Journal of Economic Behavior and Organization 231 (2025) 106899

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Framing effects in consumer expectations surveys

Lora Pavlova

ZEW - Leibniz Centre for European Economic Research, Germany

A R T I C L E   I N F O

JEL classification:
C83
D84
E31

Keywords:
Probabilistic expectations
Survey design
Household inflation expectations

A B S T R A C T

In a randomized survey experiment, I test how variations in question wording and format influence response behavior. Participants from a representative sample in Germany are divided into four groups, each receiving a different version of a question about expected inflation over the next 12 months. I compare two wordings used in leading consumer surveys: (i) the change in prices in general and (ii) the inflation rate. Additionally, I examine responses to a question about probabilistic beliefs regarding future inflation and a simpler one asking for the expected minimum, maximum, and most likely inflation rate. The findings show that response behavior varies significantly with framing. Simpler wording like ‘prices in general’ and less restrictive format produce higher mean expected inflation. While simpler wording leads to higher individual uncertainty, asking for the minimum, maximum, and mode yields lower uncertainty. The results suggest that framing in consumer expectations surveys can shape the elicited data, underscoring the importance of careful question design.

1. Introduction

Studying subjective expectations about macroeconomic and individual outcomes has gained importance as researchers have moved away from the assumption of full-information rational expectations (FIRE). Directly surveying consumers is a key method for eliciting their subjective expectations. Multiple central banks and research institutions conduct such surveys, and this practice has become more widespread in recent years.[1] At the same time, a growing strand of literature shows that studying subjective beliefs can help understand the heterogeneous choices agents make and design better monetary and fiscal policy tools (D’Acunto and Weber, 2024).

One particularly important example is inflation. Expectations about the future rate of price changes play an important role in forecasting inflation outcomes (Brandão-Marques et al., 2023) and economic activity, as well as for the transmission of monetary policy. Policymakers also carefully monitor inflation expectations for early signs of de-anchoring and potential threats to price stability (Nagel, 2022).

Charles Manski, one of the main advocates of the use of subjective expectations data, highlights the ‘importance of careful attention to question wording when eliciting expectations’ via surveys (Manski, 2018, p. 440). For instance, different wording choices may lead to variation in the data coming from different interpretations rather than true differences in beliefs (Bruine de Bruin et al., 2010). Furthermore, recent papers such as Becker et al. (2023) provide evidence on the priming effects of the probabilistic question format widely utilized for eliciting inflation expectations in several large-scale consumer surveys. The broad application of such data for both research and policy calls for re-examining in greater detail how survey design affects elicited macroeconomic expectations of consumers.

E-mail address: lora.pavlova@zew.de.

[1] For a detailed list of consumer and firm surveys worldwide and the institutions conducting them see e.g. Weber et al. (2022) and Coibion et al. (2020).

https://doi.org/10.1016/j.jebo.2025.106899
Received 22 July 2024; Received in revised form 7 January 2025; Accepted 10 January 2025
Available online 31 January 2025

Using an experiment embedded in a survey, this paper documents the causal effects of changes in framing on households' inflation expectations in Germany. More precisely, I focus on two aspects — format and wording. To the best of the author's knowledge, this is the first study to examine both aspects in the context of consumer probabilistic expectations. I find that response behavior varies strongly with framing. Acknowledging the potential shortcomings of the standard probabilistic approach to eliciting inflation expectations, I propose a simpler question asking for the minimum, maximum, and mode instead. Central measures and uncertainty computed from these two questions differ significantly. Asking for the minimum, maximum, and mode yields higher implied means by more than 1 pp on average. In contrast, individual uncertainty measured by the standard deviation of the distribution, is estimated to be 0.6 to 1 pp smaller. Expectations of those provided with the 'inflation rate' wording seem less upward-biased and more concentrated around 2%. This is consistent with prior findings for point forecasts. However, a substantial share of respondents states they think most about specific goods' prices such as those of food and gas, independent of the wording choice. Even when asked directly about expected inflation, only about one-quarter report thinking about this broader concept when making their forecast.

The remainder of the paper is structured as follows. Section 2 discusses the role of wording and format in the elicitation of consumers' subjective inflation expectations, while Section 3 lays out the experimental framework. Section 4 gives an overview of the data set and Section 5 summarizes the estimated treatment effects. I discuss some practical aspects of response behavior and their differences by treatment in Section 6. Section 7 concludes.

2. Background and related literature

Wording Longer-running surveys such as the Michigan Survey ask participants about their expectations about the 'change in prices in general' (Curtin, 1996). Although simpler and more understandable, this wording yields more extreme values and larger disagreement in the data (D'Acunto et al., 2023). While variation in inflation expectations can be (partially) attributed to factors related to personal experiences and characteristics,[2] another potential source is question interpretation. Bruine de Bruin et al. (2010) report that people tend to interpret the 'prices in general' wording as asking most for specific prices such as food and gas. Hence, consumers may report expectations for different outcome variables.

In a more recent survey, the Federal Reserve Bank of New York (FRBNY) adopted an alternative wording that directly asks about the expected inflation rate. This more precise formulation aims at reducing ambiguity and variation in responses originating from different interpretations rather than true differences in beliefs (Armantier et al., 2017). However, it potentially comes at the cost of a higher degree of complexity. Some respondents may not have a clear understanding of what the economic term 'inflation rate' entails. This wording can, therefore, lead to higher non-response rates, especially among populations with lower education and financial literacy, or limited cognitive abilities (D'Acunto et al., 2020).

Previous experimental studies indeed show that the wording choice can significantly impact point forecasts for inflation. Using a representative US sample, Bruine de Bruin et al. (2012) document that point predictions from a question using the 'prices in general' or 'prices you pay' wording exceed those produced by the 'inflation rate' wording. This also holds for respondents' perceptions of inflation. While the authors find almost equal non-response rates across wordings, survey participants do report having more difficulty understanding terms such as the 'inflation rate'.[3]

Format Together with a new wording, the FRBNY also introduced a new type of format - a probabilistic 'bin' format. While point predictions provide a view for the variable of interest expected 'on average', they contain no information about the associated individual uncertainty. One possibility to capture this uncertainty is to elicit the whole subjective distribution of the respondent.[4] This is done by asking participants to assign probabilities to a set of pre-defined, non-overlapping intervals representing possible outcome ranges for the variable of interest ('bins'). Accordingly, the probabilities should sum to 100%. In the following years, several central banks followed suit and adopted this question format in large consumer expectation surveys, among others the Bank of Canada, the Deutsche Bundesbank, and the European Central Bank. More precisely, the question about short-term inflation expectations from the SCE reads as follows:

Q9 Now we would like you to think about the different things that may happen to inflation over the next 12 months. [...] 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
the rate of inflation will be between 8% and 12%
⋮
the rate of deflation (opposite of inflation) will be 12% or higher[5]

[2] For example, differences in consumption baskets and thus exposure to different prices (D'Acunto et al., 2020), cohorts living through various inflation regimes (Malmendier and Nagel, 2016; Goldfayn-Frank and Wohlfart, 2020), socioeconomic status (Das et al., 2020), diverse financial planning horizons (Bruine de Bruin et al., 2010).

[3] For Europe, Bruine de Bruin et al. (2017) document similar findings using a nationally representative survey of Dutch consumers, albeit the differences in levels between various wordings appear less pronounced.

[4] For a more detailed discussion on point forecasts versus probabilistic forecasts see e.g. D'Acunto et al. (2023), Dräger and Lamla (2024).

[5] Respondents are asked to assign probabilities for the following intervals: $(-\infty,-12]$, $(-12,-8]$, $(-8,-4]$, $(-4,-2]$, $(-2,0]$, $(0,2]$, $(2,4]$, $(4,8]$, $(8,12]$, $(12,\infty)$. Note that the design does not specify which intervals contain the bounds. Right-closed intervals are assumed by the author to enable further processing of the data, see also Section 4. The original questionnaire is available online at https://www.newyorkfed.org/microeconomics/sce/background.html.

While highly informative, this format also suffers several disadvantages. For once, it strongly relies on the respondents' capabilities of expressing themselves via probabilities, which in turn requires a certain degree of numeracy and sophistication. While an overall willingness to convey expectations in a probabilistic manner has been documented in earlier studies (Armantier et al., 2013), several artifacts have been observed in the data so far. For example, when it comes to open-ended probability questions, studies document a high proportion of 0, 50 and 100% answers, also known as 'focal point responses' (Hurd, 2009; Dominitz and Manski, 1997). Stating 50% might indicate high epistemic uncertainty among respondents and signal they struggle with numbers and probabilities, or to form a distribution about the expected outcomes of an unknown concept (Fischhoff and Bruine de Bruin, 1999; Bruine de Bruin et al., 2000). In contrast, in the context of probabilistic 'bin' question, stating 50%–50% can be interpreted as the respondent being quite certain of the outcome. Moreover, even if people have precise probabilities in mind, they may resort to rounding to facilitate communication (Manski and Molinari, 2010; Manski, 2018). Based on the raw data alone, it is unclear which one applies.[6]

A more recent study by Becker et al. (2023) provides further experimental evidence of the large influence of the underlying response scale on the reported distributions, implied means and uncertainty. Among others, the authors show that participants tend to assign higher probability mass to a given numeric range, as the number of bins representing this range increases.[7] This bias can potentially overstate the degree of anchoring to the inflation target or to any other value, depending on how the bins are specified. Moreover, even if policymakers are interested in expectations volatility rather than their level, uncertainty derived from density forecasts also seems very prone to changes as the underlying intervals change (Becker et al., 2023). Hence, even though there are clear benefits from using a probabilistic question and it has since been adopted by multiple institutions in their surveys, these aspects call for examining a viable alternative to this format.

More generally, this paper contributes to the strand of literature focusing on survey design. In the context of consumer (or firm) inflation expectations, several aspects have been addressed so far. Bruine de Bruin et al. (2017) discuss the effect of administration mode (e.g. face-to-face vs. web-based surveys) and opportunities to revise answers similar to the practice in the Michigan Survey (Curtin, 1996). Bruine de Bruin et al. (2012) and Bruine de Bruin et al. (2017) study the effect of wording on the central tendency and disagreement of point forecasts for inflation as well as perceptions, and report significant differences between wordings. Coibion et al. (2020) do not find any systematic biases in the first and second moments based on variations in wording for firms in New Zealand. Phillot and Rosenblatt-Wisch (2024) examine the effect of question ordering on the forecast consistency of firms and report significant differences depending on whether the point or the density forecast is elicited first. Hayo and Méon (2023) assess the effect of guided vs. non-guided questions on the reported inflation expectations and non-response rates. Boctor et al. (2024) report significant differences in implied means and uncertainty derived from the probabilistic question as opposed to a scenario-based question.

3. An experiment on framing effects

Format Using an experiment embedded in a representative online survey among consumers in Germany, I test how using an alternative format for eliciting subjective distributions of expected inflation fares against the existing probabilistic 'bin' approach. The question, which I will henceforth refer to as 'min–max', reads as follows:

Question 1: What do you think the rate of inflation (or rate of deflation) is most likely to be over the next twelve months? What will the rate of inflation be as a maximum or minimum value?

The min–max question allows for eliciting a subjective distribution and hence uncertainty for the variable of interest without explicitly asking for probabilities, but only about the expected minimum, maximum and mode. It thus reduces the associated cognitive load and eliminates response scale bias. While the min–max question is more limited regarding the functional form of the distribution – a triangular distribution – it loosens the symmetry assumption. Relaxing this assumption is central for a more accurate modeling of risk perceptions of consumers (Ryngaert, 2023). Moreover, Krüger and Pavlova (2024) report that depending on the underlying variable, 30 to 40% of respondents in the SCE already use sparse histograms (with only one or two bins), for which the triangular distribution is used for further processing (Engelberg et al., 2009). Furthermore, the min–max question has the advantage that respondents can explicitly state the ends of the support of their histogram rendering assumptions on the unbounded intervals unnecessary. The fact that respondents are not presented with any numbers or ranges beforehand reduces potential priming.

By design, Question 1 does not accommodate cases that include one or more disjoint regions with positive probability mass or exhibit bi-modality.[8] Importantly, while they should not necessarily be deemed uninformative and can reflect true beliefs, being more difficult to handle by standard approaches for quantification, they are usually excluded from the analysis. Hence, if respondents are willing to convey their beliefs in the min–max format, this will lead to an increase in information. Either way, employing the min–max question would not lead to a loss of information in this regard.

Questions similar in design have been implemented in the context of households (Coibion et al., 2024; Christelis et al., 2020) and firms (Altig et al., 2022), however, in either case, the respondents are additionally asked to assign a probability to the corresponding

[6] Manski and Molinari (2010) suggest examining the set of responses provided by the survey participant in order to draw on their rounding pattern, i.e. exact or rounded reporting. However, people might adopt different patterns depending on whether the elicited outcome is personal or macroeconomic, and if the latter — how familiar they are with the concept, and so on.

[7] For a more detailed discussion see Section 5.1 of Becker et al. (2023).

[8] Although such cases are not dominant in household expectations data, they often represent a non-trivial share even in large samples. For the SCE between June 2013 and November 2020, Zhao (2023) documents less than 5% of histograms with disjoint regions and about 14% containing multiple modes.

scenario, e.g. to the 'lowest' or 'highest' outcome, or the average of the two. In contrast, the setup proposed above, which is closely related to the one used for eliciting a firm's expected sales growth in the ifo Business Survey (Bachmann et al., 2021), is completely free from numerical probabilities. Huisman et al. (2021) also use a similar framework, asking for a point forecast, a minimum, and a maximum level of the AEX index to collect stock market expectations of Dutch investors. Goldfayn-Frank et al. (2024) propose a theoretically motivated, multi-step approach to eliciting household expectations, where in the first step respondents are also asked for the expected minimum and maximum outcome.

Wording While the majority of studies so far have focused on the effect of wording in the context of point forecast, little is known whether these effects persist in the probabilistic setup. For example, the biasing effect of thinking about specific prices could be offset by the dampening effect of the response scale. Therefore, following Bruine de Bruin et al. (2010) and Bruine de Bruin et al. (2012), I test the effect of using a simpler wording such as 'prices in general'[9] on the resulting subjective distributions for short-term inflation. The experimental design allows for comparing the effect of a change in wording, conditional on format, but also testing for any potential interactions between the two.

To mitigate survey fatigue arising from asking respondents for their expectations about the same variable over the same horizon twice, I chose a between-subjects design. This should also help balance the cognitive load and minimize the impact of experimenter demand effect on the collected data (Stantcheva, 2023). To this end, the survey participants are randomly split into four treatment arms as follows:

- Group A1: probabilistic question about the inflation rate
- Group B1: min–max question about the inflation rate
- Group C1: probabilistic question about changes in prices in general
- Group D1: min–max question about changes in prices in general

Importantly, one should note that respondents who receive the 'inflation rate' wording are shown a short definition of what inflation is using an info box: *'Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as deflation.'* This has been established in previous waves of the survey. Hence, the comparison occurs between a question asking directly for the expected inflation rate with a definition of inflation against asking for expectations of changes in prices in general. This feature of the survey design could potentially reduce the asymmetry in responses across different demographic groups which stems from lack of understanding of what inflation is and therefore diminish any existing differences. It is of prime importance to gain insight whether providing a definition for inflation indeed mitigates concerns associated with more complex wording as explained in Section 2.

Forming expectations To shed more light on whether wording affects the variability of responses through the question-interpretation channel, respondents are asked what they thought most about when producing their forecast. The question is adapted from Bruine de Bruin et al. (2010) and reads as follows:

Question 2: What did you think about most when answering the question about your inflation rate expectations before?[10]

Respondents are presented with five options to can choose from: (i) prices you pay in your everyday life such as food and gas, (ii) prices paid by households in Germany, (iii) Germany's inflation rate, (iv) changes in the cost of living, (v) other specific prices, whereby the last option is open-ended. The options selected for the current experiment were the top-rated in the original one (see Bruine de Bruin et al., 2010, Table 2). One can expect that people who were randomly assigned to the 'prices in general' treatment arm, should more likely select 'everyday prices'. Accordingly, if interpreted correctly, respondents who received the 'inflation rate' wording, should most often select the corresponding option — 'Germany's inflation rate'.

4. Data set

The experiment was conducted as part of the 9th wave of the Bundesbank Online Panel – Households (BOP-HH) in September 2020. BOP-HH is an online survey launched in 2019 and currently conducted at a monthly frequency, involving a variety of topics on both personal and macroeconomic outcomes. The survey is representative of the German online population aged 16 and above. The sample size varies from 2000 to 5000 respondents and contains a panel component. The September 2020 wave of BOP-HH has a sample size of roughly 4000 respondents. Thus, in each of the four treatment arms there are about 1000 participants.[11]

As part of the core module of the survey, participants receive a number of questions specifically targeted at inflation expectations. First, they are asked about their perception of inflation over the last 12 months. Then, regarding the upcoming 12 months respondents report whether they expect inflation or deflation and at how much percent. Finally, depending on the treatment group, they are given either the probabilistic or the min–max question again asking for short-term inflation expectations. In the probabilistic treatment, participants are additionally instructed that the sum across bins should be one hundred. As they insert their answers, the current sum of the probabilities is displayed. In case this condition is not met an error message is displayed, asking the respondent

[9] Earlier experiments such as Bruine de Bruin et al. (2012) include a third alternative - 'prices you pay'. For this, however, the same arguments as for the 'prices in general apply, and in most cases, it does not produce significantly different expectations, which is why I do not consider it in the experiment.

[10] The wording is adjusted accordingly for groups C1 and D1. For the exact question wording see Appendix C.

[11] For an overview of the survey see also Beckmann and Schmidt (2020).

to correct their answer. Unless the probabilities sum to 100%, one cannot proceed. If one attempts to skip the question, two options of non-response are shown: 'Do not know' and 'No answer'. All questions are documented in detail in Appendix C.[12]

In addition to the core and project-specific questions, information on the respondent's age, education, employment status, household income and size, the number of children, and others is collected. Further details are provided in Appendix B of the Supplementary Material.

Computing moments and uncertainty from the raw probabilities resp. the minimum, maximum and mode, requires further processing of the data. For fitting a distribution to the subjective histograms I follow Engelberg et al. (2009) and fit a symmetric triangular distribution to histograms with at most two bins with positive probability and a generalized Beta distribution else. Some necessary adjustments are made, given that the method was initially designed for bins of equal width, which is not the case in the current setup.[13] For the min–max question, I fit a simple triangular distribution to the support endpoints and mode, a, b, and c. Appendix D provides further details on the data processing. In the following, I provide some descriptions and summary statistics on the elicited expectations data.

5. Effects of wording and format on consumers' inflation expectations

This section presents the empirical results. First, I briefly discuss the resulting ('raw') probabilities as well as the averages of fitted moments, percentiles, and endpoints across treatment arms in Section 5.1. Section 5.2 presents the estimated effects for implied means and uncertainty. In Section 5.3 I analyze how the choice of wording impacts the collected data via the question interpretation channel.

5.1. Comparison of subjective probabilities and summary statistics

The upper panel of Fig. 1 depicts the probabilities respondents assigned to the corresponding bins averaged across treatment arms. Blue bars represent those for participants sub-sampled into the 'inflation rate' treatment group, and green ones for the 'prices in general'. The resulting distributions are somewhat similar in shape, with probability concentrated on positive outcomes. The interval $(0, 2]$ that contains the inflation target at the time, attracts more probability in the 'inflation rate' than in the 'prices in general' treatment. Overall, higher inflation outcomes ($> 4\%$) are deemed more probable in the 'prices in general' treatment. In both cases, respondents consider deflationary trends quite likely — in the 'inflation rate' slightly more than in the 'prices in general'. The difference is statistically significant at the 1%-level.

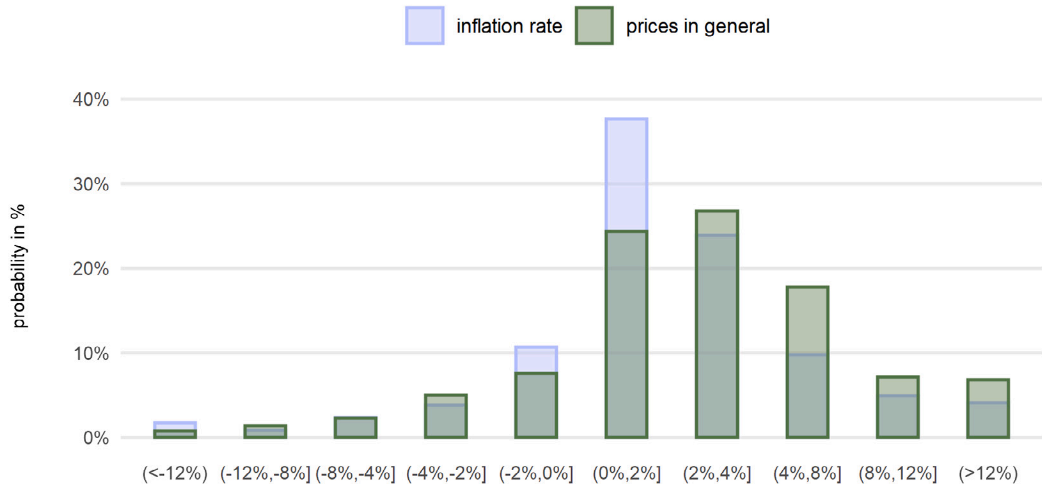
For comparability, the lower panel of Fig. 1 reports the difference $F(u_k)_i - F(u_{k-1})_i$, where u_k is the upper bound of interval k with $u_k \in \{-12, -8, -4, -2, 0, 2, 4, 8, 12, \infty\}$ averaged across respondents. $F(x)_i$ is the CDF of the triangular distribution of respondent i, based on the parameters, a, b, and c from the min–max question. In a sense, the reported distributions are 'discretized' to match the bins of the probabilistic format and then the probabilities are averaged across treatment arms.

The most striking difference is that the probability mass assigned to deflation dramatically declines compared to the upper panel of Fig. 1. It goes down from roughly 20% to less than 3% in the 'inflation rate' treatment. Similarly from 17% to 3% in the 'prices in general' treatment. Expectations that are essentially bounded by zero are well-documented in previous literature and base on the notion that positive price changes are more salient for consumers than negative ones (D'Acunto et al., 2021; Cavallo et al., 2017). Furthermore, the difference in deflation probability across formats is in line with the findings of Goldfayn-Frank et al. (2024). One possible explanation is that providing respondents with more intervals representing negative values increases the probability placed in those (Becker et al., 2023). In the case of the min–max question where respondents are not confronted with any intervals, the probability of deflation significantly drops. Apart from that, the distributions exhibit similar features with a heavy right tail, especially for the 'prices in general' wording.

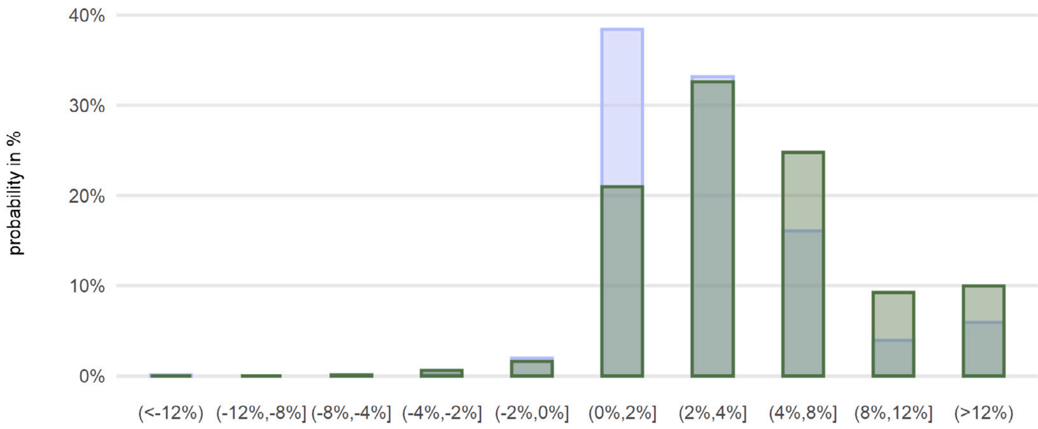
Table 1 presents average values for selected fitted moments, endpoints, and percentiles across all four treatment groups. Notably, the left endpoint, S_r and the 5%-percentile, p_5, are often negative in the probabilistic setup but positive — and slightly above 2% — in the min–max question. While S_r shows substantial variation across treatment arms, conditional on wording, p_{95} is rather stable across groups. However, caution is advised when interpreting these results, as they are sensitive to the underlying assumptions on the support of the fitted distribution in the probabilistic question. One alternative would be to use the highest and lowest reported point forecasts in the sample as support endpoints. While feasible, this approach introduces variability in the assumptions across survey waves, thus reducing comparability over time. This highlights an advantage of the min–max question: it avoids the need for such assumptions, ensuring consistency over time and across studies. Finally, irrespective of format, the 'prices in general' wording always produces (i) higher implied means and percentiles and (ii) higher standard deviations. In the following section, I examine these differences more formally.

[12] Additionally, the Deutsche Bundesbank makes all past questionnaires, including the one used in the September 2020 wave, available online at https://www.bundesbank.de/en/bundesbank/research/survey-on-consumer-expectations/questionnaires-850746.

[13] See Krüger and Pavlova (2024) and the associated software available at https://github.com/FK83/forecasthistogram.



(a) probabilistic format (default)



(b) min-max format

Fig. 1. Average subjective probabilities for different ranges for inflation from the probabilistic and min–max format. Blue bars display probabilities assigned by those who received the 'inflation rate' wording, green bars for those with the 'prices in general'.

Table 1
Averages of moments, support endpoints, and percentiles.

	$E_i[\pi_{t+1}]$ (1)	$\sigma_i^{\pi_{t+12}}$ (2)	S_l (3)	S_r (4)	$p5$ (5)	$p25$ (6)	$p50$ (7)	$p75$ (8)	$p95$ (9)
	probabilistic question								
A1: 'inflation rate'	2.11	1.92	−4.09	10.16	−0.95	0.71	2.05	3.46	5.35
C1: 'prices in general'	3.55	2.17	−2.83	13.62	0.21	1.94	3.43	5.05	7.29
	min–max question								
B1: 'inflation rate'	4.31	0.87	2.20	6.28	2.85	3.67	4.33	4.96	5.71
D1: 'prices in general'	5.51	1.34	2.24	8.54	3.25	4.53	5.54	6.52	7.67

Note: Columns (1) to (9) report the averages per treatment group for the following measures: implied mean, standard deviation, left endpoint, right endpoint, 5%-, 25%-, 50%-, 75%- and 95%-percentiles.

5.2. Effects on individual mean expectations and uncertainty

Given the experimental setup described in Section 3, one can estimate the treatment effects of changes in wording and format on the desired inflation expectations measure, y_i, as follows:

$$y_i = \alpha + \beta \text{min–max}_i + \gamma \text{prices}_i + \delta \text{joint}_i + \varepsilon_i, \tag{1}$$

Table 2
Treatment effects on individual implied means and uncertainty.

	Dependent variable:					
	$E_i[\pi_{t+1}]$			$\sigma_i^{\pi+12}$		
	(1)	(2)	(3)	(4)	(5)	(6)
default	2.218***	1.848***	0.893	2.119***	1.320***	2.730***
	(0.218)	(0.062)	(0.753)	(0.103)	(0.031)	(0.311)
min–max	2.434***	1.068***	2.580***	−1.126***	−0.649***	−1.110***
	(0.382)	(0.086)	(0.389)	(0.124)	(0.035)	(0.124)
prices	1.311***	0.998***	1.290***	0.393***	0.415***	0.362**
	(0.328)	(0.110)	(0.323)	(0.152)	(0.060)	(0.149)
joint effect	4.214***	2.326***	4.156***	−0.481***	−0.257***	−0.498***
	(0.423)	(0.108)	(0.428)	(0.161)	(0.039)	(0.166)
Observations	3,686	3,686	3,549	3,686	3,686	3,549
Robust linear		✓			✓	
Controls			✓			✓

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Robust standard errors in parentheses. Survey weights are used to ensure representativeness. Columns (1)–(3) present the results for mean expected inflation, Columns (4)–(6) - for uncertainty. 'Robust linear' refers to Huber robust regression specifications, accounting for influential observations. 'Controls' refers to the inclusion of demographic control variables such as age, gender, education, income and others. A full list of control variables is provided in Table B.1 of Appendix B.

In this study, I focus on the implied mean $E_i[\pi_{t+1}]$ and the standard deviation $\sigma_i^{\pi_{t+12}}$ as proxy for individual uncertainty. These measures are commonly used in empirical literature on household expectations (Weber et al., 2023) and for reporting current inflation developments by central banks.[14] In Eq. (1), α represents the average value of the desired measure for group A1, which receives the default 'bin' question, as outlined in Section 3. The remaining terms on the right-hand side capture the average of the implied mean (or uncertainty) for the other three treatment groups, relative to the default group. That is, the variables $min\text{–}max_i$, $prices_i$, and $joint_i$ are dummies taking unit value when respondent i is randomly assigned to group B1, C1 or D1, respectively.

In a scenario where neither wording nor format affects response behavior, the resulting measures should be consistent across treatment arms, assuming random assignment and a sufficiently large sample size. However, as outlined in Section 2, the probabilistic question tends to attract more probability mass in the middle intervals (Becker et al., 2023). Based on this, I hypothesize that moving from a probabilistic format with predefined intervals to the min–max approach, will result in an upward shift in both means and uncertainty. In other words, this would manifest as a positive β in Eq. (1). Conditional the probabilistic format, the effect of wording on inflation expectations, captured by γ, is less straightforward. If findings for point forecasts summarized in Section 2 extend to probabilistic expectations, the 'prices in general' wording should yield higher mean expected inflation. However, it is not obvious that this is the case, as the effects of wording and response scale could potentially offset each other. Similarly, there is essentially no prior empirical evidence on how wording affects individual uncertainty, leaving the sign and magnitude of γ unclear. In contrast, for the min–max format, implied means are likely higher under the 'prices in general' than under the 'inflation rate' wording, aligning with patterns observed in point forecasts. Due to the higher volatility of food and gas prices, this applies similarly for uncertainty. Hence, δ should be positive and likely larger than β.

The results are reported in Table 2. Columns (1) to (3) document the estimates for implied means, while Columns (4) to (6) - for uncertainty. Columns (2) and (4) report estimates that account for influential observations using Huber weighted regressions. To improve estimation efficiency, Columns (3) and (6) include the following demographic controls: age, education, employment status, household income and size, home ownership, residential region, and residence in East or West Germany prior to 1989.

On average, respondents in treatment A1 expect a mean inflation rate of 2.2%. The estimate for β is large and positive. With roughly 4.6%, mean expected inflation reported in the min–max treatment is more than double the one of the default group. Asking for 'changes in prices in general' increases implied means by 1.3 pp. The joint effect is highly significant, with group D1 (where both wording and format are varied) displaying expectations 4.2 percentage points higher relative to the default group. Changes in wording and format consistently increase the expected mean inflation rate, but not uncertainty. The min–max format yields about 1.1 pp lower uncertainty. In contrast, the 'prices in general' wording raises the standard deviation by about 0.4 percentage points. For group D1, the effect of the min–max format seems to prevail, resulting in a negative overall effect. For either variable, no significant interaction effects are detected except at the 10%-level in Columns (2) and (4).[15] Thus, there is little evidence of a sizeable interaction effect between format and wording beyond the sum of the individual ones. While estimated coefficients vary in magnitude when controlling for outliers in Columns (2) and (4) using Huber weighted regression, they remain statistically and economically significant.

In summary, reporting behavior is highly sensitive to changes in wording and format. The results suggest that individual mean expectations and uncertainty vary strongly with survey framing. In some cases, framing changes can more than double mean

[14] See e.g. https://www.bundesbank.de/en/bundesbank/research/survey-on-consumer-expectations-bop-hh/.

[15] Testing for interaction effects is done by a standard F-test and $H_0 : \delta = \beta + \gamma$.

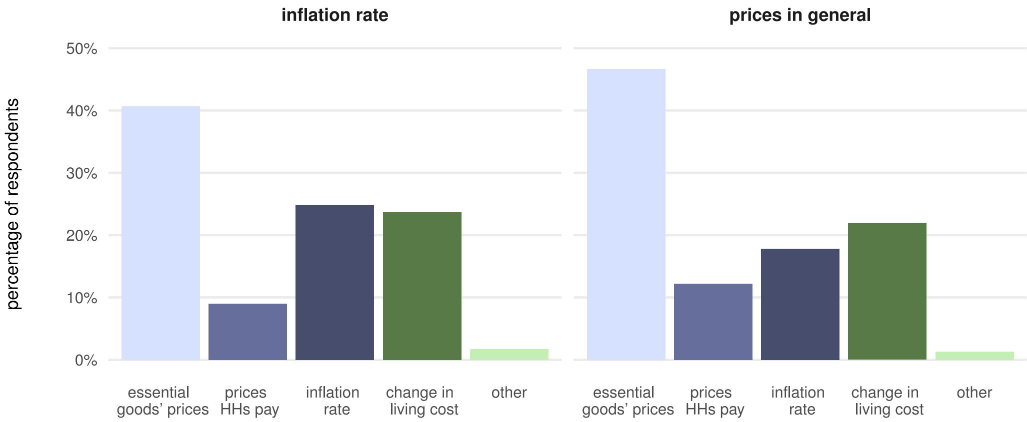


Fig. 2. Effect of wording on forming expectations. Distributions of the topics respondents thought about when answering the question about the 'inflation rate' (left panel) or 'prices in general' (right panel). Each respondent can select one of four pre-defined options or provide an individual answer via the option 'other'.

expectations. I can confirm the initial hypothesis that simpler wording and format generally lead to higher implied means. Contrary to expectations, respondents in the min–max treatment display lower uncertainty on average, even when using the 'prices in general' wording. This suggests that responses from the standard probabilistic question may be artificially more dispersed, possibly due to (i) assumptions about the histogram's support, (ii) strong response scale effects, or both.

Robustness and heterogeneity To illustrate the potential impact of support assumptions, Figures D.4 and D.5 in the Appendix, depict the estimates for implied means and uncertainty for different assumptions on the endpoints of the unbounded intervals. While the estimates for implied means remain practically unchanged, those for uncertainty vary slightly, but not in a statistically significant way.

Furthermore, the effects might vary depending on the underlying socio-demographic population. To this end Tables B.3 and B.4 report the results for different groups by gender, income, education and residence pre-1989. The estimates for implied means suggest that the probabilistic format has particularly strong effect on women and those with no college degree, whereas wording affects most those born in East Germany before 1989. The results for uncertainty are much more nuanced. Another important aspect is whether respondents are part of the panel component of the sample. Kim and Binder (2023) document that repeated survey participation can impact respondents' expectations and uncertainty. To address potential concerns regarding panel conditioning, I estimate the effects based only on the sub-sample of first-time respondents, excluding roughly 700 observations. Again, the results reported in Table B.5 remain practically unchanged and indicate that framing effects are equally strong for both first-time participants as well as panelists.

Overall, the estimates provide substantial evidence that different question formats produce very different mean expectations depending on the underlying population. Significant differences in the min–max format, respectively the lack thereof in the probabilistic one suggest that females and those with no bachelor degree adapt their forecasts more closely to the underlying response scale in the probabilistic setup.

5.3. Effect of wording on forming expectations

Next, I analyze whether different wordings induce thoughts about different underlying prices during the expectation formation process. Of particular interest is, whether respondents associate the 'prices in general' formulation most with price changes of specific goods they observe in their day-to-day shopping. Fig. 2 depicts the distributions of the topics respondents thought about when answering the question about the 'inflation rate' (left panel) or 'prices in general' (right panel).

Overall, the resulting frequencies of topics under the 'prices in general' formulation appear similar to the one reported by Bruine de Bruin et al. (2010) using the same wording, with the topic 'inflation rate', ranked third instead of fourth. The distribution observed in the German sample seems slightly more polarized with almost a majority of respondents thinking about essential goods' prices, compared to 'Prices Americans pay', which was selected by roughly 40% of US survey participants as the main topic.[16] Although for the 'inflation rate' wording, the corresponding share is somewhat smaller (41% vs. 47%), overall a substantial percentage of participants think about supermarket or gas prices when forming their expectations. Even when asked directly for their inflation rate prediction, only about one-quarter of the respondents report actually thinking about it. Considering the additional information people receive at the beginning of the survey about what inflation is and that the questions are continuously accompanied by info boxes, this share is surprisingly small. While the 'inflation rate' wording is indeed able to reduce the share of respondents producing

[16] Note that the question in Bruine de Bruin et al. (2010) is framed differently than the one conducted in BOP-HH. In their setup, one could rank multiple topics, whereas, in the current setting, one could only select one topic. However, due to concerns about survey fatigue, the question implemented in BOP-HH had to be simplified.

Table 3
Effects of question interpretation.

	Dependent variable:					
	$E_i[\pi_{t+1}]$			$\sigma_i^{\pi+12}$		
	(1)	(2)	(3)	(4)	(5)	(6)
default	0.812***	1.105***	0.029	1.814***	1.078***	2.481***
	(0.306)	(0.081)	(0.732)	(0.156)	(0.036)	(0.323)
min–max	2.462***	1.073***	2.613***	−1.013***	−0.619***	−1.007***
	(0.382)	(0.085)	(0.390)	(0.120)	(0.036)	(0.120)
prices	1.165***	0.927***	1.198***	0.364**	0.381***	0.335**
	(0.323)	(0.112)	(0.321)	(0.149)	(0.058)	(0.146)
joint effect	4.152***	2.288***	4.138***	−0.386**	−0.241***	−0.405**
	(0.413)	(0.109)	(0.420)	(0.158)	(0.040)	(0.164)
Thinking about:						
prices of essential goods	1.844***	0.893***	1.256***	0.317*	0.283***	0.164
	(0.358)	(0.092)	(0.382)	(0.163)	(0.034)	(0.168)
prices German HHs pay	1.029**	0.670***	0.382	0.121	0.198***	0.076
	(0.467)	(0.127)	(0.501)	(0.190)	(0.056)	(0.195)
changes in living costs	2.052***	1.273***	1.301***	0.217	0.298***	0.105
	(0.404)	(0.102)	(0.434)	(0.165)	(0.038)	(0.175)
other	0.903	0.706***	0.249	−0.189	0.148**	−0.290
	(0.734)	(0.255)	(0.745)	(0.223)	(0.072)	(0.252)
Observations	3,685	3,685	3,548	3,685	3,685	3,548
Robust linear		✓			✓	
Controls			✓			✓

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Robust standard errors in parentheses. Survey weights are used to ensure representativeness. Columns (1)–(3) present the results for mean expected inflation, Columns (4)–(6) - for uncertainty. 'Robust linear' refers to Huber robust regression specifications, accounting for influential observations. 'Controls' refers to the inclusion of demographic control variables such as age, gender, education, income and others. A full list of control variables is provided in Table B.1 of Appendix B.

a forecast based solely on their personal shopping experience, broader concepts of price changes remain less thought of. Two main conclusions can be drawn from this: (i) diffuse question interpretation continues to be a major source of variation, beyond true differences in beliefs, and (ii) a large share (40%) of respondents base their inflation forecast on specific prices they observe in their day-to-day shopping instead of broader concepts about price changes.

To illustrate how the question interpretation channel affects central measures and uncertainty, Table 3 reports the estimates from a regression controlling for people's reported associations from Question 2. Columns (1) to (3) report the results for implied means, Columns (4) to (6) for standard deviation. The baseline category is defined to be 'thinking about the inflation rate'. The estimates indicate that diffuse interpretations affect primarily implied means rather than uncertainty. Thinking about concepts other than the 'inflation rate', yields up to 2 pp higher mean expected inflation (see Column (1)). The estimates are fairly robust to other specifications as documented in Columns (2) and (3). However, the same does not apply for uncertainty. The lack of significant estimates in Column (4) suggests that socio-demographic variables can explain most of the variation in associations for uncertainty. This is documented in further detail in Table 3 in the Appendix, which lists predictors of selecting 'thinking about the inflation rate' besides wording.

6. Further aspects

This section discusses several practical aspects of response behavior relevant for both researchers and survey designers. In the following I summarize the most important differences across treatments. Table 4 reports the corresponding statistics.

Consistency across quantitative measures As the underlying beliefs about future inflation are inherently unobservable, one possibility of validating the results is to compare the within-subject consistency across treatments. There are several ways to measure the internal consistency of reported answers.

In the upper panel of Table 4 I report the proportion of occurrences when the point prediction does not lie in the range of the support (fitted or self-reported). Conditional on the probabilistic format, the share of such cases is 11% for the 'inflation rate' wording and 14% for the 'prices in general'. This is somewhat surprising given it was expected that the latter would yield higher internal consistency, due to its simplicity. However, it might also be the case that respondents report more extreme values for their point prediction, thinking of food and gas prices, and are thus more prone to subsequently revising their expectations downwards when confronted with the bin response scale. That is, priming effects are higher for the 'prices in general' wording. In contrast, in the min–max setup, this pattern is reversed: It is the 'prices in general' wording that produces answers more in line with peoples' initial point forecasts.

Further, in the upper panel of Table 4, I report the correlation between point forecasts and implied means. With 68%, the combination of the min–max format with the 'prices in general' wording yields the highest correlation among the four options. This

Table 4
Response patterns across treatments.

	probabilistic question		min–max question	
	inflation	*prices*	*inflation*	*prices*
Consistency				
PP not in support	10.3	13.8^{p5}	14.2^{m1}	$8.2^{p1,m1}$
Corr(PP, $E_i[\pi_{t+1}]$)	0.55	0.65	0.45	0.68
$P(PP \in (X_{pp}, Y_{pp}]) = 0$	16.3	15.6		
Contain disjoint regions	3.3	2.4		
Item non-response	6.7	6.6	13.4^{m1}	$9.6^{p1,m5}$
Bin usage				
Sparse histogram	52.2	42.2^{p1}		
At least one outer bin	19.4	26.8^{p1}		
Mean number of bins	3.2	3.5^{p1}		
Using 50%–50% responses	7.7	7.1		
Rounding				
Mode is multiple of 5			19.6	36.1^{p1}
Min is multiple of 5			16.3	24.9^{p1}
Max is multiple of 5			24.3	43.5^{p1}
Observations	944	948	879	915

Note: ^{p1}, ^{p5}, ^{p10} indicate that the corresponding measure is significantly different in the ‘prices in general’ vs. the ‘inflation rate’ wording at the 1, 5, and 10%-level.^{m1}, ^{m5}, ^{m10} indicate that the corresponding measure is significantly different in the ‘min–max’ vs. the ‘probabilistic format’ at the 1, 5, and 10%-level. Shares and probabilities are reported in percentage points. The reported differences are based on χ-squared, Wilcoxon–Mann–Whitney, or Kolmogorov–Smirnov tests, depending on the nature of the underlying variable. X_{pp} and Y_{pp} are the endpoints of the interval that contains the point prediction (PP).

re-iterates the advantage of the min–max that respondents directly state the range of the support, making answers across multiple questions more coherent. This is essential as observations that do not abide by this internal consistency rule are often filtered out of the sample. For the case of the min–max question, the proportion of data that is thrown out would be (almost) cut by half depending on the wording, compared to the probabilistic setup. Similarly to Zhao (2023), I also report the number of occasions where (i) there is zero probability assigned to the interval containing the point forecast and (ii) number of cases with positive probability assigned to non-adjacent bins. Both cases are usually considered undesirable by practitioners and are often excluded from samples. Conditional on the probabilistic setup, neither aspect seems to vary significantly with wording.

Item non-response Despite its greater complexity, the probabilistic question yields a lower non-response rate as indicated in Table 4 and there are no differences across wordings. In the min–max setup, non-response rates are significantly higher for each wording. Similar to the findings of Hayo and Méon (2023) for point forecasts, a possible explanation might be that the probabilistic question is a form of guided question. In essence, the response scale serves as a reference to participants and ‘guides’ them to an answer, they would otherwise have not provided (Hayo and Méon, 2023). That is, respondents evade answers they consider undesirable to the interviewer, researcher, or survey designer (a phenomena called ‘social desirability bias’ (SDB), see also Stantcheva (2023)). The fact that significant differences between wordings arise in the min–max setup, with the ‘inflation rate’ producing a significantly higher non-response share than the ‘prices in general’ additionally supports this argument. While higher response rates are generally desirable in surveys, such answers might simply increase the noise in the data and are less or not at all informative of the underlying true beliefs (Hayo and Méon, 2023).

In the current setting, it is not feasible to precisely determine whether the probabilistic setup introduces noise to the data, which otherwise would have resulted in non-response due to SDB. Some patterns in non-response reported in Table B.8 do provide evidence in support of this argument. More precisely, Table B.8 reports the average marginal effects of wording, and several demographic characteristics on non-response for different question formats. Roughly 80 participants did not provide a point forecast for inflation. Despite the power limitations due to small sample size, the estimation reported in Column (4) of Table B.8 suggests that these respondents are subsequently more likely to provide a forecast in the probabilistic setup than in the min–max setup.

Bin usage More than half of the respondents (52%) in the ‘inflation rate’ group, report histograms with positive probability in at most two bins (sparse histograms). The corresponding proportion for the ‘prices in general’ treatment is significantly lower at 42%. These numbers are somewhat higher compared to other consumer surveys using a probabilistic format such as the SCE (Krüger and Pavlova, 2024). Generally, the ‘prices in general’ wording seems to prompt respondents to use a higher number of bins - 3.5 vs 3.2 for the ‘inflation rate’. Albeit small the difference is statistically significant at the 1%-level. About one-fifth of the people in the ‘inflation rate’ treatment report positive probability in at least one outer bin. For the competing group this amounts to 26.8%. Overall, we can reject the null hypothesis that the distribution of the used number of bins is the same across the two wordings at the 1%-level (see also Figure A.3 in Appendix A). Finally, in both subgroups, the proportion of respondents reporting 50%–50% answers is relatively low at about 7.7% and 7.1%, respectively, and does not differ significantly across wordings.

Rounding It has long been acknowledged that rounding can convey some idea for the respondents’ uncertainty. In the context of consumer inflation expectations, Binder (2017) first introduces a rounding-based uncertainty measure utilizing point forecasts from the Michigan Survey. Additionally, some respondents may resort to rounding to simplify communication (Manski, 2018). Following the latter intuition, rounding might be viewed as an undesirable feature of the data, as it makes interpretation more difficult.

Similarly to Binder (2017), in Table 4, I report the share of minimum, maximum, and mode that are multiples of five in the min–max question. Overall, a considerable share of respondents provides rounded values for the mode with substantial differences between wording choices: 36% for the 'prices in general' vs 20% for the 'inflation rate' wording. This is intuitive given the higher volatility of food and energy prices people think of. Since, to the best of the author's knowledge, the min–max question has not yet been implemented in other large-scale consumer expectation surveys that can be used for comparison, it is difficult to judge whether these shares are unusually high or not. The highest value observed in the data (43.5% for the maximum in group D1) is close to the one observed for point predictions in the Michigan Survey - 41.4% (see Table A.1 of Binder (2017)). This suggests that the usage of the min–max question does not lead to excessive rounding in the data.

7. Conclusion

Directly eliciting consumer expectations about inflation has evolved tremendously in the past few decades. The adoption of a probabilistic format in multiple large-scale consumer surveys has furthered subjective inflation expectations research, deepened our understanding of the underlying formation process, and improved cross-country comparison. Still, several aspects in this context need to be addressed. For once, there appears to be a large framing effect on responses due to the underlying bin definitions which in some cases can act as an anchor for consumers' expectations. Additionally, the assumptions on both the expectation formation process as well as the subsequent processing of the data often appear strong. In particular, the former relies on survey participants holding precise probabilities about future events and being able to convey them in a numerical format.

Even though it yields a higher non-response rate, the min–max question proposed in the experiment provides a viable alternative to the current default format. As pointed out by Hayo and Méon (2023), a guided survey question might potentially introduce more noise to the data instead of collecting informative answers. The same argument may apply to the probabilistic format. The min–max setup is attractive for survey designers due to its simplicity and straightforwardness. The format reduces the priming effects that result from the underlying scale to a minimum as well as eliminates 'problematic' cases such as bi-modal distributions or those with positive probability in disjoint regions. Another important aspect is that the probabilistic format strongly relies on a stable, tractable set of bin definitions over time. This can prove challenging as inflation fluctuates. While changes in bin definitions could cause a break in the time series, as people assign more probability to outer intervals, the assumptions on their bounds become more important. As inflation varies, different choices for these bounds could be justified, thus diminishing comparability over time. While differences in measures derived from different formats should be interpreted with utmost caution, the evidence presented in this paper calls for examining a viable alternative to the probabilistic format. While a large-scale implementation of an alternative question to elicit subjective distributions of future inflation will be costly, using the min–max format as an occasional 'sanity check' for expectations could be advantageous, especially when the data is used for policy evaluation or recommendations.

The results on question wording confirm the findings of previous experimental studies and reiterate the importance of precise question formulation. Even a direct formulation such as the 'inflation rate' seems insufficient to invoke thoughts about inflation among the majority of consumers. More so, roughly 40% of the survey participants still mainly think about specific prices such as those of food and gas when producing their forecast. A potential avenue for future research could be to explore how to reduce this share via wording or providing respondents with additional information.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Lora Pavlova reports financial support was provided by German Research Foundation. The data used in this paper are proprietary. They were obtained from the Deutsche Bundesbank. Access to the data used can be obtained following the requirements of the Research Data and Service Centre of the Deutsche Bundesbank. Further information on the application process for the data can be found at: https://www.bundesbank.de/en/bundesbank/research/rdsc/research-data/bop-hh-757542.

Acknowledgments

I gratefully acknowledge financial support from the German Science Foundation (DFG) via grant KR 5214/1-1. I would also like to thank the participants of the 10th HKMetrics Workshop, Workshop on Economic Expectations, Survey Data, and Forecasting at Heidelberg University, NBB Workshop on Macroeconomics and Survey Data, 44th International Symposium on Forecasting, CNB Workshop on Households' Inflation Expectations, as well as Tabea Bucher-Koenen, Christian Conrad, Zeno Enders, Alexander Glas, Malte Knüppel, Fabian Krüger, Stefan Trautmann, Dmitri Vinogradov, Michael Weber, and the anonymus referee for their helpful comments and suggestions. I am grateful to the BOP-HH team for including the proposed questions in the survey. This paper uses data from the Bundesbank-Online-Panel-Households. The results published and the related observations and analysis may not correspond to the results or analysis of the data producers. The data used in this paper are proprietary and were obtained from the Deutsche Bundesbank. Access to the data used can be obtained following the requirements of the Research Data and Service Centre of the Deutsche Bundesbank. Further information on the application process for the data can be found at: https://www.bundesbank.de/en/bundesbank/research/rdsc/research-data/bop-hh-757542. The trial was pre-registered at the American Economic Association RCT Registry at https://www.socialscienceregistry.org/trials/6482.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jebo.2025.106899.

Data availability

The authors do not have permission to share data.

References

Altig, D., Barrero, J.M., Bloom, N., Davis, S.J., Meyer, B., Parker, N., 2022. Surveying business uncertainty. J. Econometrics 231 (1), 282–303, Annals Issue: Subjective Expectations & Probabilities in Economics.

Armantier, O., Bruine de Bruin, W., Potter, S., Topa, G., van der Klaauw, W., Zafar, B., 2013. Measuring Inflation Expectations. Annu. Rev. Econ. 5 (1), 273–301.

Armantier, O., Topa, G., van der Klaauw, W., Zafar, B., 2017. An overview of the Survey of Consumer Expectations. Fed. Reserv. Bank N. Y. Econ. Policy Rev. 23, 51–72.

Bachmann, R., Carstensen, K., Lautenbacher, S., Schneider, M., 2021. Uncertainty and change: survey evidence of firms' subjective beliefs. Working Paper 29430, National Bureau of Economic Research.

Becker, C., Duersch, P., Eife, T., 2023. Measuring Inflation Expectations: How the Response Scale Shapes Density Forecasts. AWI Discussion Paper Series No. 727, Heidelberg University.

Beckmann, E., Schmidt, T., 2020. Bundesbank Online Pilot Survey on Consumer Expectations. Technical Paper 01/2020, Deutsche Bundesbank.

Binder, C.C., 2017. Measuring uncertainty based on rounding: New method and application to inflation expectations. J. Monetary Econ. 90, 1–12.

Boctor, V.R., Coibion, O., Gorodnichenko, Y., Weber, M., 2024. On eliciting subjective probability distributions of expectations. Working Paper 32406, National Bureau of Economic Research.

Brandão-Marques, L., Gelos, G., Hofman, D., Otten, J., Pasricha, G.K., Strauss, Z., 2023. Do household expectations help predict inflation?. IMF Working Paper 2023/224, International Monetary Fund.

Bruine de Bruin, W., Fischhoff, B., Millstein, S.G., Halpern-Felsher, B.L., 2000. Verbal and numerical expressions of probability:"It's a fifty–fifty chance". Organ. Behav. Hum. Decis. Process. 81 (1), 115–131.

Bruine de Bruin, W., van der Klaauw, W., Downs, J.S., Fischhoff, B., Topa, G., Armantier, O., 2010. Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. J. Consum. Aff. 44 (2), 381–402.

Bruine de Bruin, W., van der Klaauw, W., Topa, G., Downs, J.S., Fischhoff, B., Armantier, O., 2012. The effect of question wording on consumers' reported inflation expectations. J. Econ. Psychol. 33 (4), 749–757.

Bruine de Bruin, W., van der Klaauw, W., van Rooij, M., Teppa, F., de Vos, K., 2017. Measuring expectations of inflation: Effects of survey mode, wording, and opportunities to revise. J. Econ. Psychol. 59, 45–58.

Cavallo, A., Cruces, G., Perez-Truglia, R., 2017. Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. Am. Econ. J.: Macroecon. 9 (3), 1–35.

Christelis, D., Georgarakos, D., Jappelli, T., van Rooij, M., 2020. Trust in the Central Bank and Inflation Expectations. Int. J. Central Bank. 16 (6), 1–37.

Coibion, O., Georgarakos, D., Gorodnichenko, Y., Kenny, G., Weber, M., 2024. The effect of macroeconomic uncertainty on household spending. Amer. Econ. Rev. 114 (3), 645–677.

Coibion, O., Gorodnichenko, Y., Kumar, S., Pedemonte, M., 2020. Inflation expectations as a policy tool? J. Int. Econ. 124, 103–297.

Curtin, R., 1996. Procedure to estimate price expectations. Mimeo. University of Michigan Survey Research Center.

D'Acunto, F., Hoang, D., Paloviita, M., Weber, M., 2020. Effective policy communication: Targets versus instruments. KIT Working Paper Series in Economics No. 147, Karlsruher Institut für Technologie (KIT), Institut für Volkswirtschaftslehre (ECON), Karlsruhe.

D'Acunto, F., Malmendier, U., Ospina, J., Weber, M., 2021. Exposure to grocery prices and inflation expectations. J. Polit. Econ. 129 (5), 1615–1639.

D'Acunto, F., Malmendier, U., Weber, M., 2023. What do the data tell us about inflation expectations? In: Handbook of Economic Expectations. Elsevier, pp. 133–161.

D'Acunto, F., Weber, M., 2024. Why survey-based subjective expectations are meaningful and important. Annu. Rev. Econ. 16, 329–357.

Das, S., Kuhnen, C.M., Nagel, S., 2020. Socioeconomic status and macroeconomic expectations. Rev. Financ. Stud. 33 (1), 395–432.

Dominitz, J., Manski, C.F., 1997. Using expectations data to study subjective income expectations. J. Amer. Statist. Assoc. 92, 855–867.

Dräger, L., Lamla, M.J., 2024. Consumers' macroeconomic expectations. J. Econ. Surv. 38 (2), 427–451.

Engelberg, J., Manski, C.F., Williams, J., 2009. Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters. J. Bus. Econom. Statist. 27, 30–41.

Fischhoff, B., Bruine de Bruin, W., 1999. Fifty–fifty=50%? J. Behav. Decis. Mak. 12 (2), 149–163.

Goldfayn-Frank, O., Kieren, P., Trautmann, S., 2024. A choice-based approach to the measurement of inflation expectations. AWI Discussion Paper Series No. 742, Heidelberg University.

Goldfayn-Frank, O., Wohlfart, J., 2020. Expectation formation in a new environment: Evidence from the German reunification. J. Monetary Econ. 115, 301–320.

Hayo, B., Méon, P.-G., 2023. Measuring household inflation perceptions and expectations: The effect of guided vs non-guided inflation questions. J. Macroecon. 78, 103558.

Huisman, R., Van der Sar, N.L., Zwinkels, R.C., 2021. Volatility expectations and disagreement. J. Econ. Behav. Organ. 188, 379–393.

Hurd, M.D., 2009. Subjective probabilities in household surveys. Annu. Rev. Econ. 1, 543–562.

Kim, G., Binder, C., 2023. Learning-through-survey in inflation expectations. Am. Econ. J.: Macroecon. 15 (2), 254–278.

Krüger, F., Pavlova, L., 2024. Quantifying subjective uncertainty in survey expectations. Int. J. Forecast. 40 (2), 796–810.

Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. Q. J. Econ. 131 (1), 53–87.

Manski, C.F., 2018. Survey measurement of probabilistic macroeconomic expectations: progress and promise. Natl. Bur. Econ. Res. Macroecon. Annu. 32, 411–471.

Manski, C.F., Molinari, F., 2010. Rounding probabilistic expectations in surveys. J. Bus. Econom. Statist. 28 (2), 219–231.

Nagel, J., 2022. Return of inflation: What do we learn from survey data?. URL https://www.bundesbank.de/en/press/speeches/return-of-inflation-what-do-we-learn-from-survey-data--893196. Speech at the Joint Spring Conference on Monetary Policy and Expectations of Households and Firms (last accessed: October 14, 2022).

Phillot, M., Rosenblatt-Wisch, R., 2024. Order matters: An experimental study on how question ordering affects survey-based inflation forecasts. Int. J. Central Bank. 20 (3), 63–114.

Ryngaert, J.M., 2023. Balance of risks and the anchoring of consumer expectations. J. Risk Financ. Manag. 16 (2), 79.

Stantcheva, S., 2023. How to run surveys: A guide to creating your own identifying variation and revealing the invisible. Annu. Rev. Econ. 15, 205–234.

Weber, M., Candia, B., Afrouzi, H., Ropele, T., Lluberas, R., Frache, S., Meyer, B.H., Kumar, S., Gorodnichenko, Y., Georgarakos, D., et al., 2023. Tell me something I don't already know: Learning in low and high-inflation settings. Working Paper 31485. National Bureau of Economic Research.

Weber, M., D'Acunto, F., Gorodnichenko, Y., Coibion, O., 2022. The subjective inflation expectations of households and firms: Measurement, determinants, and implications. J. Econ. Perspect. 36 (3), 157–184.

Zhao, Y., 2023. Internal consistency of household inflation expectations: Point forecasts vs. density forecasts. Int. J. Forecast. 39 (4), 1713–1735.