// NO.24-036 | 06/2024

DISCUSSION PAPER

// LORA PAVLOVA

Framing Effects in Consumer Expectations Surveys

Framing Effects in Consumer Expectations Surveys

Lora Pavlova

ZEW - Leibniz Centre for European Economic Research*

Abstract

In a randomized experiment embedded in a survey, I test the effects of variations in question wording and format on consumer response behavior and the corresponding inflation expectations. To this end, survey participants from a representative sample of German consumers are broken down into four treatment groups and presented with different versions of a question asking for their subjective distribution for inflation over the next 12 months. As part of the experiment, two competing wordings, previously known from leading consumer surveys, are considered: (i) the change in prices in general or (ii) the inflation rate. In addition, I compare the responses to a question asking for consumers' probabilistic beliefs about future inflation, to those from a simpler one asking for the expected minimum, maximum, and most likely inflation rate over the short term. I find that response behavior varies strongly with framing. Simpler wording such as 'prices in general' and a less restrictive format lead to higher mean expected inflation, on average. While simpler wording increases individual uncertainty derived from the subjective histograms, asking for minimum, maximum and mode leads to lower uncertainty about expected inflation.

Keywords: probabilistic expectations, survey design, household inflation expectations

JEL classification: C83, D84, E31

*Contact address: ZEW - Leibniz Centre for European Economic Research, L 7, 1, 68161 Mannheim, Germany. Email address: lora.pavlova@zew.de. 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 10^{th} HKMetrics Workshop, Workshop on Economic Expectations, Survey Data, Forecasting at Heidelberg University, NBB Workshop on Macroeconomics and Survey Data, as well as Christian Conrad, Alexander Glas, Fabian Krüger, Stefan Trautmann, and Michael Weber for their helpful comments and suggestions. This paper uses data from the Bundesbank-Online-Panel-Households. The results published and the related observations and analysis may not correspond to results or analysis of the data producers.

1 Introduction

In times of economic turmoil, it has become increasingly important that policymakers receive the most reliable and timely information on consumers' expectations about the future development of both macroeconomic and idiosyncratic outcomes. In the context of the broader public, one of the most effective ways of eliciting expectations in economics is directly surveying consumers. Multiple central banks and research institutions conduct such surveys[1], and this practice has become more widespread in recent years. 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 in the transmission mechanism of monetary policy. Policymakers also carefully monitor inflation expectations for early signs of erosion of central bank credibility and de-anchoring of expectations that can lead to dangerous price-wage spirals (Nagel, 2022).

For decades, it has been common practice to ask consumers for a value of the expected 'change in prices in general'. A leading example is the Reuters/ University of Michigan Survey of Consumers (henceforth Michigan Survey), which elicits consumer expectations about price changes in the form of point forecasts since the late 1970s (Curtin, 1996). In a more recent survey, the Federal Reserve Bank of New York (FRBNY) adopted an alternative question wording that directly asks consumers about the expected inflation rate. This change aims for a more precise formulation to reduce ambiguity and variation in responses that originate from different interpretations rather than true differences in beliefs (Armantier et al., 2017). However, this comes at the cost of a potentially higher degree of complexity. Therefore, the effect of a change in wording might not be symmetric across demographic groups. While more educated and financially literate people may adjust their answers to better conform with economic terminology, for their counterparts this might lead to a higher non-response rate (D'Acunto et al., 2020).

Together with a new wording, the FRBNY also introduced a new type of question to its survey, previously mostly utilized in surveys of professional forecasters such as the ECB and US SPF. While point predictions provide a view of the inflation rate expected 'on average', they contain no information about the individual uncertainty associated with the forecast. One possibility to capture uncertainty is to elicit the whole subjective distribution of the respondent[2]. This is done by asking participants to assign probabilities to a set of non-overlapping intervals representing possible outcome ranges for the variable of interest (subjective histogram). While highly informative, this format also suffers several disadvantages. For once, it strongly relies on the respondent's capabilities of expressing herself via probabilities, which in turn requires a certain degree of numeracy and sophistication. For example, in the context of open-ended probability questions, studies document a high proportion of 50% answers as well as an over-proportional usage of 0 and 100%, also known as 'focal point responses' (Hurd, 2009; Dominitz and Manski, 1997). Stating 50% might indicate high epistemic uncertainty among respondents and signal that they struggle with numbers and probabilities, or to form a distribution about the expected outcomes of an unknown concept

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

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

(Fischhoff and Bruine de Bruin, 1999; Bruine de Bruin et al., 2000). In contrast, in the context of probabilistic expectations, stating 50-50% can be interpreted as the respondent being quite certain of the outcome for the variable of interest. Even if people have precise probabilities about future outcomes in mind, they may resort to rounding to facilitate communication (Manski and Molinari, 2010). A more recent study by Becker et al. (2023) provides some experimental evidence of the large influence of the underlying response scale on the reported distributions and subsequently derived inflation expectations and uncertainty measures.

Using a nationally representative consumer expectations survey in Germany, this paper estimates the causal effects of a change in question wording and format on the reported inflation expectations, both jointly and separately. To the best of the author's knowledge, this is the first study to examine these aspects in the context of consumer probabilistic expectations. I find that response behavior varies strongly with framing. Substituting the current probabilistic question with a simpler one asking for the minimum, maximum, and mode leads to an increase in mean expected inflation by more than 1 pp. In contrast, the individual uncertainty of the respondent measured by the standard deviation of their distribution, is estimated to be 0.6 to 1 pp smaller, suggesting that standard assumptions on the endpoints of the outer intervals might be less precise. As to the effect of wording on reported short-run inflation expectations, I can largely confirm findings from previous literature that the expectations of those provided with the *inflation rate* wording seem less upward-biased and more concentrated around 2%. 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 producing their forecast.

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 of Consumers (Curtin, 1996). Bruine de Bruin et al. (2012) and Bruine de Bruin et al. (2017) study, among others, the effect of wording on the central tendency and disagreement of inflation expectations and perceptions, and report differences between wording versions such as the 'inflation rate', 'prices in general' or 'prices you pay'. 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 (2018) 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. Becker et al. (2023) focus exclusively on how the response scale shapes expectations derived from subjective histograms of households and provide evidence for substantial sensitivity to the underlying bin definitions. Another recent paper closely related to the current study is the one by Hayo and Méon (2022), in which the authors assess the effect of guided vs. non-guided questions on the reported inflation expectations and non-response rates.

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

2 An Experiment on Framing Effects

The current study aims to test framing effects when eliciting responses about short-term inflation expectations of a representative sample of consumers in Germany.[3] Section 2.1 and Section 2.2 discuss recent practises and lay out how the effects of format and wording are analyzed within the experimental framework. Section 2.3 discusses an additional part of the experiment aimed at shedding more light on the process of forming expectations.

2.1 Format

While the Michigan Survey asks about the expected rate of price changes, or in other words a point prediction, another important component of respondents' expectations is their uncertainty. Instead, one could elicit the participant's whole subjective distribution of possible inflation outcomes over the horizon of interest. In 2013, the FRBNY launched a large-scale consumer survey on expectations - 'The Survey of Consumer Expectations' (henceforth SCE) - and introduced a probabilistic question format for the rate of inflation, average home price, and personal wage over different horizons (Armantier et al., 2017). In the following years, several central banks followed suit and adopted this question format in further large consumer expectation surveys, among others the Bank of Canada, Deutsche Bundesbank, and the European Central Bank. For example, the question about short-term inflation expectations from the SCE is specified 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 inflation will be between 4% and 8%

the rate of inflation will be between 2% and 4%

the rate of inflation will be between 0% and 2%

the rate of deflation (opposite of inflation) will be between 0% and 2%

the rate of deflation (opposite of inflation) will be between 2% and 4%

the rate of deflation (opposite of inflation) will be between 4% and 8%

the rate of deflation (opposite of inflation) will be between 8% and 12%

the rate of deflation (opposite of inflation) will be 12% or higher

[3] The trial was pre-registered at the American Economic Association RCT Registry at `https://www.socialscienceregistry.org/trials/6482`.

In essence, respondents are asked to assign probabilities based on their beliefs to different ranges ('bins') of possible inflation outcomes such that they sum up to 100%. Such a type of question allows for computing a measure of individual uncertainty such as the variance or IQR, or some other measure of the spread after fitting a continuous distribution to the discrete histogram (Engelberg et al., 2009).[4]

However, an important caveat is that the collected data highly depends on both people's preference to think and their ability to express their expectations and beliefs using numerical probabilities (Manski, 2018). While an overall willingness to convey expectations in a probabilistic manner has been exhibited by respondents (Armantier et al., 2013), several artifacts have been observed in the data so far. For instance, Bruine de Bruin et al. (2000) find that when asked to estimate the risk of an event happening and report the corresponding probability, uncertain respondents often resort to using the number 50.[5] More generally, the use of round numbers when reporting inflation expectations is found to be a proxy for subjective uncertainty (Binder, 2017), but can also be used by consumers to facilitate communication (Manski, 2018). Based on the raw data alone, it is unclear which one applies.[6]

Another important aspect is shown in more recent work by Becker et al. (2023), where they investigate the effects of changes in the response scale on density forecasts. 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, and hence leads to biased answers.[7] This artifact in the data has the potential to overstate the degree of anchoring to the inflation target or any other value for that matter, depending on how the bin definitions 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, according to 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.

Using an experiment embedded in a representative online survey among consumers in Germany, I test the effects of using an alternative format on the resulting subjective distributions of expected inflation. To allow for eliciting a subjective distribution of the variable of interest, while minimizing cognitive load and response scale bias, in the experiment I include a question on inflation expectations which is completely free of probabilities. The question, which I will henceforth refer to as 'min-max', is 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?

[4]While highly informative, the standard method of quantification needed to compute an uncertainty measure often imposes rather restrictive assumptions on the underlying subjective distribution and is sensitive to small changes in the reported probabilities (see Krüger and Pavlova (2023) for a more detailed discussion).

[5]The answer can be a sign that the participant struggles to express their feelings as a number or experiences high epistemic uncertainty, rather than an intended use of 50% (Fischhoff and Bruine de Bruin, 1999).

[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).

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 and presented with different versions of a question asking for their subjective distribution of expected inflation over the next twelve months. The treatment arms are as follows:

- Group A1: probabilistic question about the inflation rate (henceforth 'default')
- 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

The idea behind this treatment is to compare the subjective distributions that result from the two question formats, but also analyze how they interact with variations in wording. While the question above is more limited regarding the functional form of the distribution, it has the advantage that respondents can explicitly state the ends of the support of their histogram making additional assumptions unnecessary. The fact that respondents are not presented with any numbers or ranges beforehand reduces the possibility of them altering their answers to better fit the survey setting, e.g. by placing probability symmetrically around zero or assigning more probability on lower numerical values if their point prediction lies in an outer bin. By design, Question 1 also reduces the share of 'problematic cases' in the data such as responses that include one or more disjoint regions with positive probability mass or exhibit bi-modality. While such cases are not dominant in household expectations data, they do represent a non-trivial share[8] and are more difficult to handle by standard methodology and therefore excluded. Overall, the degree of complexity of the question is reduced and the information respondents receive beforehand is minimized.

Questions similar in design have been implemented in the context of households (Coibion et al., 2022; Christelis et al., 2020) and firms (Altig et al., 2022), however, in either case, the respondents are asked to additionally assign a probability to a corresponding 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 inflation expectations, where in the first step respondents are also asked for the expected minimum and maximum outcome.

[8]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.

2.2 Wording

Manski highlights the "importance of careful attention to question wording when eliciting expectations" via surveys (Manski, 2018, p. 440). Longer-running surveys such as the Michigan Survey ask participants about their expectations about the 'change in prices in general'. While simpler in design, the responses to this question have frequently exhibit extreme values, overstating realized inflation, and display large disagreement (see D'Acunto et al. (2023) for an overview and discussion). While variation in inflation expectations can be, at least partially, attributed to factors related to personal experiences[9], it can also stem from differences in the question interpretation. That is, whether people *think* about person-specific financial experiences or price changes on a broader scale such as those reflected in the consumer price index. Bruine de Bruin et al. (2010) report that people tend to interpret the 'prices in general' wording as asking most for expectations of specific prices such as those of food and gasoline. On the other hand, not all respondents may have a clear understanding of what economic terms such as the inflation rate, used in more recent surveys such as the SCE, entail. Hence, this may 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 have indeed shown that question wording significantly impacts elicited point forecasts. Using a representative US sample, Bruine de Bruin et al. (2012) document that both means and medians of the aggregate point predictions distribution from a question using a 'prices in general' or 'prices you pay'-wording exceed those produced by an 'inflation rate'-wording. This also holds true for respondents' perceptions of inflation. While the authors find almost equal non-response rates across wordings, they do report participants having more difficulty understanding terms such as the 'inflation rate'.[10]

While the majority of studies has focused on the effect of wording in the context of (aggregate) point forecast distributions of households, little is known whether these effects will persist in the probabilistic set-up now widely utilized across consumer surveys. It can well be the case that the biasing effect of thinking about specific prices is mitigated by the features of the probabilistic question format. For example, reporting extreme values caused by the 'prices in general' wording could be offset by the dampening effect of the response scale in the probabilistic question.

Therefore in the experiment, I test the effect of using simpler wording such as 'prices in general' on the resulting subjective distributions of consumers. More precisely, I extend the experimental setup used in Bruine de Bruin et al. (2010) and Bruine de Bruin et al. (2012) to the context of probabilistic expectations. Each two sub-samples receive the *inflation rate* (A1, B1) or *change in prices in general* (C1, D1) wording. Earlier experiments such as Bruine de Bruin et al. (2012) include a third alternative, namely '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.

[9]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).

[10]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.

2.3 Forming expectations

Furthermore, 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), where it was successfully implemented as part of a web-based survey among members of the RAND's American Life Panel. Respondents are presented with five options they can choose from, whereby one 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). The question is as follows:

Question 2: What did you think about most when answering the question about your inflation rate expectations before?[11]

- Prices you pay in your everyday life such as food and gasoline
- Prices Germans pay
- Germany's inflation rate
- Changes in the cost of living
- Other specific prices (*please name*)

One can expect that people who were randomly assigned to the *prices in general* treatment arm, should be more likely to select 'everyday prices'. Accordingly, if interpreted correctly, respondents from the *inflation rate* treatment group, should most often select the corresponding option - 'Germany's inflation rate'. Furthermore, it would be important to determine whether there is any systematic heterogeneity by demographic characteristics, e.g. whether populations with lower financial literacy think more often of personal experiences, while simultaneously controlling for a wording effect.

3 Data Set

The experiment was conducted as part of the 9th wave of the Bundesbank Online Panel - Households (henceforth 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 sample size varies from 2,000 to 5,000 respondents and contains a panel component. The survey is representative of the German online population aged 16 and above. In addition to the core and project-specific answers, information on the respondent's age, education, employment, income, household size, the number of children, and others is also provided. For further information on the survey see Beckmann and Schmidt (2020).

[11] Again, the wording is adjusted accordingly for groups C1 and D1. For the exact question wording see Appendix C.

In the several months prior to the September wave, Germany experienced low inflation rates[12] This was also reflected in the short-run inflation expectations of BOP-HH respondents, which after an upward shift at the beginning of the COVID-19 pandemic, gradually subsided to pre-pandemic levels in the autumn of 2020. While expectations remained relatively stable, uncertainty over the first year of the pandemic rose significantly.[13] Such a phenomenon was also observed in other consumer expectation surveys such as the SCE, shown for example in Armantier et al. (2021).

The September 2020 wave of BOP-HH, in which the experiment was conducted, has a sample size of roughly 4,000 respondents. Thus, in each of the four treatment arms, there are about 1,000 participants. Table B.1 in Appendix B shows the distribution of selected socio-demographic characteristics across the four treatment groups and Table B.2 reports the probabilities of being selected in each treatment arm based in the respondent's characteristics. Respondents who receive the *inflation rate* wording are subject to an information treatment, established in previous waves of the survey. More precisely, they are shown the following short definition of what inflation is: *'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.'*

As part of the core module of the survey, participants receive a number of questions specifically targeted at their inflation expectations.[14] First, they are asked about their perception of the inflation rate development over the last 12 months. Then, regarding the upcoming 12 months in the following order respondents report whether they expect inflation or deflation, and if so, at how much percent (point prediction). Finally, depending on the treatment group, they are presented with either the probabilistic or the question asking for the mode, minimum, and maximum. For the probabilistic question, participants are additionally instructed that the sum across probabilities should be one hundred. As they insert their answers, the current sum of the probabilities is displayed. If one attempts to skip the question, two options of non-response are shown: 'Don't know' and 'No answer'. All relevant questions on inflation expectations are documented in detail in Appendix C.

For fitting a distribution to the subjective histograms produced by the probabilistic question I broadly 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 the fact that the method was initially designed for bins of equal width, which is not the case in the current setup[15]. Following Armantier et al. (2017) I impose an upper bound of the open intervals at ± 38.

In the experiment, the other half of the participants are presented with the min-max

[12]The monthly, year-on-year inflation rate for June, July, and August 2020 was at 0.8, 0.0, and -0.1%. Source: Statistisches Bundesamt (Destatis).

[13]Summary statistics on inflation expectations and uncertainty are published monthly at https://www.bundesbank.de/en/bundesbank/research/survey-on-consumer-expectations/inflation-expectations-848334.

[14]A detailed overview of the survey questions is documented in Appendix C. Additionally, 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.

[15]See Krüger and Pavlova (2023) and the associated software available at https://github.com/FK83/forecasthistogram.

question. In essence, by answering Question 1 the respondent reports both ends of the support of their subjective distribution, a and b, as well as the mode of the distribution, c. Using this information one can fit a simple triangular distribution to the subjective probabilities, which is a common practice in empirical literature.[16] Loosening the symmetry assumption on the form of the distribution allows for more flexibility compared to the standard practice. It follows that the mean and the variance of the distribution are:

$$\mathrm{E}[\pi_{t+12}] = \frac{a+b+c}{3}$$
$$\sigma^2_{\pi_{t+12}} = \frac{a^2+b^2+c^2-ab-bc-ac}{18}$$

Excluding respondents who did not provide an answer leaves 1,892 responses for the probabilistic question and 1,794 for the min-max question. Generally, one would like to further leave out observations that either (i) do not sum up to 100% for the probabilistic question or (ii) have a reported mode outside of bounds in the min-max setup. Such cases of inconsistency do not occur in the data. In the following, I provide some descriptions and summary statistics on the elicited expectations data.

4 Effects of Wording and Format on Consumers' Inflation Expectations

This section presents the empirical results. First, I briefly discuss the average values for fitted moments, percentiles, and endpoints across the different treatment arms in Section 4.1. Section 4.2 lays out the estimated effects on mean expected inflation and individual uncertainty. In Section 4.4 I analyze how the choice of wording impacts the collected data via the question interpretation channel. Section 4.3 exploits additional socio-demographic variables to determine, whether the effect is heterogeneous across various groups. Section 4.5 concludes with a brief comparison of point forecasts and perceptions by wording.

4.1 Derived Measures

Table 1 reports the average values for the distributions of fitted moments, endpoints, and percentiles across all four treatment arms. I observe that both the left endpoint and the p_5 are (often) negative on average for the probabilistic setup, whereas for the min-max question, both of them are positive and even slightly above 2%. While there is little difference in the p_{95}, the one in the right endpoint is striking, it is about $2-3$ pp larger in the probabilistic format than in the min-max setup. It seems that irrespective of the question format, the *prices in general* wording produces (i) higher expected values for the inflation rate, and (ii) higher uncertainty in terms of the standard deviation. The distributions of the fitted means are additionally plotted in Figure A.1 in the Appendix. In all four cases, across wording and format variations I can reject the null hypothesis of a two-sample Kolmogorov-Smirnov test that the distributions are equivalent at the 1%-level. Nonetheless, caution is advised when interpreting these results, since they are very sensitive to the underlying assumptions

[16] It is common to fit an isosceles triangle distribution, which is a special case of the triangular distribution (Engelberg et al., 2009).

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

	$E[\pi_{t+12}]_i$	$\sigma_i^{\pi_{t+12}}$	left endpoints	right endpoints	$p5$	$p25$	$p50$	$p75$	$p95$
			probabilistic question						
A1: *inflation rate*	2.09	1.84	−3.31	8.49	−0.87	0.74	2.06	3.41	5.14
C1: *prices in general*	3.50	2.07	−2.25	11.13	0.22	1.96	3.44	4.99	6.97
			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

on the support of the fitted distribution for the case of the probabilistic question. Another possibility in the quantification procedure is to use the highest and lowest reported point forecasts as endpoints for the histograms. While feasible, it is not clear that this approach is more desirable than the one used in the current paper or others, for that matter, as assumptions on the support will change with each survey wave and diminish comparability over time. This contributes to the argument that the min-max question has an advantage over the probabilistic setup, namely, the lack of necessity for assumptions on the support of the histogram, thus ensuring comparability over time and across studies.

4.2 ATEs on Individual Mean Expectations and Uncertainty

Given the experimental setup described in Section 2, one can estimate the treatment effects of a change in wording and format in a simple linear regression framework as follows:

$$y_i = \alpha + \beta \text{format}_i^{min-max} + \gamma \text{wording}_i^{prices} + \delta \text{joint}_i + \varepsilon_i, \quad (1)$$

where y_i is the inflation expectations measure of interest of person i. In the study, I focus mostly on measures of central tendency and uncertainty such as $E[\pi_{t+12}]_i$ and $\sigma_i^{\pi_{t+12}}$ derived from the reported subjective histograms. α represents the respective measure for group A1 which receives the default version of the question as described in Section 2.1. The remaining terms on the right-hand side capture the mean expected inflation (or uncertainty) for the other three treatment groups relative to the default group. That is, the variables $format_i^{min-max}$, $wording_i^{prices}$, and $joint_i$ are dummies taking unit value when respondent i is assigned to group B1, C1 or D1, respectively.

In an ideal scenario where wording and format do not influence response behavior, the resulting four distributions should be identical across measures. While more precise in design, a pure probabilistic question causes respondents to place more probability mass in the middle intervals (Becker et al., 2023). Therefore, I conjecture that moving from a probabilistic formulation with predefined intervals to a less restrictive setup such as the min-max, one should observe an upward shift in expectations coupled with a larger variation in the individual responses, or put differently a positive β in Equation (1). As this should be true irrespective of the wording choice, one can expect a positive δ as well.

Table 2: ATEs on individual mean inflation expectations and uncertainty

	Dependent variable:					
	$E[\pi_{t+12}]_i$			$\sigma_i^{\pi+12}$		
	(1)	(2)	(3)	(4)	(5)	(6)
default	2.193***	1.838***	1.746**	2.011***	1.286***	2.806***
	(0.217)	(0.062)	(0.720)	(0.099)	(0.028)	(0.307)
min-max	2.459***	1.076***	2.614***	−1.018***	−0.620***	−1.008***
	(0.381)	(0.086)	(0.393)	(0.120)	(0.033)	(0.122)
prices	1.285***	0.984***	1.338***	0.387***	0.406***	0.367**
	(0.326)	(0.107)	(0.321)	(0.147)	(0.055)	(0.144)
joint effect	4.239***	2.334***	4.221***	−0.373**	−0.230***	−0.389**
	(0.422)	(0.107)	(0.428)	(0.158)	(0.038)	(0.163)
Observations	3,686	3,686	3,549	3,686	3,686	3,549
Robust linear		X			X	
Controls			X			X

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.

Conditional on the probabilistic format, the effect of wording on inflation expectation measures captured by γ is less clear. If previous findings for point forecasts as summarized in Section 2.2 can be translated to probabilistic expectations, then one should expect higher mean expected inflation in the 'prices in general' wording. However, it is not obvious that this is the case, e.g. due to the effects of wording and response scale biasing expectations in opposite directions and eventually offsetting each other. The same applies to individual uncertainty for which there is essentially no prior empirical evidence, to the best of the author's knowledge. Finally, combining simpler wording and format in treatment group D1 further allows for estimating potential interaction effects that go beyond the separate effects of wording and format.

Based on Equation (1), I run a simple linear regression to estimate the average treatment effects (ATEs) of changes in wording and format for different measures derived from probabilistic responses. Columns (1) to (3) of Table 2 report the estimated coefficients for mean expected inflation, while Columns (4) to (6) present the results for uncertainty, derived as described in Section 3. 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 from treatment arm A1 expect an inflation rate of roughly 2.2%. However, adopting the min-max format more than doubles people's mean expected inflation. Substituting current wording with the *prices in general* one adds an additional 1.3 pp to the already elevated expectations. The joint effect is highly significant meaning the expectations of the final treatment group, where I vary both wording and format exceed the one of the default by 4.2 pp. However, it rarely exceeds the sum of the other two coefficients on format and wording. When testing formally whether combining the two treatments in group D1 has a larger effect than implementing them separately, one cannot reject the null hypothesis, except in specification (2), where the lack of difference can be rejected at the 10%-level. This means that there is little evidence for a sizeable interaction effect beyond the separate estimates for format and wording. While the estimated coefficients are prone to changes in magnitude as soon as I control for the presence of outliers using Huber weighted regression in specification (2), they remain statistically and economically significant.

While the changes in wording and format always lead to an increase in the predicted inflation rate on average, that is not the case for individual uncertainty measured by the standard deviation. The usage of the *prices in general* wording leads to significantly higher uncertainty. Whereas a change in wording increases uncertainty by roughly 0.4 pp, switching to a min-max format reduces it by 0.6 to 1 pp, depending on the regression specification reported in Columns (4) to (6) of Table 2. Regarding the interaction of the two in treatment group D1, the influence of the question format seems to prevail and the overall estimated effect is negative. In column (4), one can reject the null hypothesis that the interaction term equals the sum of the remaining two treatment effects at the 10%-level. Again, controlling for outliers in column (5) of Table 2 reduces the absolute value of the coefficients but does not change their sign or significance.

To sum up, it appears that reporting behavior is very prone to changes induced by variations in wording and format. The results suggest that individual mean expectations and uncertainty vary strongly with survey framing. In some instances, changes in framing can lead to an increase in mean expectations of more than 100% compared to the initial level. One can confirm the initial assumption that simpler wording and less restrictive format would lead to higher expected inflation on average. Contrary to what was expected, respondents in the min-max treatment group appear much less uncertain on average based on the standard deviation of their histograms. This suggests that responses produced by the standard probabilistic question might be artificially more spread out, for instance, due to (i) assumptions on the support of the histogram, (ii) strong framing effects of the probabilistic format, or a combination of both.

4.3 Heterogeneity of Treatment Effects

Another important aspect is whether the estimated treatment effects are symmetric across socio-demographic groups. One can expect that this is not the case, since answering a probabilistic question about the inflation rate requires a certain degree of numeracy and financial literacy. Populations with lower financial literacy generally have more difficulty expressing their expectations in terms of probabilities (Fischhoff and Bruine de Bruin, 1999; Bruine de Bruin et al., 2000) and a less clear understanding of the economic term 'inflation' (Lusardi and Mitchell, 2014)[17]. Those less financially knowledgeable might be thus more susceptible to framing.

Unfortunately, while the core module of the survey collects information on multiple respondents' characteristics such as age, income, employment, profession, and others, there is no direct measure of one's financial literacy. Instead, I focus on the reported educational

[17]One of the questions used to measure financial literacy in a standardized framework worldwide, so-called 'Big 3', is a question asking for the respondent's understanding of inflation (Lusardi and Mitchell, 2014).

background and household income category as main proxies. Additionally, as many studies document differences in expectations by gender, e.g. originating from different shopping habits (D'Acunto et al., 2021b; Bryan and Venkatu, 2001), I also assess the treatment effects for men and women separately. Finally, literature reports differences in expectations based on experiences that have a long-lasting or 'imprinting' effect on some populations such as the German reunification (Goldfayn-Frank and Wohlfart, 2020), I therefore compare the effects for respondents who lived in East Germany before 1990 to those who did not.

I evaluate the heterogeneity in framing effects by repeatedly adding interaction terms of the treatment variables and the socio-demographic characteristics to Equation 1.[18] The corresponding estimates for mean expected inflation and uncertainty are reported in Tables B.3 and B.4 in Appendix B. Interestingly, asymmetry in the effect of the alternative question format is observed for mean expectations by gender and education, whereby the increasing effect is stronger for females and those with no college degree by roughly 1.5 to 2 pp, relative to their counterparts. The joint effect of wording and format on mean expected inflation is larger for below-median income households than for the top-50% in the income distribution - there is a 2.5 pp difference in expectations on average between these two groups. There are no asymmetries observed for wording, except for those living in East Germany prior to 1989, for which the *prices* option increases average expectations by roughly 2.5 pp. The latter also exhibit larger uncertainty when asked for their expectations for changes in prices in general. Overall, the estimates provide substantial evidence that different question formats produce very different mean expectations depending on the underlying population. Females and those with no college degree appear to adapt their forecasts for inflation more closely to the underlying response scale in the probabilistic question.

Finally, another important aspect to consider is whether respondents are part of the panel component of the sample. Kim and Binder (2023) document that repeated survey participation can impact the level of expectations and uncertainty of the respondents. Therefore I run the model of Equation 1 based only on the sub-sample of first-time respondents, which results in the exclusion of roughly 700 observations. The results reported in Table B.5 remain practically unchanged.

4.4 Effect of Wording on Forming Expectations

Next, I move on to Question 2 and analyze whether differences observed between the two wording treatments can be accounted for by differences in interpretation. 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. The distributions of the topics respondents self-reportedly thought about when answering inflation expectation questions are depicted in Figure 1. 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 data seems slightly more polarized with the majority of respondents thinking about their personal shopping experience, compared to 'Prices

[18] For ease of exposition, some variables such as household income and education are redefined to be binary. That is, I compare those with college degrees to those without, the top 50% of income distribution in the sample to the bottom 50%.

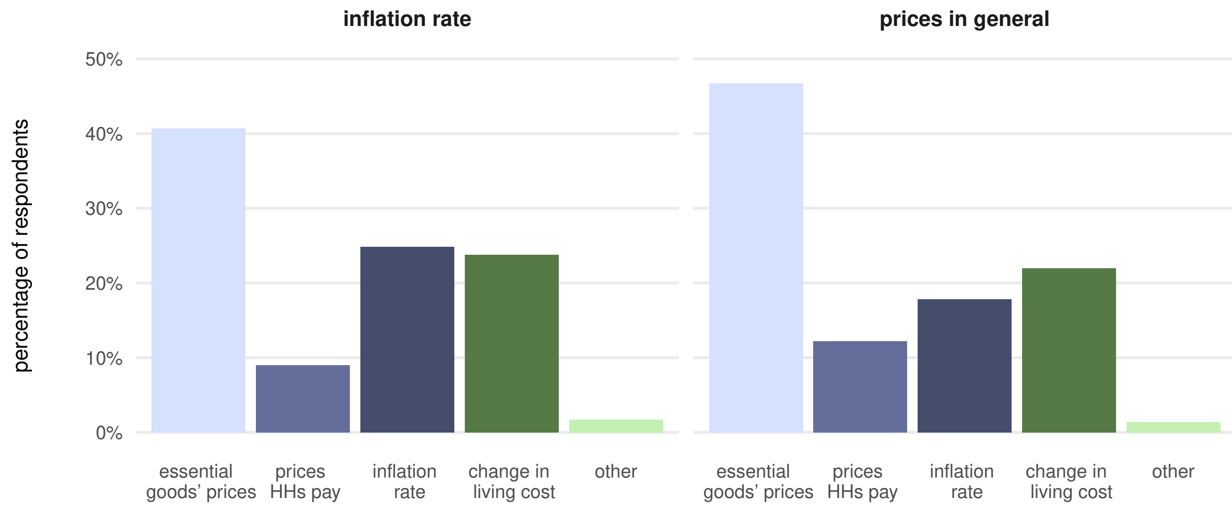


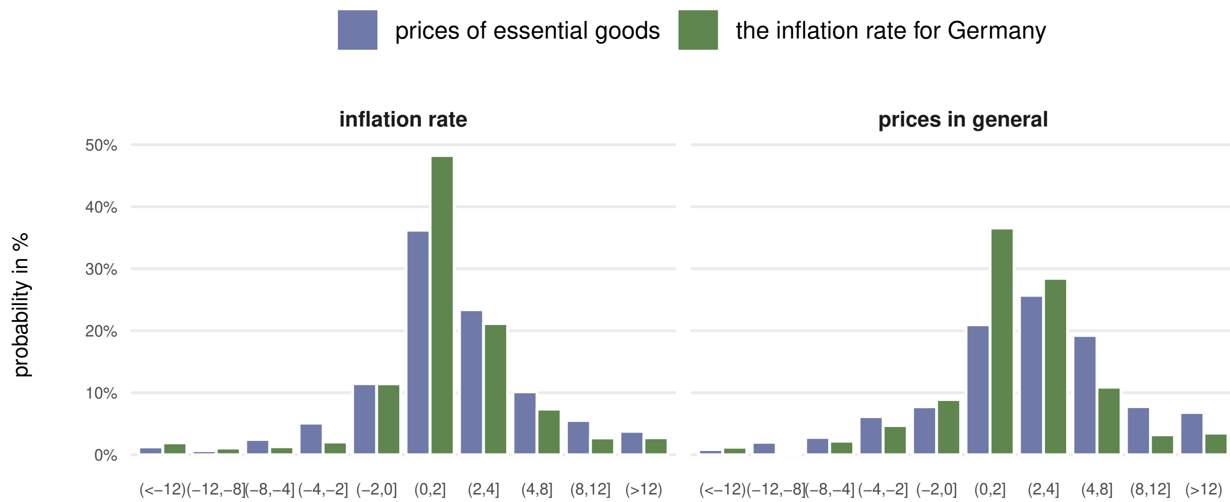
Figure 1: Effect of wording on forming expectations. Each respondent can select one of five options or provide an individual answer.

Americans pay', which was selected by roughly 40% of US survey participants as the main topic[19].

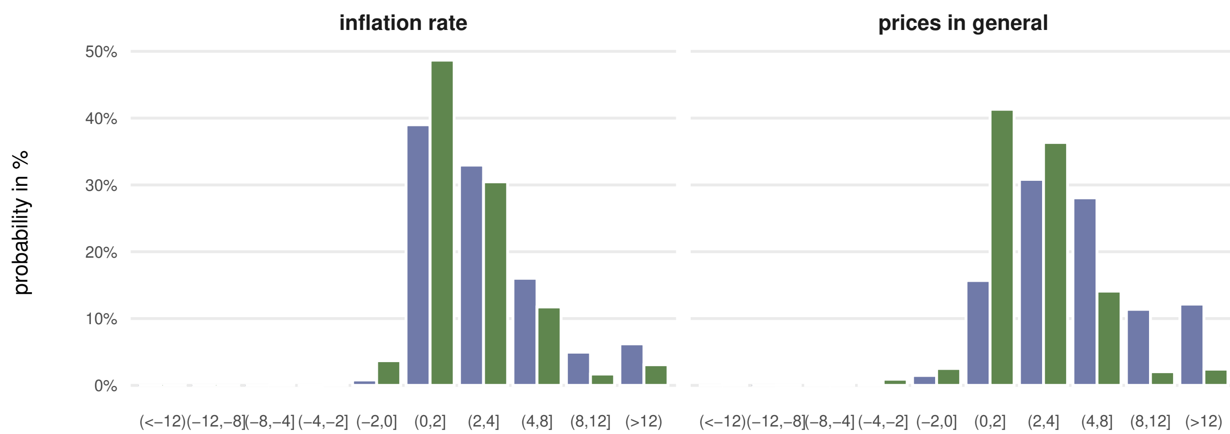
Although for the *inflation rate* wording, the corresponding share is somewhat smaller (41% vs. 47%), overall a substantial percentage of participants think about supermarket or gas prices when forming their expectations. Even when asked directly for their inflation rate prediction, only about one-quarter 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 a forecast based solely on their personal shopping experience, broader concepts of price changes remain less thought of. Hence: (i) diffuse question interpretation continues to be a major source of variation, beyond true differences in beliefs, and (ii) people base their inflation forecast on specific prices they observe in their day-to-day shopping instead of broader concepts about price changes.

To better illustrate the effect mentioned in (i) above and highlight the importance of clear and precise wording, in Figure 2 I plot the average subjective probabilities for people who reported thinking of 'prices of essential goods' vs. 'the inflation rate for Germany' for the four treatment arms. Even though the respondents within each treatment group received the exact same questions, the resulting pairs of distributions are very distinct. More formally, using a HotellingsT2 test one can reject the null that the pairs of subjective probabilities are the same within each wording and format combination. Expectations of people thinking about the inflation rate are clearly more anchored and contain less probability mass in the

[19]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.



(a) probabilistic format (default)



(b) min-max format

Figure 2: Average subjective probabilities by wording and association. Left panels represent groups who received the *inflation rate* wording, right panels - *prices in general* wording. Blue bars represent the subjective probabilities reported by participants who thought most about the prices of essential goods such as food and gasoline. Green bars depict the probabilities for those who report thinking about the inflation rate in Germany.

right tail than those reported thinking about essential goods' prices across all treatments.

Table B.8 in the Appendix further documents summary statistics for mean expected inflation and uncertainty for different groups of respondents, based on which concepts they reported thinking about. Interestingly, in most cases, there appear to be differences by wording even if participants report thinking about the same concept. For example, in the probabilistic question (resp. min-max question), those who base their forecast primarily on observed essential goods' prices report higher expected inflation in the *prices in general* wording 3.4% (6.5%) than in the *inflation rate* wording 2.2% (4.6%) on average. The same is

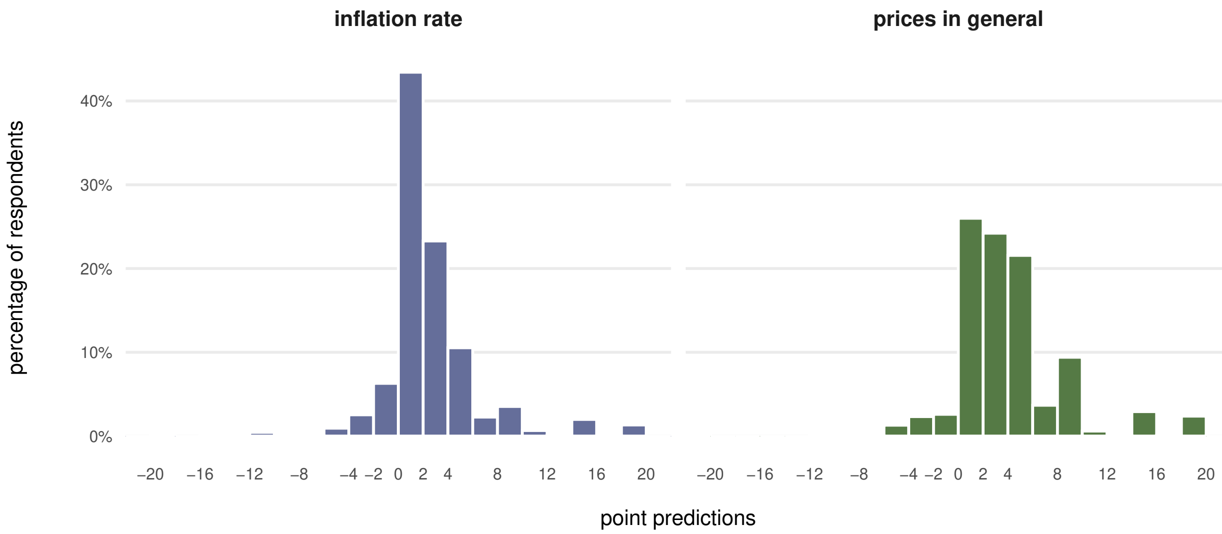


Figure 3: Distribution of point forecasts across wordings. Point predictions exceeding 100% in absolute value are removed. Values above 20% and below −20% are not shown but accounted for in the graph (roughly 2% of all observations). Observations are not weighted.

also true for individual uncertainty. However, the influence of wording choice is not present for the group reporting thinking about the inflation rate. This could be interpreted in a sense that the broader concept of 'inflation rate' not only acts as a numerical anchor but can potentially reduce more general framing effects.

Finally, regarding differences across demographic groups, e.g. by education, we observe the following: for each wording and each education group, the leading choice is 'prices you pay' as shown in Figure A.2 in the Appendix. However, the proportion of respondents selecting this option declines as education increases. Also, people with lower education tend to think less about the actual inflation rate, especially in the *prices in general* wording. Again, this suggests substantial asymmetry in the effects of wording for different demographic groups.

4.5 Point Forecasts and Perceptions

One can further compare the distributions of point forecasts which result from providing respondents with different wordings. These are depicted in Figure 3. Summary statistics can additionally be found in Table B.9 in the Appendix. The findings are in line with those from existing literature. The distribution of point forecasts from the *inflation rate* wording in the left panel of Figure 3 is much more concentrated on values between 0 and 4%, with more than 40% of respondents reporting values in the interval containing the inflation target. For *prices in general*, the majority of reported point forecasts lie in the bins covering values between 0 and 6%, and they appear roughly equally distributed across the three intervals in the range. While higher inflation values seem more likely for the *prices in general* treatment, there are more participants expecting deflation in the counterpart treatment. A spike in reported point predictions occurs for bins containing values that are multiples of five in both distributions, phenomena indicating uncertainty as documented by Binder (2017), is present in both distributions. Overall, one can reject the null hypothesis that the distributions are the same at the 1%-level using a Mann-Whitney test. Notably, the difference of more than 2 pp in mean point forecasts across wordings appears to be much more pronounced than those documented in previous literature, e.g. Bruine de Bruin et al. (2017).

One important policy indicator when it comes to aggregate distributions is the disagreement of point forecasts, that is the standard deviation across the latter (Boero et al., 2011). In each scenario, the *prices in general* wording yields point forecasts that are significantly more dispersed - 7.08 vs. 6.66 pp, confirmed by a Fligner-Killeen Test. In Figure A.3 in the Appendix I provide a more detailed overview of the point forecasts' distributions by both wording and interpretation, the corresponding measures are reported in Table B.9. To illustrate the importance of question interpretation as a source of heterogeneity, again I compare the distributions of individuals who report thinking about the same topic across wordings. It seems that the *prices* wording produces forecasts that (i) exhibit a higher upward bias, consistent with previous literature (Dräger and Fritsche, 2013), and (ii) higher disagreement, however only when people focus on personal shopping experiences. Another takeaway from Table B.9, is that similar to derived means, the effect of wording seems to disappear when survey respondents actually focus on the inflation rate when forming their point forecasts. This reiterates the importance of aiming for wording that increases the share of respondents thinking about the inflation rate. The bottom panel of Table B.9 documents the same measures for the inflation perceptions of respondents, for which, to a large extent, the analogous patterns apply.

5 Additional Aspects

This section discusses several aspects of response behavior starting by comparing the raw probabilities from the probabilistic question to those implied from the min-max question in Section 5.1. In the remainder of the Section, I look at further relevant dimensions such as item non-response (Section 5.2), rounding of reported values (Section 5.3), bin usage (Section 5.4), and consistency across quantitative measures (Section 5.5).

5.1 Comparison of Subjective Probabilities

The upper panel of Figure 4 depicts the distributions of the respondents' probabilistic expectations. The left panel shows the distributions for participants sub-sampled into the *inflation rate* treatment group, the right one for the *prices in general*. These are somewhat similar in shape, with probability assigned predominantly to positive outcomes. The interval $(0, 4]$, which contains the official inflation target as the midpoint, attracts more probability in the *inflation rate* treatment than in the *prices in general* treatment - 62% vs. 51%. With about one-third probability mass, respondents assigned to the latter group deem higher inflation outcomes ($> 4\%$) a lot more probable. In both cases, respondents consider deflationary trends quite likely - in the *inflation rate* slightly more than in the *prices in general*, with the difference being statistically significant at the 1%-level.

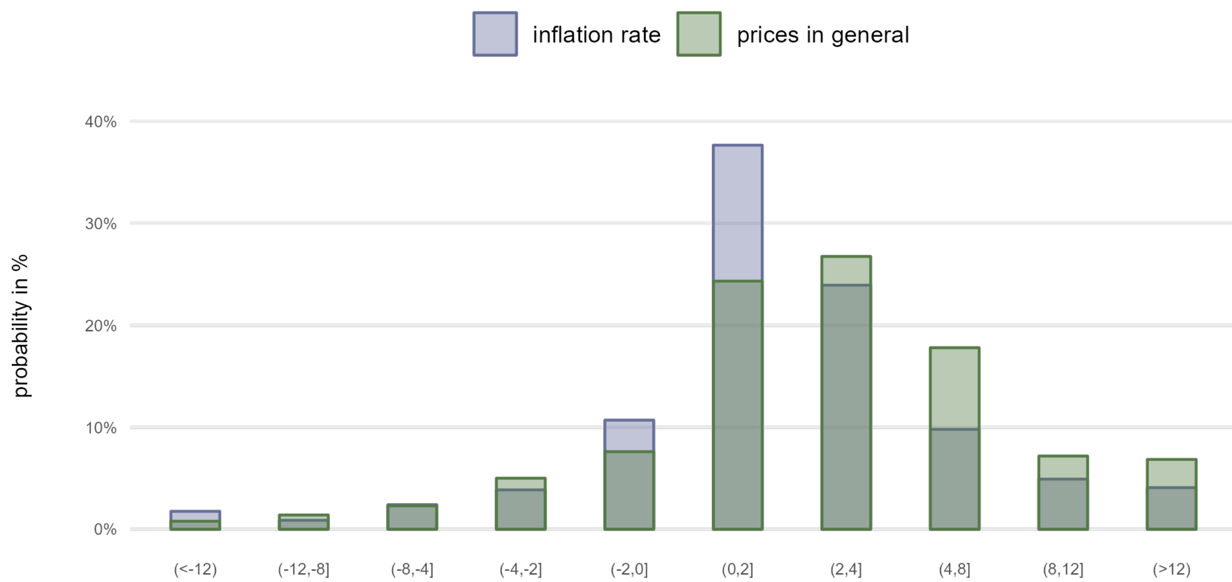
For comparability, the lower panel of Figure 4 reports the difference $F(u_k)_i - F(u_{k-1})_i$, where u is the upper bound of the interval k with $u \in \{-12, -8, -4, -2, 0, 2, 4, 8, 12, \infty\}$

Table 3: Response patterns across treatments

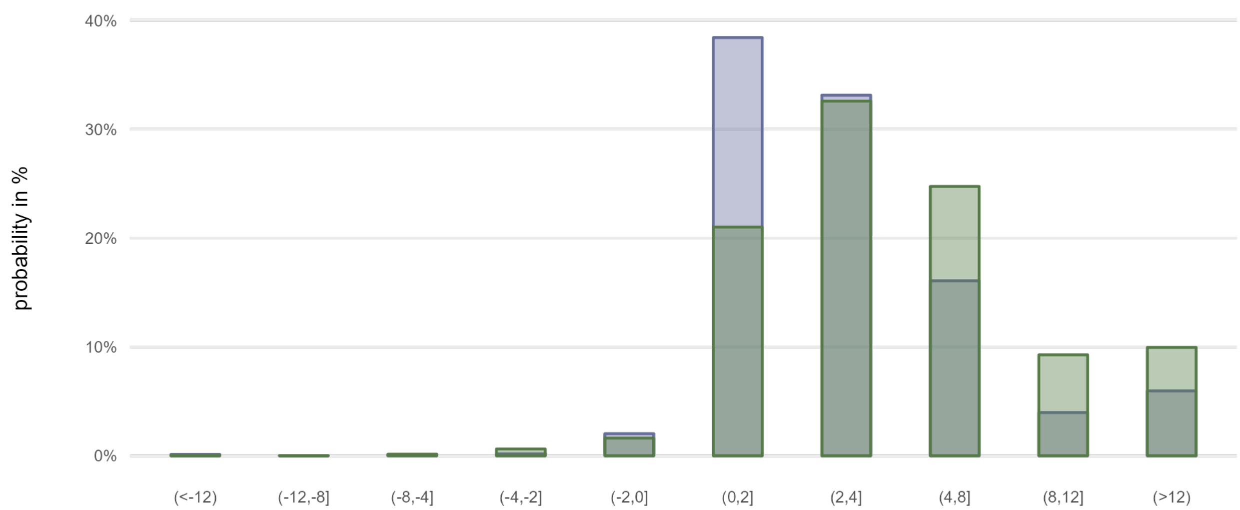
	probabilistic question		min-max question	
	inflation rate	*prices*	*inflation rate*	*prices*
Non-response				
item non-response	6.7	6.6	13.4^{m1}	$9.6^{p1,m5}$
Uncertainty and rouding				
sparse histogram	52.3	42.2^{p1}		
using 50-50% responses	7.7	7.1		
at least one outer bin	19.4	26.8^{p1}		
mean number of bins	3.2	3.5^{p1}		
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}
Consistency				
PP not in support	11.0	15.1^{p5}	14.2^{m1}	$8.2^{p1,m1}$
$q_5 <$ PP $> q_{95}$	24.8	27.0	37.3^{m1}	$31.5^{p5,m5}$
$P((X_{pp}, Y_{pp}]) = 0$	16.3	15.6		
contain disjoint regions	3.3	2.4		
Observations	944	948	879	915

Note:^{p1}, ^{p5}, ^{p10} indicate that the corresponding measure is significantly different in the *prices in general* from 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* from 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).

and $F(x)_i$ is the CDF of the triangular distribution of respondent i, based on the reported parameters, a, b, and c from the min-max question. In a sense, the reported distributions are 'discretized' to match the intervals of the probabilistic format and then averaged over all respondents. The most striking feature of the data is the substantially lower probability mass assigned to inflation outcomes below zero. Expectations that are predominantly positive are consistent with previous findings that positive price changes attract more of the consumers' attention and influence expectations more than negative ones (D'Acunto et al. (2021a), Cavallo et al. (2017)). The average probability assigned to negative inflation rates plummets from 19.6% to 2.4%, and 17.1% to 2.4% in each wording variation. This is in line with the findings of Becker et al. (2023) that providing respondents with more intervals representing negative values of inflation increases the probability placed in those. In the more extreme case of the min-max question where respondents are not confronted with any intervals, the probability of deflation dramatically declines. Apart from that, the distributions exhibit similar features with more probability mass concentrated in the right tail, especially for the *prices in general* wording.



(a) probabilistic format (default)



(b) min-max format

Figure 4: Average subjective probabilities for different ranges for inflation from the probabilistic and min-max format. Left panels display probabilities assigned by those who received the *inflation rate* wording, right ones for those with the *prices in general* wording.

5.2 Item Non-response

Despite its greater complexity, the probabilistic question yields lower non-response rates as indicated in Table 3. With almost 7%, the item non-response is low and there are no differences across wordings. In the min-max setup, non-response rates are significantly higher for each wording. Even if we account for respondents who put 100% probability in the very first bin (0.6%), this would not make up for the difference between formats.

Similar to the findings of Hayo and Méon (2022) for point forecasts, a possible explanation for this result 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, 2022), e.g. due to social desirability bias (SDB). That is, respondents evade answers they consider undesirable to the interviewer, researcher, or survey designer (Stantcheva, 2023). The fact that significant differences between wordings do arise in the min-max setup, with *inflation rate* producing a significantly higher non-response share than the *prices in general* additionally supports this argument. While participants experience more difficulty answering the *inflation rate* question, providing them with a guided question seems to offset the difference between wordings. Generally higher response rates are desirable in surveys, but as Hayo and Méon (2022) point out, such answers might simply increase the noise in the data and are less or not at all informative of the true beliefs of the respondent.

In the current setting, it is not feasible to precisely determine whether the probabilistic set-up 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.6 do provide some evidence for this argument. More precisely, Table B.6 reports the average marginal effects of format, wording, and several demographic characteristics on non-response for the point forecast (Column (1)), probabilistic (Column (2)), and min-max (Column (3)) questions. Non-response seems to be less likely, among others, for those who hold a bachelor's degree or higher and there is a clear pattern for household income - higher-income respondents are also more likely to provide an answer - in the probabilistic but not in the min-max question. Furthermore, roughly 80 participants did not provide a point forecast. Despite the power limitations due to sample size, the estimation reported in Column (4) of Table B.6 suggests that these respondents are subsequently more likely to provide a forecast in the probabilistic set-up than in the min-max set-up, even though they did not provide a value for the short-term inflation prognosis in the previous question.

5.3 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 3, I report the share of minimum, maximum, and mode that are multiples of five in the min-max question. Overall, a considerable share of

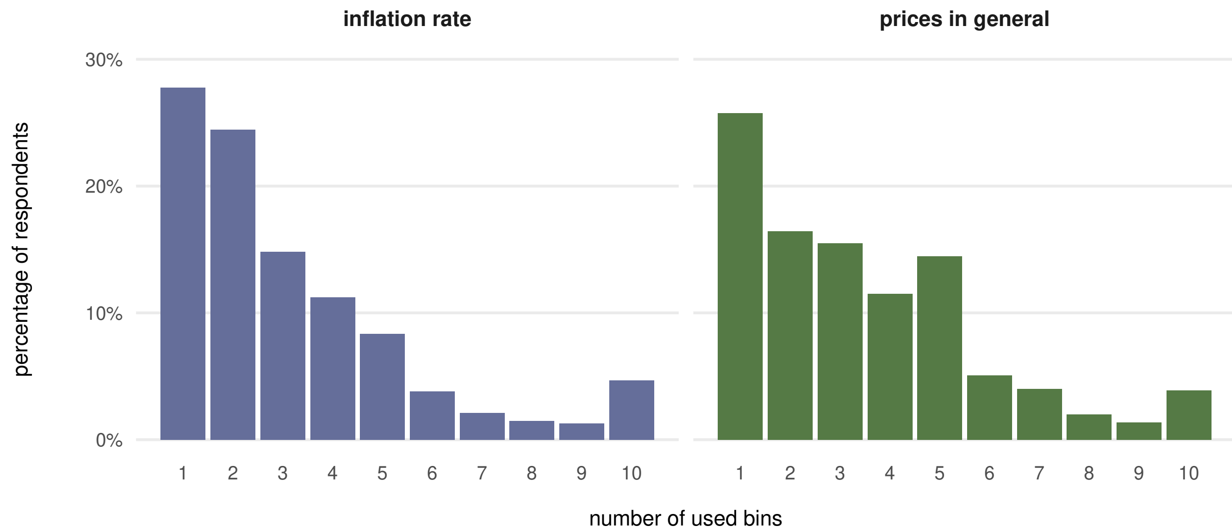


Figure 5: Share of respondents using number of bins $n \in \{1, ..., 10\}$ in the probabilistic question by wording

respondents provides rounded values with substantial differences between wording choices: the *prices in general* wording produces much more uncertain responses, which is intuitive given the higher volatility of food and energy prices people think of. Notably, it is the maximum value for expected inflation that is most likely a multiple of five, followed by the mode and minimum. 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 means that the usage of the min-max question does not lead to unusually high rounding in the data.

5.4 Bin Usage

More than half of the respondents 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 lower at 42%. This is comparable with other consumer surveys using a probabilistic format such as the SCE (Krüger and Pavlova, 2023). Generally, the *prices in general* wording seems to prompt respondents to use a higher number of bins as depicted in Figure 5. While the number of used bins gradually declines in the left panel of Figure 5, with a final peak at 10, for the *prices in general* wording, three peaks are visible: at 1, 5, and 10. To quantify better the effect of wording on the number of bins with positive probability in Table B.7 in the Appendix I report the estimates of a Poisson and quasi-Poisson regression of the number of used bins on the wording and a set of demographic controls. On average, receiving the *prices in general* formulation increases the number of bins with positive probability by 0.12 (0.13 respectively).

About one-fifth of the people who received the *inflation rate* treatment report positive probability in at least one outer bin, for the competing group this amounts to 26.8%. This result can be interpreted as being in line with previous literature which has found that when confronted with the *prices in general* wording, people tend to extrapolate based on their personal shopping experiences (Bruine de Bruin et al., 2010). As food and energy prices are more volatile than the overall inflation rate and often suffer from large hikes, respondents would more often use the outer bins in their prognosis. 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. 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.

5.5 Consistency across Quantitative Measures

Internal consistency across multiple inflation expectations measures is desirable for several reasons. From a survey designer's perspective, low internal consistency might indicate miscommunication. This could mean that the question does not accurately elicit the beliefs or expectations of interest, or in other words, has low *face validity*. Obviously, different interpretations on the side of the respondents, as well as their personal characteristics, could influence the degree of consistency found in the data.[20] Nonetheless, it is an important aspect to consider, especially when the data is used for policy evaluation and recommendations.

There are several ways to measure the internal consistency of reported answers. In the lower panel of Table 3 I report the proportion of occurrences when the point prediction (i) does not lie in the range of the support (fitted or self-reported), or (ii) lies below p_5 or above p_{95}, estimated based on the reported probabilities. For subjective histograms, in the *inflation rate* treatment, the first share is at 11%, whereas for *prices in general* at 15%. 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 subsequently more prone to revising their expectations downwards when confronted with the bin response scale. That is, priming effects are higher for the *prices* 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 and yields higher internal consistency of responses.

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) the number of cases of positive probability in disjoint regions. An example for (i) is a participant reporting an expected inflation rate of 5% as a point forecast, but then assigning 50% probability to the interval $[0\%, 2\%)$ and the remaining 50% to $[2\%, 4\%)$. Hence, the interval that contains the point forecast, $[4\%, 8\%)$ does not contain any probability. In turn, (ii) would mean that a respondent assigns for instance 50% probability to $[0\%, 2\%)$ and another 50% to $[4\%, 8\%)$ - intervals which are not adjacent. The latter case does occur in the data, but rather seldom in about 3% of all observations in either wording. Also, roughly 16% of respondents assign

[20] For a more detailed discussion on the internal consistency of household inflation expectations see e.g. Zhao (2023).

zero probability to the interval that contains their point prediction, but again, there are no significant differences across wordings.

What is striking is the fact that irrespective of treatment, roughly 30% to 40% of the participants in each subgroup report point forecasts that are either smaller than q_5 or exceed q_{95}. In contrast to professional forecasters whose point predictions are known to represent some central measure of tendency such as the mean, many consumers seem to provide a point prediction that reflects a tail outcome. This is an important implication as point forecasts and derived means from distributions are often used as substitutes in the empirical and experimental literature. To this end, in Figure A.4 in the Appendix, I plot the point predictions against the means derived from the probabilistic and min-max questions and the corresponding correlations. With 68%, the combination of the min-max format with the *prices in general* wording yields the highest correlation among the four options. This again re-iterates the advantage of the min-max over the probabilistic question that people directly state the range of the support, making answers across multiple questions more coherent. This is an essential point 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.

6 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 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 (2022), 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 increases. 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 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.

References

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

Armantier, O., W. Bruine de Bruin, S. Potter, G. Topa, W. van der Klaauw, and B. Zafar (2013): "Measuring Inflation Expectations," *Annual Review of Economics*, 5, 273–301.

Armantier, O., G. Koşar, R. Pomerantz, D. Skandalis, K. Smith, G. Topa, and W. van der Klaauw (2021): "How economic crises affect inflation beliefs: Evidence from the Covid-19 pandemic," *Journal of Economic Behavior & Organization*, 189, 443–469.

Armantier, O., G. Topa, W. van der Klaauw, and B. Zafar (2017): "An overview of the Survey of Consumer Expectations," *Federal Reserve Bank of New York Economic Policy Review*, 23, 51–72.

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

Becker, C., P. Duersch, and T. Eife (2023): "Measuring Inflation Expectations: How the Response Scale Shapes Density Forecasts," Working paper, University of Heidelberg, Department of Economics.

Beckmann, E. and T. Schmidt (2020): "Bundesbank online pilot survey on consumer expectations 01/2020," Tech. rep., Bundesbank Technical Paper.

Binder, C. C. (2017): "Measuring uncertainty based on rounding: New method and application to inflation expectations," *Journal of Monetary Economics*, 90, 1–12.

Boero, G., J. Smith, and K. F. Wallis (2011): "Scoring rules and survey density forecasts," *International Journal of Forecasting*, 27, 379–393.

Brandão-Marques, L., G. Gelos, D. Hofman, J. Otten, G. K. Pasricha, and Z. Strauss (2023): "Do Household Expectations Help Predict Inflation?" .

Bruine de Bruin, W., B. Fischhoff, S. G. Millstein, and B. L. Halpern-Felsher (2000): "Verbal and numerical expressions of probability:"It's a fifty–fifty chance"," *Organizational behavior and human decision processes*, 81, 115–131.

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

Bruine de Bruin, W., W. van der Klaauw, G. Topa, J. S. Downs, B. Fischhoff, and O. Armantier (2012): "The effect of question wording on consumers' reported inflation expectations," *Journal of Economic Psychology*, 33, 749–757.

Bruine de Bruin, W., W. van der Klaauw, M. van Rooij, F. Teppa, and K. de Vos (2017): "Measuring expectations of inflation: Effects of survey mode, wording, and opportunities to revise," *Journal of Economic Psychology*, 59, 45–58.

Bryan, M. F. and G. Venkatu (2001): "The curiously different inflation perspectives of men and women," *Economic Commentary*.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017): "Inflation expectations, learning, and supermarket prices: Evidence from survey experiments," *American Economic Journal: Macroeconomics*, 9, 1–35.

Christelis, D., D. Georgarakos, T. Jappelli, and M. van Rooij (2020): "Trust in the Central Bank and Inflation Expectations," *International Journal of Central Banking*, 16, 1–37.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, G. Kenny, and M. Weber (2022): "The effect of macroeconomic uncertainty on household spending," Chicago Booth Research Paper No. 22-17, University of Chicago, Becker Friedman Institute for Economics.

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

Curtin, R. (1996): "Procedure to estimate price expectations," Tech. rep., Ann Arbour, MI: University of Michigan Survey Research Center.

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2020): "Effective policy communication: Targets versus instruments," KIT Working Paper Series in Economics No. 147.

D'Acunto, F., U. Malmendier, and M. Weber (2023): "What do the data tell us about inflation expectations?" in *Handbook of economic expectations*, Elsevier, 133–161.

Das, S., C. M. Kuhnen, and S. Nagel (2020): "Socioeconomic status and macroeconomic expectations," *The Review of Financial Studies*, 33, 395–432.

Dominitz, J. and C. F. Manski (1997): "Using expectations data to study subjective income expectations," *Journal of the American Statistical Association*, 92, 855–867.

Dräger, L. and U. Fritsche (2013): "Don't Worry, Be Right! Survey Wording Effects on In flation Perceptions and Expectations," DEP (Socioeconomics) Discussion Papers-Macroeconomics and Finance Series No.8/ 2013.

Dräger, L. and M. J. Lamla (2023): "Consumers' macroeconomic expectations," *Journal of Economic Surveys*.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021a): "Exposure to grocery prices and inflation expectations," *Journal of Political Economy*, 129, 1615–1639.

D'Acunto, F., U. Malmendier, and M. Weber (2021b): "Gender roles produce divergent economic expectations," Proceedings of the National Academy of Sciences 21.

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

Fischhoff, B. and W. Bruine de Bruin (1999): "Fifty–fifty= 50%?" *Journal of Behavioral Decision Making*, 12, 149–163.

Goldfayn-Frank, O., P. Kieren, and S. Trautmann (2024): "A Choice-Based Approach to the Measurement of Inflation Expectations," AWI Discussion Paper Series No. 742, Heidelberg University.

Goldfayn-Frank, O. and J. Wohlfart (2020): "Expectation formation in a new environment: Evidence from the German reunification," *Journal of Monetary Economics*, 115, 301–320.

Hayo, B. and P.-G. Méon (2022): "Measuring household inflation perceptions and expectations: The effect of guided vs non-guided inflation questions," Working Paper No. 11-2022, MAGKS Joint Discussion Paper Series in Economics.

Huisman, R., N. L. Van der Sar, and R. C. Zwinkels (2021): "Volatility expectations and disagreement," *Journal of Economic Behavior & Organization*, 188, 379–393.

Hurd, M. D. (2009): "Subjective probabilities in household surveys," *Annual Review of Economics*, 1, 543–562.

Kim, G. and C. Binder (2023): "Learning-through-survey in inflation expectations," *American Economic Journal: Macroeconomics*, 15, 254–278.

Krüger, F. and L. Pavlova (2023): "Quantifying subjective uncertainty in survey expectations," *International Journal of Forecasting*.

Lusardi, A. and O. S. Mitchell (2014): "The economic importance of financial literacy: Theory and evidence," *American Economic Journal: Journal of Economic Literature*, 52, 5–44.

Malmendier, U. and S. Nagel (2016): "Learning from inflation experiences," *The Quarterly Journal of Economics*, 131, 53–87.

Manski, C. F. (2018): "Survey measurement of probabilistic macroeconomic expectations: progress and promise," *National Bureau of Economic Research Macroeconomics Annual*, 32, 411–471.

Manski, C. F. and F. Molinari (2010): "Rounding probabilistic expectations in surveys," *Journal of Business & Economic Statistics*, 28, 219–231.

Nagel, J. (2022): "Return of inflation: What do we learn from survey data?" Speech at the Joint Spring Conference on Monetary Policy and Expectations of Households and Firms (last accessed: October 14, 2022).

Phillot, M. and R. Rosenblatt-Wisch (2018): "Inflation expectations: The effect of question ordering on forecast inconsistencies," SNB Working Paper Series 11/2018, Swiss National Bank.

Stantcheva, S. (2023): "How to run surveys: A guide to creating your own identifying variation and revealing the invisible," *Annual Review of Economics*, 15, 205–234.

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

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

Appendix

A Additional Graphs

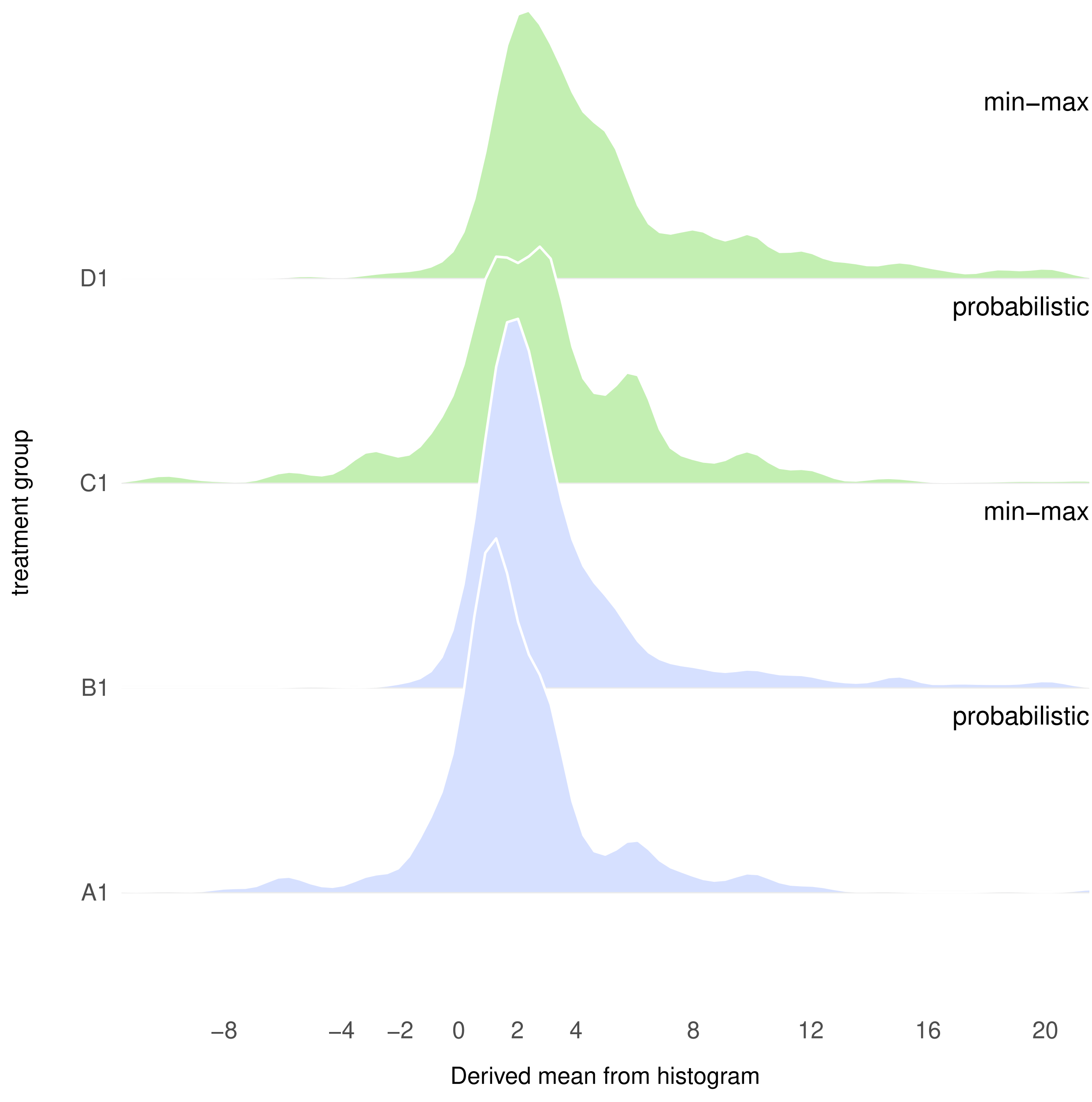
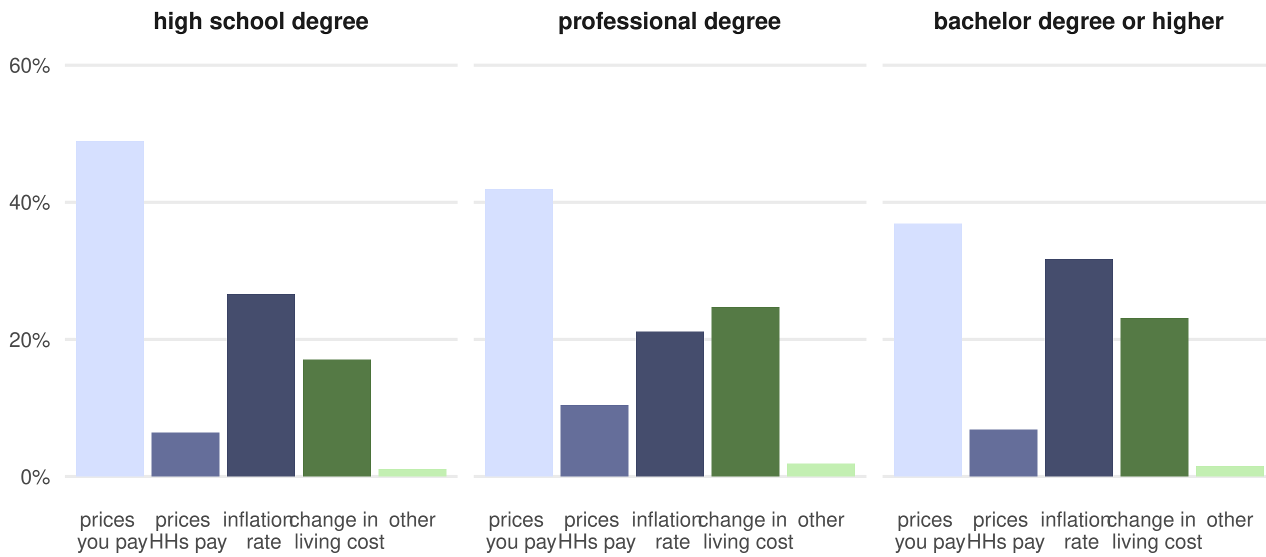
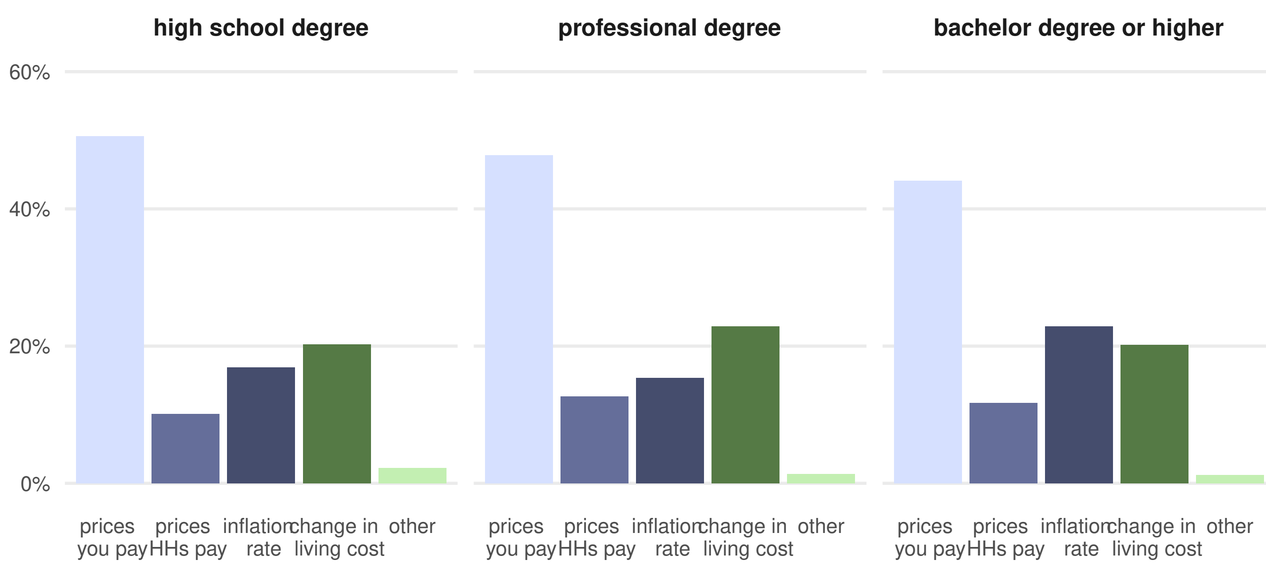


Figure A.1: Distribution of derived means from histograms. Treatment groups A1 and B1 receive the *inflation rate* wording (depicted in blue), and the remaining groups, C1 and D1, receive the *prices in general* (depicted in green).



(a) inflation rate



(b) prices in general

Figure A.2: Effect of wording on forming expectations for different demographic groups. Each respondent can select one of four pre-defined options or provide an individual answer. Upper panel: treatment groups A1 and B1, which received the *inflation rate* wording. Lower panel: groups C1 and D1 who answered a *prices-in-general* question.

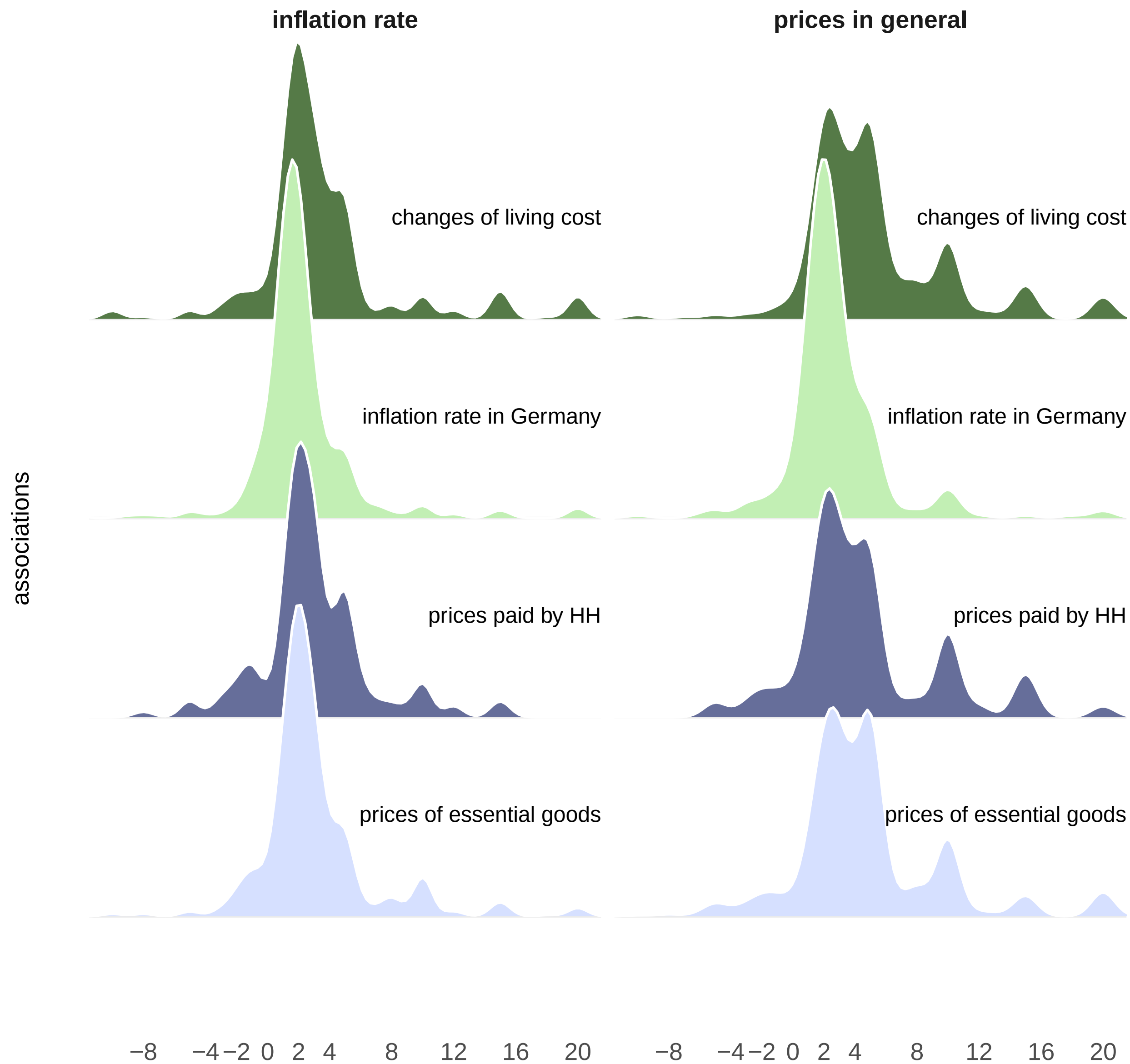


Figure A.3: Distributions of point forecasts by wording and association. The left panels depict distributions from groups that receive the *inflation rate* wording, right panels - *prices in general*.

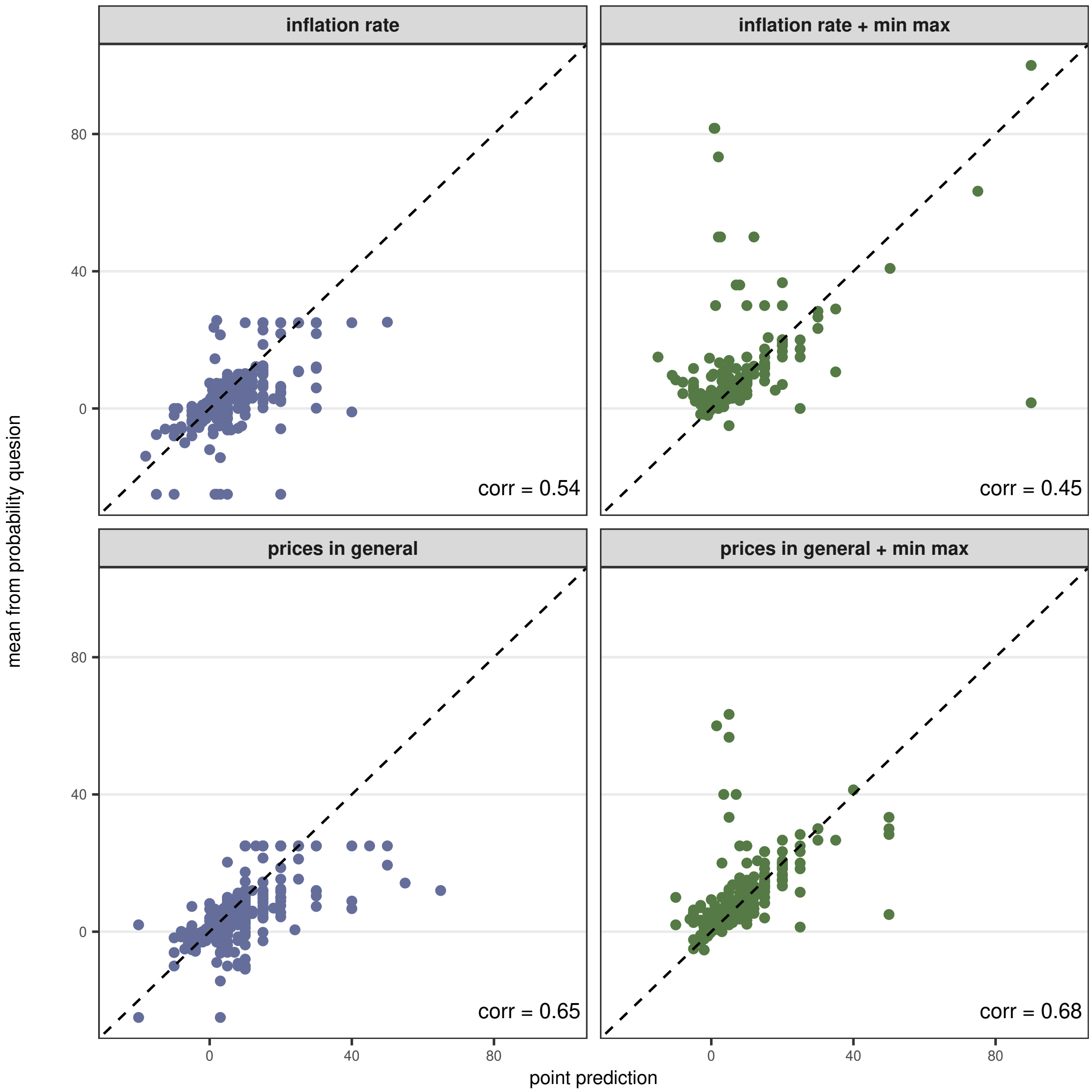


Figure A.4: Point predictions and means from derived or self-reported distributions together with correlation coefficients.

B Additional Tables

Table B.1: Summary statistics by treatment group

group	A1			B1			C1			D1		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
age	913	54	16	841	54	16	918	55	16	877	54	16
male	913	0.61	0.49	841	0.59	0.49	918	0.56	0.5	877	0.59	0.49
education	913			841			918			877		
high school	46	5%		38	5%		41	4%		42	5%	
professional degree	571	63%		519	62%		573	62%		563	64%	
bachelor or higher	296	32%		284	34%		304	33%		272	31%	
employed	913	0.60	0.49	841	0.59	0.49	918	0.57	0.5	877	0.59	0.49
HH income category	913			841			918			877		
HH income below €2,500	262	29%		244	29%		250	27%		273	31%	
HH income €2,500 to €3,500	219	24%		191	23%		240	26%		209	24%	
HH income €3,500 to €5,000	247	27%		227	27%		243	26%		231	26%	
HH income above €5,000	185	20%		179	21%		185	20%		164	19%	
HH size	913	2.3	1.1	841	2.3	1.1	918	2.3	1.1	877	2.3	1.1
region	913			841			918			877		
North	155	17%		154	18%		158	17%		170	19%	
West	247	27%		232	28%		224	24%		232	26%	
South	346	38%		298	35%		374	41%		321	37%	
East	165	18%		157	19%		162	18%		154	18%	
born in East Germany pre-1989	913	0.17	0.38	841	0.17	0.38	918	0.18	0.39	877	0.16	0.37
homeowner	913	0.68	0.47	841	0.66	0.47	918	0.67	0.47	877	0.69	0.46

Table B.2: Probability of selection into treatment

	group A1	group B1	group C1	group D1
	(1)	(2)	(3)	(4)
age (in years)	0.001	−0.0005	0.001	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
professional degree	−0.035	0.019	−0.006	0.021
	(0.038)	(0.037)	(0.038)	(0.037)
bachelor or higher	−0.033	0.026	−0.001	0.009
	(0.038)	(0.039)	(0.039)	(0.039)
employed	0.024	−0.008	−0.013	−0.003
	(0.019)	(0.019)	(0.019)	(0.019)
female	−0.003	−0.008	0.038*	−0.025
	(0.021)	(0.021)	(0.022)	(0.020)
HH income €2,500 to €3,500	−0.001	0.011	0.018	−0.028
	(0.022)	(0.022)	(0.023)	(0.021)
HH income €3,500 to €5,000	−0.004	0.025	0.026	−0.044*
	(0.026)	(0.026)	(0.026)	(0.024)
HH income €5,000 or more	0.003	−0.004	0.003	−0.002
	(0.008)	(0.008)	(0.008)	(0.008)
HH size	0.020	0.006	−0.008	−0.018
	(0.023)	(0.022)	(0.023)	(0.022)
West	0.014	−0.019	0.033	−0.027
	(0.021)	(0.020)	(0.021)	(0.020)
South	0.023	0.013	−0.023	−0.011
	(0.032)	(0.030)	(0.030)	(0.030)
East	−0.009	−0.015	0.044	−0.019
	(0.028)	(0.027)	(0.029)	(0.027)
born in East Germany pre-1989	0.006	−0.013	−0.021	0.029*
	(0.017)	(0.017)	(0.018)	(0.017)
Observations	3,549	3,549	3,549	3,549
Log Likelihood	−2,073.992	−1,989.057	−2,072.701	−2,056.774

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses. Additional control variables are omitted for brevity.

Table B.3: Heterogeneity in treatment effects

	Dependent variable:			
	$E[\pi_{t+12}]_i$			
	(1)	(2)	(3)	(4)
default	1.182	0.769	0.521	1.161
	(0.780)	(0.812)	(0.816)	(0.801)
min-max	1.944***	3.141***	3.147***	2.469***
	(0.406)	(0.473)	(0.562)	(0.372)
prices	1.310***	1.392***	1.476***	0.904***
	(0.347)	(0.400)	(0.425)	(0.332)
joint effect	4.233***	4.792***	5.350***	4.360***
	(0.520)	(0.534)	(0.600)	(0.487)
college	−1.790***	−0.701	−1.809***	−1.767***
	(0.276)	(0.455)	(0.276)	(0.274)
female	1.237***	1.686***	1.705***	1.624***
	(0.470)	(0.309)	(0.305)	(0.308)
$d_{highincome}$	−1.001***	−1.018***	0.024	−1.004***
	(0.318)	(0.319)	(0.497)	(0.318)
East pre 1989	1.961***	1.951***	1.892***	1.094*
	(0.504)	(0.509)	(0.516)	(0.615)
min-max ×female	1.529*			
	(0.807)			
prices ×female	0.067			
	(0.656)			
joint ×female	0.246			
	(0.887)			
min-max ×college		−2.112***		
		(0.764)		
prices ×college		−0.288		
		(0.602)		
joint ×college		−1.975***		
		(0.748)		
min-max ×$d_{highincome}$			−1.229	
			(0.751)	
prices ×$d_{highincome}$			−0.383	
			(0.653)	
joint ×$d_{highincome}$			−2.528***	
			(0.835)	
min-max ×East pre 1989				0.984
				(1.456)
prices ×East pre 1989				2.596**
				(1.035)
joint ×East pre 1989				−0.212
				(0.947)
Observations	3,549	3,528	3,549	3,549
R^2	0.111	0.110	0.108	0.113
Adjusted R^2	0.106	0.105	0.104	0.108

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses. Additional control variables are omitted for brevity.

Table B.4: Heterogeneity in treatment effects for uncertainty

	Dependent variable:			
	$\sigma_i^{\pi_{t+12}}$			
	(1)	(2)	(3)	(4)
default	2.530***	2.647***	2.573***	2.736***
	(0.310)	(0.317)	(0.323)	(0.317)
format	−0.947***	−1.032***	−1.012***	−1.107***
	(0.119)	(0.148)	(0.162)	(0.135)
wording	0.478***	0.357*	0.365*	0.216
	(0.174)	(0.187)	(0.201)	(0.167)
joint	−0.069	−0.323	−0.139	−0.423**
	(0.228)	(0.208)	(0.250)	(0.191)
college	−0.547***	−0.414**	−0.555***	−0.546***
	(0.092)	(0.186)	(0.092)	(0.092)
female	0.781***	0.453***	0.461***	0.443***
	(0.213)	(0.106)	(0.104)	(0.105)
$d_{highincome}$	−0.093	−0.094	0.116	−0.095
	(0.111)	(0.113)	(0.202)	(0.112)
East pre 1989	0.280**	0.267*	0.247*	−0.032
	(0.139)	(0.139)	(0.139)	(0.232)
format ×female	−0.232			
	(0.252)			
wording ×female	−0.353			
	(0.299)			
joint ×female	−0.713**			
	(0.327)			
format ×college		−0.071		
		(0.210)		
wording ×college		−0.149		
		(0.251)		
joint ×college		−0.321		
		(0.268)		
format ×$d_{highincome}$			−0.094	
			(0.233)	
wording ×$d_{highincome}$			−0.106	
			(0.288)	
joint ×$d_{highincome}$			−0.643**	
			(0.299)	
format × East pre 1989				0.370
				(0.246)
wording ×East pre 1989				0.660**
				(0.308)
joint ×East pre 1989				0.164
				(0.313)
Observations	3,528	3,528	3,528	3,528
R^2	0.081	0.079	0.081	0.080
Adjusted R^2	0.077	0.074	0.077	0.075

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses. Additional control variables are omitted for brevity.

Table B.5: Treatment effects for first-time respondents

	Dependent variable:					
	$E[\pi_{t+12}]_i$			$\sigma_i^{\pi+12}$		
	(1)	(2)	(3)	(4)	(5)	(6)
default	2.156***	1.817***	1.533*	2.074***	1.307***	2.700***
	(0.251)	(0.070)	(0.804)	(0.119)	(0.034)	(0.355)
min-max	2.558***	1.118***	2.719***	−1.090***	−0.624***	−1.091***
	(0.444)	(0.098)	(0.457)	(0.138)	(0.038)	(0.141)
prices	1.423***	1.135***	1.484***	0.324*	0.401***	0.274*
	(0.349)	(0.123)	(0.355)	(0.167)	(0.062)	(0.164)
joint effect	4.259***	2.531***	4.324***	−0.459***	−0.202***	−0.471***
	(0.453)	(0.120)	(0.469)	(0.169)	(0.042)	(0.180)
Observations	2,964	2,964	2,848	2,964	2,964	2,848
Robust linear		X			X	
Controls			X			X

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

Table B.6: Non-response behavior

	Dependent variable: d_nan			
	point forecasts	probabilistic	min-max	NA in PP
	(1)	(2)	(3)	(4)
age (in years)	0.0002	0.002***	0.002***	−0.004
	(0.0002)	(0.001)	(0.001)	(0.004)
professional degree	−0.010	−0.042	0.018	0.126
	(0.011)	(0.030)	(0.037)	(0.147)
bachelor or higher	−0.010	−0.067***	−0.013	0.238*
	(0.008)	(0.020)	(0.039)	(0.123)
employed	−0.005	−0.027*	−0.015	0.084
	(0.005)	(0.014)	(0.019)	(0.109)
male	−0.019***	−0.038***	−0.051***	0.203**
	(0.005)	(0.011)	(0.015)	(0.090)
HH income €2,500 to €3,500	−0.001	−0.049***	−0.031*	−0.178
	(0.005)	(0.011)	(0.017)	(0.113)
HH income €3,500 to €5,000	−0.009*	−0.059***	−0.028	−0.271*
	(0.005)	(0.011)	(0.018)	(0.145)
HH income €5,000 or more	−0.019***	−0.066***	−0.049**	0.240
	(0.004)	(0.010)	(0.019)	(11.106)
HH size	0.002	0.012*	0.020**	0.016
	(0.002)	(0.006)	(0.008)	(0.048)
West	0.006	−0.007	−0.026	−0.391***
	(0.008)	(0.016)	(0.019)	(0.151)
South	0.006	0.008	−0.010	−0.353***
	(0.007)	(0.016)	(0.019)	(0.109)
East	0.004	0.028	0.006	−0.256
	(0.008)	(0.021)	(0.023)	(0.193)
born in East Germany pre-1989	0.039	0.068	−0.039	−0.436**
	(0.033)	(0.064)	(0.065)	(0.204)
homeowner	−0.007	0.011	−0.023	−0.121
	(0.005)	(0.012)	(0.017)	(0.088)
min-max				0.224**
				(0.091)
'prices in general'	−0.008*	−0.004	−0.035**	−0.053
	(0.004)	(0.011)	(0.014)	(0.097)
Observations	4,000	2,004	1,996	74
Log Likelihood	−390.071	−454.912	−643.596	−28.562
Akaike Inf. Crit.	812.142	941.823	1,319.193	91.124

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

Table B.7: Effect of wording on the number of used bins

	Dependent variable: number of used bins	
	Poisson	*quasi-Poisson*
	(1)	(2)
intercept	1.876***	1.497***
	(0.067)	(0.075)
prices in general wording	0.122***	0.132***
	(0.025)	(0.028)
age (in years)	−0.015***	−0.012***
	(0.001)	(0.001)
professional degree	−0.056	0.048
	(0.043)	(0.052)
bachelor or higher	−0.069	0.122**
	(0.045)	(0.053)
employed	−0.016	−0.063*
	(0.029)	(0.032)
male	0.038	0.046
	(0.025)	(0.028)
HH income €2,500 to €3,500	0.053	0.061
	(0.036)	(0.039)
HH income €3,500 to €5,000	0.094**	0.027
	(0.038)	(0.042)
HH income €5,000 or more	0.093**	0.053
	(0.042)	(0.047)
HH size	−0.015	0.037***
	(0.013)	(0.014)
West	0.095**	0.033
	(0.041)	(0.045)
South	0.036	0.024
	(0.039)	(0.043)
East	−0.013	0.003
	(0.053)	(0.057)
born in East Germany pre-1989	−0.016	−0.015
	(0.048)	(0.051)
homeowner	0.039	0.003
	(0.028)	(0.031)
Observations	1.831	1.753
Log Likelihood	-3,892.672	-3,259.051
Akaike Inf. Crit.	7,817.345	6,550.103

Note:*p<0.1; **p<0.05; ***p<0.01. Robust standard errors in parentheses.

Table B.8: Summary statistics on derived means and std. deviation

	$E[\pi_{t+12}]$		$\sigma^{\pi_{t+12}}$	
	probabilistic format			
	inflation rate	prices in general	inflation rate	prices in general
prices of essential goods	2.20	3.41^{p1}	1.79	2.18^{p1}
prices paid by HHs in Germany	1.82	3.73^{p5}	1.89	2.01
inflation rate	1.38	2.15	1.88	1.60
changes in living cost	2.77	4.55^{p1}	1.90	2.20
	$E[\pi_{t+12}]$		$\sigma^{\pi_{t+12}}$	
	min-max format			
	inflation rate	prices in general	inflation rate	prices in general
prices of essential goods	4.60	6.49^{p1}	0.89	1.56^{p1}
prices paid by HHs in Germany	4.71	4.90	0.87	1.21^{p10}
inflation rate	3.55	2.92	0.76	0.99
changes in living cost	4.45	6.10^{p5}	0.96	1.30^{p5}

Note: ^{p1}, ^{p5}, ^{p10} indicate that the corresponding measure is significantly different in the *prices in general* from the *inflation rate* wording at the 1, 5, and 10%-level. To test for significant differences in means I use *t*-tests.

Table B.9: Summary statistics of point forecasts and perceptions

	π^{pp}_{t+12}		$\sigma^{\pi^{pp}_{t+12}}$	
	inflation rate	prices in general	inflation rate	prices in general
prices of essential goods	3.40	5.61^{p1}	5.69	7.57^{p1}
prices paid by HHs in Germany	3.23	4.76^{p5}	7.82	4.78^{p10}
inflation rate	2.56	3.07	6.57	6.22
changes in living cost	4.19	6.09^{p1}	7.78	7.40^{p1}
all	3.36	5.15^{p1}	6.66	7.08^{p1}
	π^{pp}_{t-12}		$\sigma^{\pi^{pp}_{t-12}}$	
	inflation rate	prices in general	inflation rate	prices in general
prices of essential goods	3.64	5.64^{p1}	6.86	7.05^{p1}
prices paid by HHs in Germany	4.37	4.87	8.82	5.58^{p1}
inflation rate	3.02	3.05	7.76	3.95^{p1}
changes in living cost	3.92	6.15^{p1}	5.28	7.20^{p1}
all	3.60	5.20^{p1}	6.93	6.58^{p1}

Note: ^{p1}, ^{p5}, ^{p10} indicate that the corresponding measure is significantly different in the *prices in general* from the *inflation rate* wording at the 1, 5, and 10%-level. To test for significant differences in means I use *t*-tests, to test for differences in disagreement - Fligner-Killeen Tests.

C Survey details

Table C.10: Selected survey questions from BOP-HH September 2020 wave

The inflation rate-Intro	
`Group filter: drandom1` $= 1 \| = 2$	
Now we would like you to think more carefully about the inflation rate.	
The inflation rate:	
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'.	
Inflation development	
`Group filter: drandom1` $= 1 \| = 2$	`Group filter: drandom1` $= 3 \| = 4$
903A: What do you think the rate of inflation or deflation in Germany was over the past twelve months?	**903B:** By what percent do you think prices in Germany, in general, have increased or decreased over the past twelve months?
Note: If it is assumed that there was deflation, please enter a negative value.	Note: If it is assumed that prices have fallen, please enter a negative value.
Values may have a maximum of one decimal separator.	Values may have a maximum of one decimal separator.
Please use a full stop rather than a comma as the decimal separator.	Please use a full stop rather than a comma as the decimal separator.
Please enter a value here:	Please enter a value here:
`[Input field]` percent	`[Input field]` percent
`Group filter: drandom1` $= 1 \| = 2$	`Group filter: drandom1` $= 3 \| = 4$
005A1: Do you think inflation or deflation is more likely over the next twelve months?	**005A2:** Do you think prices in general, are more likely to increase or decrease over the next twelve months?
Note: Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as 'deflation'.	
Please select one answer:	Please select one answer:
1 = Inflation more likely	1 = More likely to increase
2 = Deflation more likely	2 = More likely to decrease

Table C.10: Selected survey questions from BOP-HH September 2020 wave (Continued)

Inflation expectations quantitative	
`Group filter: drandom1` $= 1 \| = 2$	`Group filter: drandom1` $= 3 \| = 4$
`If 005A1` $= 1 \| -9997 \| -9998$	`If 005A2` $= 1 \| -9997 \| -9998$
005B1: What do you expect the rate of inflation in Germany to roughly be over the next twelve months?	**005B2:** By roughly what percentage do you expect prices in general, to increase over the next twelve months?
`If 005A1` $= 2$	`If 005A2` $= 2$
005B1: What do you expect the rate of deflation in Germany to roughly be over the next twelve months?	**005B2:** By roughly what percentage do you expect prices in general to decrease over the next twelve months?
Note: Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as 'deflation'.	
Please enter a value in the input field (values may have one decimal place).	Please enter a value in the input field (values may have one decimal place).
`[Input field]` percent	`[Input field]` percent
Inflation expectations qualitative	
`Group filter: drandom1` $= 1 \| = 2$	`Group filter: drandom1` $= 3 \| = 4$
005A1: Do you think inflation or deflation is more likely over the next twelve months?	**005A2:** Do you think prices in general, are more likely to increase or decrease over the next twelve months?
Note: Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as 'deflation'.	
Please select one answer:	Please select one answer:
1 = Inflation more likely	1 = More likely to increase
2 = Deflation more likely	2 = More likely to decrease
Inflation expectations quantitative	
`Group filter: drandom1` $= 1 \| = 2$	`Group filter: drandom1` $= 3 \| = 4$
`If 005A1` $= 1 \| -9997 \| -9998$	`If 005A2` $= 1 \| -9997 \| -9998$
005B1: What do you expect the rate of inflation in Germany to roughly be over the next twelve months?	**005B2:** By roughly what percentage do you expect prices in general, to increase over the next twelve months?

Table C.10: Selected survey questions from BOP-HH September 2020 wave (Continued)

`If 005A1` = 2	`If 005A2` = 2
005B1: What do you expect the rate of deflation in Germany to roughly be over the next twelve months?	**005B2:** By roughly what percentage do you expect prices in general to decrease over the next twelve months?
Note: Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as 'deflation'.	
Please enter a value in the input field (values may have one decimal place).	Please enter a value in the input field (values may have one decimal place).
`[Input field]` percent	`[Input field]` percent
Inflation expectations probabilistic	
`Group filter: drandom1` = 1	`Group filter: drandom1` = 3
904A1: In your opinion, how likely is it that the rate of inflation will change as follows over the next twelve months?	**904A2:** In your opinion, how likely is it that prices in general will change as follows over the next twelve months?
Note: The aim of this question is to determine how likely you think it is that something specific will happen in the future. You can rate the likelihood on a scale from 0 to 100, with 0 meaning that an event is completely unlikely and 100 meaning that you are absolutely certain it will happen. Use values between the two extremes to moderate the strength of your opinion. Please note that your answers to the categories must add up to 100.	Note: The aim of this question is to determine how likely you think it is that something specific will happen in the future. You can rate the likelihood on a scale from 0 to 100, with 0 meaning that an event is completely unlikely and 100 meaning that you are absolutely certain it will happen. Use values between the two extremes to moderate the strength of your opinion. Please note that your answers to the categories must add up to 100.
The rate of deflation (opposite of inflation) will be 12% or higher	Prices will decrease by 12% or more
The rate of deflation (opposite of inflation) will be between 8% and 12%	Prices will decrease between 8% and 12%
The rate of deflation (opposite of inflation) will be between 4% and 8%	Prices will decrease between 4% and 8%
The rate of deflation (opposite of inflation) will be between 2% and 4%	Prices will decrease between 2% and 4%
The rate of deflation (opposite of inflation) will be between 0% and 2%	Prices will decrease between 0% and 2%

Table C.10: Selected survey questions from BOP-HH September 2020 wave (Continued)

The rate of inflation will be between 0% and 2%	Prices will increase between 0% and 2%
The rate of inflation will be between 2% and 4%	Prices will increase between 2% and 4%
The rate of inflation will be between 4% and 8%	Prices will increase between 4% and 8%
The rate of inflation will be between 8% and 12%	Prices will increase between 8% and 12%
The rate of inflation will be 12% or higher	Prices will increase by 12% or more
`Group filter: drandom1` $= 2$	`Group filter: drandom1` $= 4$
904B1: 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 and minimum value?	**904B2** By what percentage do you think prices in general are most likely to increase or decrease over the next twelve months? What will the price change be as a maximum and minimum value?
Note: If it is assumed that there will be deflation, please enter a negative value.	Note: if it is assumed that prices will fall, please enter a negative value.
Values may have a maximum of one decimal place.	Values may have a maximum of one decimal place.
Please use a full stop rather than a comma as the decimal separator.	Please use a full stop rather than a comma as the decimal separator.
Most likely inflation or deflation rate `[Input field]` percent	Most likely change `[Input field]` percent
Minimum `[Input field]` percent	Minimum `[Input field]` percent
Maximum `[Input field]` percent	Maximum `[Input field]` percent

Download ZEW Discussion Papers:

https://www.zew.de/en/publications/zew-discussion-papers

or see:

https://www.ssrn.com/link/ZEW-Ctr-Euro-Econ-Research.html
https://ideas.repec.org/s/zbw/zewdip.html

// IMPRINT

ZEW – Leibniz-Zentrum für Europäische Wirtschaftsforschung GmbH Mannheim

ZEW – Leibniz Centre for European Economic Research

L 7,1 · 68161 Mannheim · Germany
Phone +49 621 1235-01
info@zew.de · zew.de

Discussion Papers are intended to make results of ZEW research promptly available to other economists in order to encourage discussion and suggestions for revisions. The authors are solely responsible for the contents which do not necessarily represent the opinion of the ZEW.