distribution, rather than to subjective beliefs. We designed an experimental task where participants sequentially sampled 20 numbers from a set of 100 (positive and negative) numbers without replacement. The distribution was either symmetric or right-skewed, in a between-subject design. Participants then report their full distribution of beliefs about another draw, from the full distribution of 100 numbers, of which they have seen 20. They use either the Midpoint Endogenous or the Bins SCE method, again between-subject. The experiment uses abstract framing, not referring to inflation. Experimental details are in Online Appendix H.

Our quality criterion is how well the participants' expectations reflect the true underlying distributions in terms of matching their mean, standard deviation, and probability of a negative number (similar to deflation probability). Results are displayed in Table 6. For the Midpoint method, the implied mean closely matches the true mean for the symmetric, and slightly overestimates it for the skewed distribution. In contrast, the Bins SCE method significantly underestimates the true mean of both the symmetric and skewed distribution, with deviations from the underlying mean much greater in absolute terms than the deviations in the Midpoint method. Testing for differences between the methods in the balance of forecasts on different sides of the true value (i.e., those who under- or overestimate), we find that the Midpoint method is better at approximating the true underlying mean of the distribution ($p < 0.001$ for both comparisons).

Table 6: Quality Checks Sampling Experiment

	Implied Mean		**Uncertainty**		**Probability Negative Number (in %-points)**	
	Symmetric	*Skewed*	*Symmetric*	*Skewed*	*Symmetric*	*Skewed*
Midpoint	0.31	0.84***	-0.10	0.04	-3.22**	-9.67***
Bins SCE	-1.10***	-1.42***	0.58***	0.77***	11.00***	11.00***
p-value of difference	<0.001	<0.001	0.075	0.015	0.060	0.688

Note: Displayed are the median differences between the implied mean, uncertainty, and the probability of observing a negative number and their respective counterpart from the underlying true distribution, at the individual level. Differences are tested against the null that the implied statistics do not differ from the underlying distribution, with ***, **, and * denoting significance at the 1%, 5%, and 10% level (two-sided t-test). The final row reports p-values of tests for the equality in sample proportions of positive and negative deviations from true values, between both methods.

For forecast uncertainty, the Midpoint method closely matches the true underlying standard deviation. The Bins SCE method yields a forecast uncertainty significantly larger than the sample standard deviation for both the symmetric and skewed distribution. Testing for differences between the methods in the balance of forecasts on different sides of the true value, we find that the Midpoint method is better at approximating the true underlying standard deviation of the distribution ($p = 0.075$ and $p = 0.015$ for symmetric and skewed, respectively).

Finally, comparing both elicitation methods in terms of the implied probability of a negative number, the Midpoint method appears to underestimate and the Bins SCE method appears to overestimate this probability. The true probabilities are 21% and 26% for the symmetric and asymmetric distributions, respectively. The Midpoint method underestimates these probabilities by 3.22% and 9.67%, the Bins SCE overestimates them by 11.00%. All deviations are significantly different from zero, and the Bins method is significantly more imbalanced then the Midpoint method for the symmetric distribution ($p = 0.06$). The strong overestimation in the Bins method is driven by the inappropriate use of the "negative" bins, while the underestimation in the Midpoint method derives from a reluctancy to state negative numbers when indicating the minimum expected number. These results are consistent with the survey evidence on the implied deflation probability. The severity of the biases is much smaller in the Midpoint than the Bins method, but it will depend on the underlying inflationary regime.

6. Conclusion

We propose a novel method to elicit the full distribution of subjective inflation expectations in surveys. Our method builds on decision theoretic research and presents a potential solution to many of the existing problems with density forecasts. Relative to the probabilistic question format of the SCE, the elicitation process is driven by the respondent and involves no anchoring on exogenously provided frames. The method therefore does not depend on the current state of the economy and works in both low and high inflation regimes, as well as 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, and a stronger correlation with the point predictions. 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). 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.

The method we propose is able to resolve several of the problems associated with density forecasts. The 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). In surveys that include several of these variables, there is then no need to provide several bin structures tailored to each variable: the Midpoint approach can directly be applied to each variable, with a, typically different, initial range provided by the survey respondent. This reduces potential confusion on the side of the survey respondents, makes measurements directly comparable, and allows to study these variables in conjunction.

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., Bruine de Bruin, W., Potter, S., Topa, G., Van Der Klaauw, W., & Zafar, B. (2013). Measuring inflation expectations. *Annual Review of Economics*, 5(1), 273-301.

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.

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.

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

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., Gorodnichenko, Y., & Weber, M. (2022). Monetary policy communications and their effects on household inflation expectations. *Journal of Political Economy*, 130(6), 1537-1584.

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.

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

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.

Duca-Radu, I., Kenny, G., & Reuter, A. (2021). Inflation expectations, consumption and the lower bound: Micro evidence from a large multi-country survey. *Journal of Monetary Economics*, 118, 120-134.

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2020). Effective Policy Communication: Targets versus Instruments. Becker Friedman Institute Working Paper 2020–148.

D'Acunto, F., Hoang, D., & Weber, M. (2022). Managing households' expectations with unconventional policies. *The Review of Financial Studies*, 35(4), 1597-1642.

Engelberg, J., Manski, C. F., & Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business & Economic Statistics*, 27(1), 30-41.

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.

Georgarakos, D., Gorodnichenko, Y., Coibion, O., & Kenny, G. (2024). The Causal Effects of Inflation Uncertainty on Households' Beliefs and Actions. NBER w33014.

Guiso, L., Japelli, T. & Pistaferri, L. (2002). An Empirical Analysis of Earnings and Employment Risk. *Journal of Business and Economic Statistics*, 20, 241-53.

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*, 145, 103568.

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

Kumar, S., Gorodnichenko, Y., & Coibion, O. (2023). The effect of macroeconomic uncertainty on firm decisions. *Econometrica*, 91(4), 1297-1332.

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. (2025). Framing Effects in Consumer Expectation Surveys. *Journal of Economic Behavior and Organization*, 231, 106899.

Potter, S., Del Negro, M., Topa, G., & Van der Klaauw, W. (2017). The advantages of probabilistic survey questions. *Review of Economic Analysis*, 9(1), 1-32.

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.

Ryngaert, J. M. (2023). Balance of risks and the anchoring of consumer expectations. *Journal of Risk and Financial Management*, 16(2), 79.

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.