FAU Discussion Papers in Economics

No. 01/2023

Households' probabilistic inflation expectations in high-inflation regimes

Christoph Becker
Heidelberg University

Peter Dürsch
University of Mannheim

Thomas Eife
Heidelberg University

Alexander Glas
FAU Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/discussion-papers/

Households' probabilistic inflation expectations in high-inflation regimes*

Christoph Becker[†] Peter Duersch[‡] Thomas Eife[§] Alexander Glas[¶]

January 12, 2023

Abstract

Central bank surveys frequently elicit households' probabilistic beliefs about future inflation. The responses provide only a coarse picture of inflation beliefs further away from zero. Using data from the Bundesbank household panel, we show that the current high-inflation environment induces respondents to allocate considerable probability to the rightmost open interval. This pile-up of probabilities negatively affects estimates of histogram moments and leads to a divergence between average expected inflation measured by probabilistic and point forecasts. The consistency of predictions can be improved by using an alternative design of the response scale that allows respondents to state more detailed beliefs for higher inflation ranges.

Keywords: Probabilistic expectations, inflation, survey data

JEL classification: D84, E31, E58

*We are grateful to the Bundesbank for including our questions in the Bundesbank Online Panel Households. Our research was improved by very helpful comments and suggestions of Jonas Dovern, Johannes Frank, Norbert Schwarz and Michael Weber.

[†]Heidelberg University, christoph.becker@awi.uni-heidelberg.de.

[‡]University of Mannheim, duersch@xeeron.de.

[§]Heidelberg University, thomas.eife@awi.uni-heidelberg.de.

[¶]Corresponding author. Friedrich-Alexander-Universität Erlangen-Nürnberg, alexander.glas@fau.de.

1 Introduction

Survey data are a popular source of information about the macroeconomic expectations of experts, households and firms. In addition to point forecasts, many surveys provide probabilistic expectations which are typically elicited by asking respondents to assign probabilities to pre-defined outcome intervals ('bins'). These probability distributions offer important insights into how survey participants assess the uncertainty, skewness and tail risk associated with their predictions (Manski, 2004).

In this paper, we analyze the quality of the probabilistic inflation expectations measured in the Bundesbank Online Panel Households (BOP-HH) in light of the recent surge of inflation in Germany and the euro area as a whole. In particular, we assess whether adjusting the bin definitions improves the consistency between the point forecasts and the probabilistic expectations by conducting a randomized experiment where some of the participants in Wave 30 (June 2022) receive the original bin design, while others receive an alternative design where the center of the intervals is closer to—but still below—the actual German inflation rate.

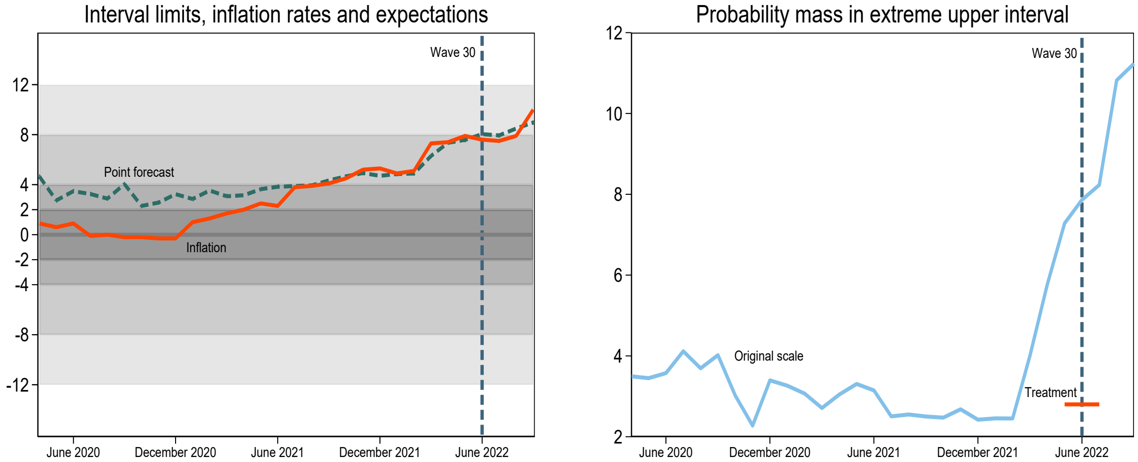
Our central finding is that the alternative design leads to considerably more consistent responses with the probabilistic expectations closely matching both actual inflation and point forecasts. This improved match between point forecasts and probabilistic expectations is driven by the fact that the original scale offers respondents a relatively small set of reasonable choices at times when inflation is very low or very high. For example, respondents who expect inflation rates of eight percent or higher only have two intervals at their disposal. This forces them either to provide inconsistent answers or to assign probabilities in extreme, marginal intervals, which is something that many respondents tend to avoid (Becker et al., 2023). Our finding is relevant for all surveys that employ scales similar to the one used in BOP-HH.[1]

As illustrated in the left plot in Figure 1, the question about probabilistic expectations consists of ten bins which are centered around an inflation rate of 0%. The interior bins cover the range from -12% to $+12\%$. The two exterior bins are half-open. A major advantage of using this responce scale is that it allows for a comparison of results both within surveys (across time) and between surveys (across different geographical locations). The red line shows the monthly German inflation rate based on the consumer price index. Before 2021, inflation rates were close to the center of the response scale. Inflation began to rise during the COVID-19 pandemic and further accelerated after the Russian invasion of Ukraine in February 2022 and the associated energy crisis. The inflation rate in June 2022, when our experiment was conducted, was 7.9%, which is just slightly below the lower bound of the rightmost interior bin. By September 2022, inflation further increased to 10%. The green line shows average inflation expectations in the BOP-HH. Clearly, households take notice of this development and adjust their point forecast accordingly.

The increase in households' point predictions is accompanied by an upward shift in their probabilistic inflation expectations. The blue line in the right plot shows the average probability mass assigned to the rightmost (half-open) bin. Before February 2022, the

[1] The baseline definition used in the BOP-HH was originally designed for the Federal Reserve Bank of New York's Survey of Consumer Expectations. See Armantier et al. (2017) for an overview. Other examples include the European Central Bank's Consumer Expectations Survey and similar surveys conducted by the central banks of Canada, France, the Netherlands, Ukraine, and the United Kingdom.

Figure 1: Probabilistic inflation expectations and interval definitions



Notes: The left plot shows monthly German consumer price inflation (red line). The dashed green line depicts the average inflation expectations of German households (trimmed by 1% from bottom and top in each month). The shaded gray areas correspond to the original bin definitions in the BOP-HH. The dashed blue line indicates the June 2022 wave of the BOP-HH to which we contributed an alternative bin design. The right plot shows the average probability mass in the highest bin based on participants presented with the original bin design (blue line). The red bar shows the corresponding average probability mass for the individuals presented our alternative bin design.

average probability fluctuated at relatively low levels between 2% and 4%. Consistent with the higher average point forecasts, we observe a steep increase in the average probability since the Russian invasion of Ukraine. The average probability in the rightmost bin was 7.9% in June 2022 and rose even further to more than 11% by September 2022. Since it is unknown what respondents consider a likely upper bound for inflation, the information provided by the open interval is limited. One has to make an assumption about the upper bound to derive a belief distribution from the answers. Thus, the evidence in Figure 1 puts into question the reliability of moments derived from the probabilistic expectations based on the original survey design.

We contributed an alternative bin design to Wave 30 of the BOP-HH where the center of the intervals is shifted from 0% to 4%, while keeping the relative bin width identical to the original design. As a result, the interior bins in the alternative treatment cover a range from −8% to +16%. The red bar in the right plot shows that for this treatment group, the average probability mass assigned to the rightmost bin is 2.8%, which is much more in line with the figures observed in earlier survey waves. These respondents use more bins, report higher histogram means that are more consistent with their point forecasts and report lower uncertainty than those in the baseline group. We conclude that the distortion of moments of the obtained belief distribution can be reduced by adjusting the bin definitions at times when inflation is unusually high.

Our research relates to the literature that explores how households form their macroeconomic expectations. Important covariates include households' socioeconomic characteristics such as gender, income and education (Bruine de Bruin et al., 2010; Das et al., 2020), their sources of information about monetary policy and the state of the economy (Coibion et al., 2022; Conrad et al., 2022) as well as individual and macroeconomic lifetime experiences (Malmendier and Nagel, 2011, 2016; D'Acunto et al., 2021). Using the BOP-HH data, Conrad et al. (2022) show that households' quantitative inflation expectations are related to the information channels that households use to inform themselves about monetary policy. In contrast, their qualitative expectations, i.e., the expected future direction of inflation, is more closely related to an individuals' lifetime inflation experiences. While these studies focus on households' point forecasts, we consider probabilistic expectations. Using the Michigan Survey of Consumers, Bruine de Bruin et al. (2011) show that consumers are generally willing and able to provide meaningful probability distributions that are consistent with the point predictions. Similarly, Zhao (2022) finds that the point forecasts of US households in the Federal Reserve Bank of New York's Survey of Consumer Expectations tend to be well-aligned with their probabilistic expectations. We contribute to the literature by analyzing whether the quality of the probabilistic expectations is related to the formulation of the corresponding question in the survey questionnaire in high-inflation regimes. As such, our analysis also relates to the literature that analyze how specifics of the survey design influence the responses. Here, Schwarz (2010) gives a good overview in general while Becker et al. (2023) and Weber et al. (2022) discuss this point in the context of inflation expectations.

The rest of this paper is organized as follows. Section 2 explains the data and discusses the competing designs of the question used for the probabilistic inflation expectations. Section 3 presents the results. We discuss our findings in Section 4. Section 5 concludes.

2 Bundesbank Online Panel Households

We use data from the BOP-HH, a representative online survey of German households operated by the Bundesbank. The survey targets individuals aged 16 years or older (see Beckmann and Schmidt, 2020, for details on the elicitation process). Among other questions, participants are asked to state their inflation expectations and socioeconomic characteristics. The survey started in 2019 with three pilot surveys. Starting with Wave 4 (April 2020), the BOP-HH is issued on a monthly basis. We focus on the responses from Wave 30 (June 2022) to which we contributed alternative formulations for the question on the probabilistic inflation expectations. In Section 3.4, we consider revisions of inflation expectations by comparing the responses from Wave 30 to those in Wave 29 (May 2022) and 31 (July 2022).

In total, 4,460 households participated in Wave 30. We remove observations from the sample whenever the household did not report probabilistic inflation expectations or if information for any of the socioeconomic characteristics is missing. We also exclude one respondent who did not state whether her point forecast represents a deflation rate or an inflation rate. This leaves 4,094 observations in our sample for Wave 30.

2.1 Probabilistic inflation expectations

BOP-HH participants receive the following question on their probabilistic expectations:[2]

[2]All questions related to inflation include an info box that informs respondents that inflation is defined as the percentage change in the general price level as measured by the consumer prices index. They also receive the information that deflation is the opposite of inflation.

```
CM004:  In your opinion, how likely is it that the rate of inflation will
change as follows over the next twelve months?
```

- `The rate of deflation (opposite of inflation) will be 12% or higher.`
- `The rate of deflation ([...])  will be between 8% and less than 12%.`
- `The rate of deflation ([...])  will be between 4% and less than 8%.`
- `The rate of deflation ([...])  will be between 2% and less than 4%.`
- `The rate of deflation ([...])  will be between 0% and less than 2%.`
- `The rate of inflation will be between 0% and less than 2%.`
- `The rate of inflation will be between 2% and less than 4%.`
- `The rate of inflation will be between 4% and less than 8%.`
- `The rate of inflation will be between 8% and less than 12%.`
- `The rate of inflation will be 12% or higher.`

Respondents are asked to rate the probability of inflation falling into each bin on a scale from 0 to 100, with 0 meaning that this outcome is completely unlikely and 100 meaning that they are absolutely certain it will happen. They also receive a notification that probabilities should add up to 100%. As mentioned above, the ten bins are centered around an inflation rate of 0%. Motivated by the recent surge in inflation rates, we contributed the following alternative bin design to the questionnaire of Wave 30:

```
P3001A: In your opinion, how likely is it that the rate of inflation will
change as follows over the next twelve months?
```

- `The rate of deflation (opposite of inflation) will be 8% or higher.`
- `The rate of deflation ([...])  will be between 4% and less than 8%.`
- `The rate of deflation ([...])  will be between 0% and less than 4%.`
- `The rate of inflation will be between 0% and less than 2%.`
- `The rate of inflation will be between 2% and less than 4%.`
- `The rate of inflation will be between 4% and less than 6%.`
- `The rate of inflation will be between 6% and less than 8%.`
- `The rate of inflation will be between 8% and less than 12%.`
- `The rate of inflation will be between 12% and less than 16%.`
- `The rate of inflation will be 16% or higher.`

In the new formulation, the center of the bins is shifted upwards by four percentage points. As a result, the bins are centered around an inflation rate of 4%, which is closer to—but still below—the actual inflation rate in May 2022 (7.9%) relative to the baseline design. We leave the number of bins as well as their widths unchanged.[3]

The sample in Wave 30 was split into three randomly assigned groups. Approximately one third of the sample (1,356 observations) was presented with the baseline design used in all previous waves. Another third of the sample (1,377 observations) was presented with the alternative design which we refer to as the 'mean-shift' setting. The remaining 1,361 observations were presented with another bin design which we do not use in our analysis.[4] Thus, our analysis focuses on the 2,733 households in the baseline group and the mean-shift group.

In the analysis below, we analyze the impact of the alternative response scale on households' probabilisitic expectations. We are particularly interested in potential differences in the shape of the histograms between the baseline group and the mean-shift group. Figure 2 shows the average responses of the individuals in both subsamples. The plot on the left depicts the average probability mass assigned to each bin across all respondents while the plot on the right shows the corresponding histogram by reporting densities instead of probabilities. The aggregate distributions clearly differ across treatments.

To assess the differences in the probabilistic expectations on an individual basis, we define the dummy variable *meanshift* that equals one if the individual belongs to the mean-shift group, and zero else. Next, we calculate the number of bins with nonzero probability (*bins*) and the probability mass assigned to the rightmost bin (*phigh*). We also define the dummy variable *multipeak* which equals one if the histogram has multiple modes, and zero else. Table A.1 in the Appendix provides details on the construction of all variables.

Finally, we compute the first four moments of each histogram. To do so, we follow Conrad et al. (2022) and assume that the probability in each bin is located at the midpoint.[5] To close the exterior bins, we assume that they have equal width to the adjacent bins, i.e., four percentage points.[6] Based on the 'mass-at-midpoint' approach, mean

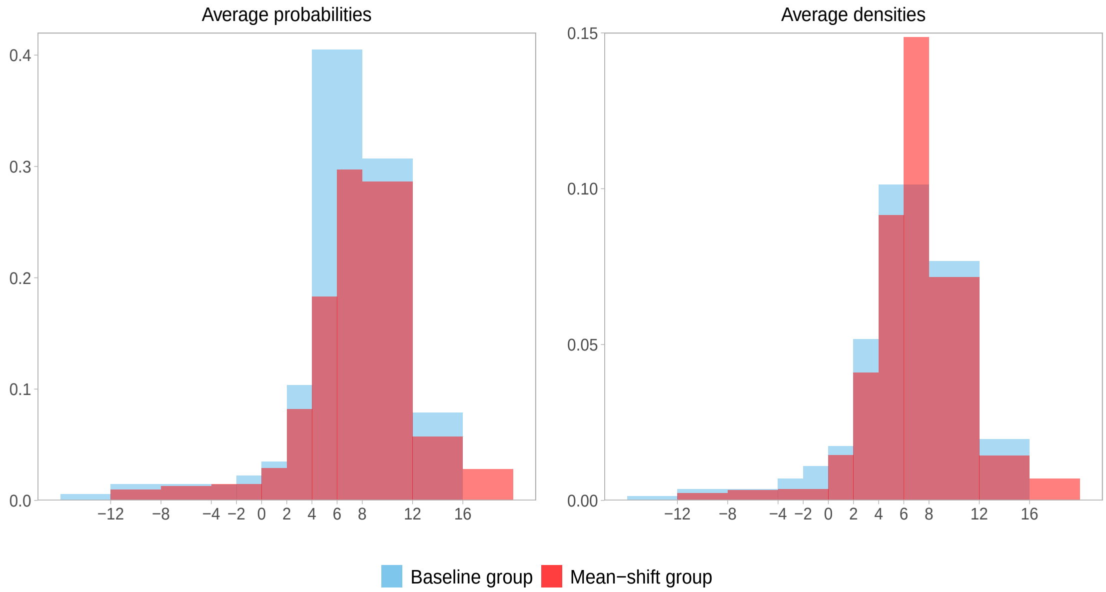
[3]Expert surveys such as the Survey of Professional Forecasters operated by the ECB and the Federal Reserve Bank of Philadelphia cover a relatively narrow outcome range. As a result, their operators frequently adjust the bin definitions in a way similar to our proposed mean-shift desgin. This is usually done in reponse to macroeconomic shocks such as the Great Recession where a considerable pile-up of probabilities in the lowest bin for GDP growth was observed in the ECB-SPF. During the Coronavirus pandemic, the ECB-SPF introduced bins with unequal width. Glas and Hartmann (2022) show that this can have an impact on the mismatch between ex-ante uncertainty as measured by the histogram variance and ex-post uncertainty based on the variability of forecast errors.

[4]This design retains the centering around 0% but includes a more granular definition of the interior bins. See Becker et al. (2023) for the motivation behind this approach. Since a takeaway from our study is that centering around an inflation rate of 0% is not appropriate in the current high-inflation regime, we do not use these observations in our analysis. However, all tables and figures for this alternative treatment are available upon request from the authors.

[5]Other alternatives include assuming uniformly distributed probabilities or fitting a continuous distribution as in Engelberg et al. (2009). However, Glas (2020) shows that this choice has little impact on estimates of the mean or the standard deviation. Moreover, Becker et al. (2022) show that fitting continuous distributions can lead to misleading results in the presence of varying interval widths.

[6]Armantier et al. (2017) and Zhao (2022) use ±38% as the bounds for the exterior bins in their analyses of the Survey of Consumer Expectations. This choice is based on historically observed inflation rates in the US. For Germany, such extreme inflation rates have not been observed. Zhao (2022) mentions in his

Figure 2: Average probabilistic expectations by treatment status



Notes: The subfigures show the average responses for the individuals in the baseline group and the mean-shift group. The left plot depicts the average probability mass in each bin while the right plot shows the histograms by reporting densities instead of probabilities.

(μ), standard deviation (σ), skewness (γ) and kurtosis (κ) of the histogram reported by household $i = 1, \ldots, n$ are calculated as follows:

$$\mu_i = \sum_{k=1}^{K} m_k \times p_{i,k} \tag{1}$$

$$\sigma_i = \sqrt{\sum_{k=1}^{K} (m_k - \mu_i)^2 \times p_{i,k}} \tag{2}$$

$$\gamma_i = \frac{\sum_{k=1}^{K} (m_k - \mu_i)^3 \times p_{i,k}}{\sigma_i^3} \tag{3}$$

$$\kappa_i = \frac{\sum_{k=1}^{K} (m_k - \mu_i)^4 \times p_{i,k}}{\sigma_i^4} \tag{4}$$

In Eqn. (1)-(4), the index $k = 1, \ldots, K$ denotes the different bins, m_k is the midpoint of the k-th bin and $p_{i,k}$ is the probability assigned to this particular bin by household i.

Panel A of Table 1 presents summary statistics for all histogram characteristics by treatment status. For skewness and kurtosis, we consider only the responses of participants who use at least three bins. On average, the individuals in the mean-shift group use

footnote 14 that he also considered ±16% for the bounds and that this choice did not affect his findings. Our choice for the bounds also makes it more difficult to detect potentially existing differences between histogram means across treatments and when comparing histogram means to point forecasts.

Table 1: Summary statistics for Wave 30 of the BOP-HH

	Baseline group					Mean-shift group				
	Obs.	Mean	SD	Min	Max	Obs.	Mean	SD	Min	Max
Panel A: Probabilistic inflation expectations										
bins	1,356	2.97	1.96	1.00	10.00	1,377	3.27	2.21	1.00	10.00
phigh	1,356	7.86	19.90	0.00	100.00	1,377	2.80	12.77	0.00	100.00
multipeak	1,356	0.08	0.26	0.00	1.00	1,377	0.08	0.27	0.00	1.00
μ_i	1,356	6.57	3.82	−14.00	14.00	1,377	7.24	3.41	−10.00	18.00
σ_i	1,356	2.02	1.74	0.00	11.34	1,377	1.82	1.74	0.00	11.72
γ_i	695	0.07	0.84	−4.14	5.02	778	0.11	0.82	−3.77	4.14
κ_i	695	3.61	2.44	1.22	29.62	778	3.53	2.24	1.08	25.73
Panel B: Point forecasts										
$\hat{\pi}_i^P$	443	6.63	2.53	0.00	20.00	435	6.67	2.55	0.00	20.00
$\hat{\pi}_i^E$	1,328	8.11	3.53	-2.00	30.00	1,350	8.14	3.32	-2.00	30.00
$\lvert\hat{\pi}_i^E - \mu_i\rvert$	1,328	2.17	3.40	0.00	25.40	1,350	1.60	2.54	0.00	26.35
Panel C: Socioeconomic characteristics										
age	1,356	56.98	14.35	17.00	80.00	1,377	56.83	14.62	16.00	80.00
east	1,356	0.17	0.38	0.00	1.00	1,377	0.17	0.37	0.00	1.00
female	1,356	0.36	0.48	0.00	1.00	1,377	0.39	0.49	0.00	1.00
fullemploy	1,356	0.44	0.50	0.00	1.00	1,377	0.42	0.49	0.00	1.00
hhsize	1,356	2.20	1.04	1.00	6.00	1,377	2.20	1.07	1.00	6.00
income	1,356	3.98	2.01	0.25	11.00	1,377	3.94	1.95	0.25	11.00
yoe	1,356	11.55	1.67	7.00	18.00	1,377	11.51	1.69	7.00	18.00

Notes: This table shows summary statistics for the probabilistic inflation expectations (Panel A), point forecasts (Panel B) and socioeconomic characteristics (Panel C) of participants in Wave 30 of the BOP-HH. For skewness and kurtosis, we focus on responses where nonzero probability is assigned to at least three bins. The samples for $\hat{\pi}_i^P$ and $\hat{\pi}_i^E$ are trimmed by 1% from top and bottom. Household income is expressed in 1,000 euro.

more bins, assign lower probability to the right-most bin, report higher histogram means and lower standard deviations.

2.2 Point forecasts

In addition to the probabilistic expectations, the BOP-HH elicits point forecasts on households' perceptions of current inflation ($\hat{\pi}_i^P$) and their expectations of inflation over the coming year ($\hat{\pi}_i^E$). In the next section, we analyze the consistency of point and probabilistic expectations via the difference between $\hat{\pi}_i^E$ and μ_i. Since it has been shown that there exists a tight link between perceived and expected inflation (Jonung, 1981; D'Acunto et al., 2021), we also consider $\hat{\pi}_i^P$, although only one third of the participants in Wave 30 were asked for their perception of the current inflation rate over the previous twelve months. To reduce the impact of outliers, we trim the top and bottom 1% of inflation perceptions/expectations. For the remaining individuals, Figure 1 above shows average inflation expectations across survey waves along with actual inflation.

Panel B of Table 1 presents summary statistics for the point forecasts by treatment status. In contrast to the probabilistic expectations, the figures for perceived and expected inflation are very similar across the two treatment groups. Notably, the average point forecast exceeds the average histogram mean in both cases. However, due to the higher average histogram mean for the mean-shift group relative to the baseline group, the average absolute deviation between point forecasts and histogram means is markedly lower for this particular group.

The average perceived inflation rate (calculated as the weighted average across the two groups) is 6.65%. For comparison, the most recent inflation figure available to Wave 30 participants was the German inflation rate in May 2022 (7.9%) since all responses were collected between 15 June and 29 June and the May 2022 inflation rate was released by the German statistical office on 14 June. Only one response was elicited on 29 June when the first estimate of the inflation rate in June was released (7.6%). Thus, the average participant in Wave 30 *underestimates* current inflation. This finding contrasts the evidence in Conrad et al. (2022) who find that BOP-HH participants in Wave 3 overestimated inflation in May 2019. Their results are consistent with our data before June 2021 (see Figure 1). Weber et al. (2022) list a 'systematic upward bias' as a stylized fact of households' inflation perceptions/expectations. Our finding suggests that this may not generally be the case in high-inflation regimes or that households are slow to adjust their beliefs. However, the weighted average of expected inflation is 8.12%. Thus, households appear to take notice of the surge in inflation rates. This is supported by the upward trend in inflation expectations shown in Figure 1. The correlation between perceived and expected inflation is 0.46.

2.3 Socioeconomic characteristics

In addition to households' inflation expectations, we use information about their socioeconomic status. We consider age (*age*), gender (*female*), employment status (*fullemploy*), whether the individual lives in East or West Germany (*east*), household size (*hhsize*), income (*income*) and years of education (*yoe*). These variables have been shown to be robust predictors of households' macroeconomic expectations and uncertainty thereof (Bruine de Bruin et al., 2010, 2011; Das et al., 2020). In all regressions below, we use the natural logarithm of income as a covariate. We include these characteristics to improve the efficiency of the estimates.

Panel C of Table 1 presents summary statistics for the socioeconomic characteristics of the participants in Wave 30 by treatment status. The average respondent in Wave 30 is 57 years old and has almost 12 years of education. 38% of the individuals are female, 43% are full-time employed and 17% live in East Germany.

As with the point forecasts, socioeconomic characteristics are distributed similarly in both treatment groups, suggesting that the treatment is indeed randomly assigned. We confirm that this is the case by running a linear regression of the *meanshift*-dummy on the socioeconomic variables. The baseline is the group of households that were presented with the original bin design. Table A.2 in the Appendix presents the results. As expected, none of the coefficients are significantly different from zero, which suggests that the random assignment of treatments was successful.

Table 2: Inflation expectations and socioeconomic characteristics: baseline group

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
age	-0.03***	-0.08*	0.00	0.02**	-0.02***	-0.00	0.02***	0.01	-0.01
	(0.00)	(0.04)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
east	-0.10	4.10**	-0.01	0.76***	-0.11	0.01	0.12	0.73***	-0.09
	(0.15)	(1.74)	(0.02)	(0.27)	(0.13)	(0.09)	(0.36)	(0.26)	(0.23)
female	-0.16	4.88***	0.06***	0.50**	0.12	-0.17**	0.03	0.93***	0.60***
	(0.11)	(1.24)	(0.02)	(0.24)	(0.11)	(0.07)	(0.20)	(0.22)	(0.22)
fullemploy	-0.27**	0.57	-0.01	0.28	-0.28**	0.03	0.13	0.51**	0.24
	(0.13)	(1.30)	(0.02)	(0.25)	(0.11)	(0.07)	(0.21)	(0.25)	(0.24)
hhsize	0.02	0.97	-0.01	0.16	-0.03	-0.07**	-0.02	0.18	0.07
	(0.06)	(0.71)	(0.01)	(0.13)	(0.05)	(0.03)	(0.09)	(0.13)	(0.12)
ln(*income*)	0.07	-2.59**	-0.01	-0.45*	0.03	0.09	0.08	-0.89***	-0.41*
	(0.11)	(1.30)	(0.01)	(0.25)	(0.09)	(0.07)	(0.21)	(0.22)	(0.22)
yoe	0.01	-0.86**	-0.01***	-0.06	-0.04	0.06***	0.10*	-0.14**	-0.10*
	(0.03)	(0.40)	(0.00)	(0.07)	(0.03)	(0.02)	(0.06)	(0.06)	(0.06)
Constant	4.05***	38.69***	0.26**	9.11***	3.46***	-1.12**	0.60	15.50***	6.59***
	(0.93)	(9.65)	(0.11)	(2.14)	(0.82)	(0.56)	(1.63)	(1.67)	(1.80)
Observations	1,356	1,356	1,356	1,356	1,356	695	695	1,328	1,328
$\bar{R}^2$	0.03	0.03	0.02	0.01	0.02	0.02	0.01	0.04	0.01

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on socioeconomic characteristics. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

3 Results

This section presents our empirical findings. We briefly consider the relationship between inflation expectations and socioeconomic status before analyzing in which aspects the inflation expectations differ between the baseline and the mean-shift group. Next, we assess the implications of our results for the consistency between point forecasts and histogram means and explore potential heterogeneity in the estimated treatment effects. Lastly, we consider revisions in inflation expectations from Wave 29 (before the treatment) to Wave 30 as well as revisions from Wave 30 to Wave 31 (after the treatment).

3.1 Histogram characteristics and socioeconomic status

In a first step, we relate the histogram characteristics and point forecasts of BOP-HH participants to their socioeconomic status. Table 2 presents the estimates for the individuals in the baseline group. Columns (1)-(3) show the results for the number of bins with nonzero probability, the probability mass in the rightmost bin and the indicator for multimodal histograms. Columns (4)-(7) present the estimates for the histogram moments. Columns (8)-(9) show the findings for the point forecasts and the absolute deviations between point forecasts and histgram means. All regressions are estimated with heteroskedasticity-consistent standard errors.

Consistent with Armantier et al. (2021), we find that older respondents have significantly higher histogram means and lower inflation uncertainty as a result of using fewer bins. In addition, kurtosis increases with age. The *east*-dummy has a significantly positive effect on histogram means and point forecasts. This is in line with Goldfayn-Frank and Wohlfart (2020) who show that East Germans have higher inflation expectations than West Germans—especially at times when inflation is unusually high—due to the inflationary shock experienced after reunification. Next, we find that women assign more probability mass to the rightmost bin and have higher inflation expectations both in terms of histogram means and point forecasts. These findings square with similar evidence in Bruine de Bruin et al. (2011), Armantier et al. (2021) and Conrad et al. (2022). In addition, the probability of reporting a multi-peaked probability distribution is significantly higher for women, the histograms of women are more left-skewed than those of men and their point forecasts and histogram means tend to deviate more strongly. Full-time employed individuals use fewer bins, have lower uncertainty and higher point forecasts (but not histogram means). Household size appears to matter little beyond a negative effect on skewness. Higher income is associated with a lower probability mass in the rightmost bin (as in Armantier et al., 2021), lower point forecasts and a higher degree of consistency between point forecasts and histogram means (see Zhao, 2022). Lastly, higher education is associated with less probability in the rightmost bin, a lower probability of stating a multi-peaked distribution, higher skewness and kurtosis, lower point forecasts and smaller deviations between point forecasts and histogram means. The findings that high-income households and highly educated individuals have lower point forecasts are consistent with Bruine de Bruin et al. (2010).

Overall, our results are in line with typical findings in the literature (Das et al., 2020). In the following analyses, we use each respondents' socioeconomic characteristics as control variables in all regressions. Since treatment assignment is unrelated to socioeconomic characteristics (see Table A.2), these variables are included primarily to increase the efficiency of the estimates.

Table A.3 in the Appendix shows that the relationship between inflation expectations and socioeconomic status is similar for the individuals in the mean-shift group with a few exceptions. For example, the coefficients on the *east*-dummy in Columns (2) and (4) are insignificant for the mean-shift group, while the coefficient in Column (8) is significant only at the 10% level. Similarly, education does not have a significant effect on histogram characteristics or point forecasts. Finally, the estimated effects of household income are larger and more significant in the regressions for the mean-shift group. These findings may hint at potential cross-sectional heterogeneity in the response of individuals when confronted with the alternative bin design. We analyze this issue in Section 3.3.

3.2 Differences in inflation expectations by treatment status

Having established the role of socioeconomic characteristics for inflation expectations, we now consider differences in expectations between the baseline group and the mean-shift group. Table 3 presents the estimates from linear regressions of inflation expectations on treatment status and socioeconomic characteristics (the latter are not shown) for the pooled sample of observations from both bin designs.

3.2.1 Differences in histogram characteristics

The histogram characteristics in Columns (1)-(7) are potentially affected by the alternative bin design. Indeed, we observe some noticeable differences in the histogram char-

Table 3: Differences in inflation expectations across treatments

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.30***	-5.16***	-0.00	0.65***	-0.20***	0.05	-0.07	0.01	-0.59***
	(0.08)	(0.64)	(0.01)	(0.14)	(0.07)	(0.04)	(0.12)	(0.13)	(0.12)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status and socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

acteristics of both groups. In particular, we find that those in the mean-shift group use significantly more bins, assign a considerably lower probability mass to the rightmost bin, report higher histogram means and have lower inflation uncertainty than those in the baseline group. These effects are also economically significant. For example, Column (2) shows that the average probability mass in the rightmost bin is more than five percentage points lower for the mean-shift group than for the baseline group. This corresponds to the vertical difference in the right plot of Figure 1. Similarly, Column (4) shows that the histogram means in the mean-shift group are, on average, 0.65 percentage point higher than those in the baseline group.

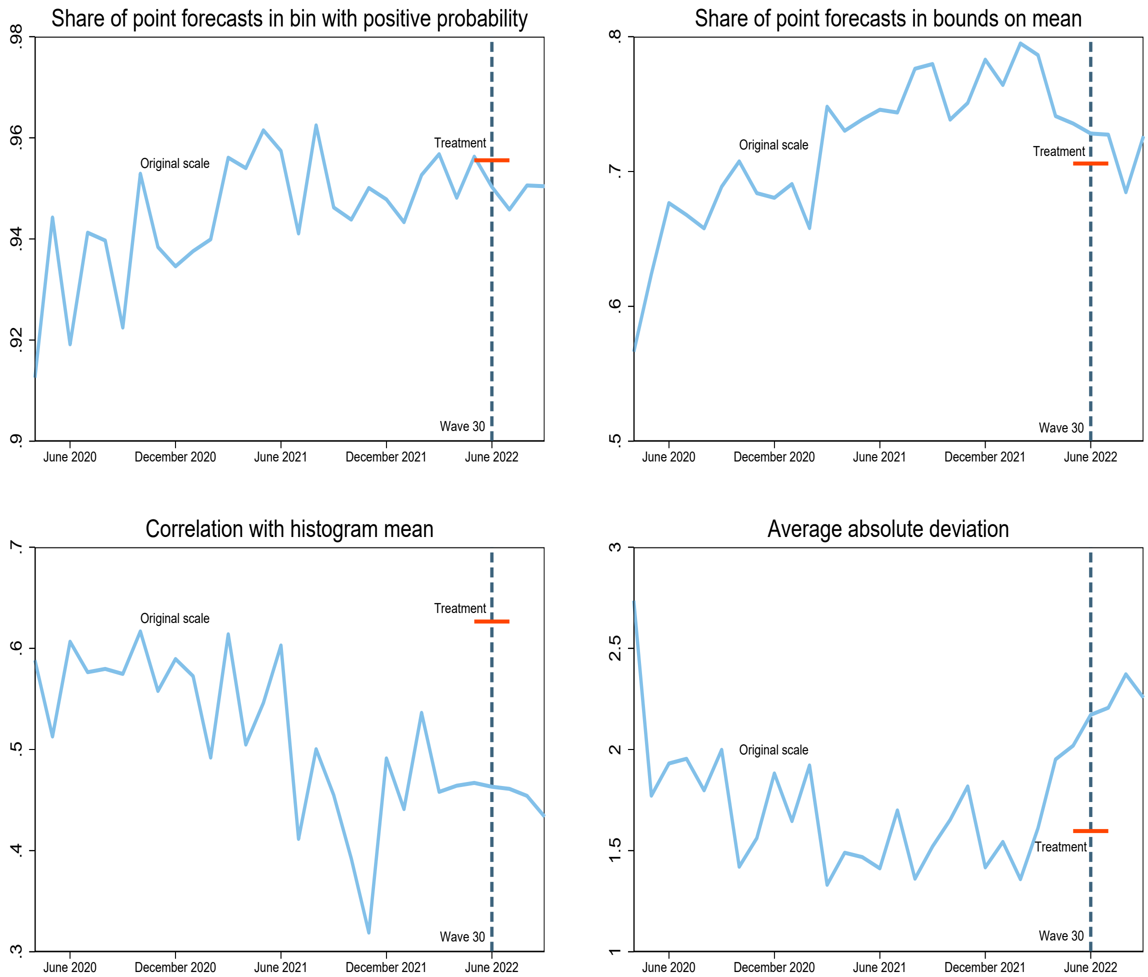
Given that all other factors such as the macroeconomic environment or the remaining questions in the survey questionnaire were identical for all respondents, the observed differences in the histogram characteristics are either due to genuinely higher expectations, different discretization biases or framing effects. However, the upward shift in the average histogram mean remains below the upward shift in the bin definitions (0.65 percentage point versus four percentage points), suggesting that participants do not simply relocate their subjective distributions around the new center of the bin design. We provide a more detailed discussion of these issues in Section 4.

3.2.2 Consistency of point forecasts and probabilistic expectations

Bruine de Bruin et al. (2011) and Zhao (2022) find that the point forecasts of US households are well aligned with measures of central tendency such as the histogram mean. To assess whether this also is the case for German households, Figure 3 shows various measures of consistency between point forecasts and histogram means.

Around 95% of households report point forecasts that fall into a bin to which the respondents assigns nonzero probability. In line with the findings in Zhao (2022), approximately 70% of point forecasts lie within the individual bounds on the histogram mean, which are calculated by replacing the midpoint m_k in Eqn. (1) with the lower bound l_k and the upper bound u_k (see Engelberg et al., 2009, for details). These findings imply that point forecasts and probabilistic expectations of German households are relatively well-aligned and supplement the evidence for the US. However, the correlation between point forecasts and histogram means exhibits a declining trend over time. Similarly, the

Figure 3: Consistency between point forecasts and probabilistic expectations



Notes: For each BOP-HH wave, the upper-left plot shows the share of point forecasts that fall into a bin to which the respondent assigns nonzero probability. The upper-right plot depicts the share of point forecasts that lie within the bounds on the histogram mean. The lower-left plot presents correlations between point forecasts and histogram means. The lower-right plot shows the average absolute deviation between point forecasts and histogram means. Point forecasts are trimmed by 1% from top and bottom. The red bars are the corresponding figures for the mean-shift group in Wave 30.

average absolute deviation between $\hat{\pi}_i^E$ and μ_i has increased in recent waves. These results suggest that the alignment between point forecasts and histograms suffers at times when households are forced to assign considerable probablity to the exterior bins as was the case in recent BOP-HH waves (see Figure 1). At the same time, the last two subfigures show a much higher degree of consistency for the mean-shift group in Wave 30. For example, the correlation between $\hat{\pi}_i^E$ and μ_i in Wave 30 is 0.63 for the mean-shift-group but only 0.46 for the baseline group. In light of these findings, we consider differences in the point forecasts and their alignment with the histogram means across treatment groups in the next step.

Column (8) of Table 3 shows that the point forecasts of individuals in the baseline group and the mean-shift groups are not significantly different from each other. In fact, the estimated coefficient on the *meanshift*-dummy is essentially zero. This is to be expected as the point forecast is elicited before the probabilistic expectation and respondents cannot return to this question later, this provides another confirmation that the randomization of treatments was successful. Our combined findings of significantly higher histogram means for the mean-shift group and stable point forecast across both groups suggest that the consistency between point forecasts and probabilistic expectations may be higher for one of the two groups. Indeed, Column (9) shows that the average absolute deviation between point forecasts and histogram means is significantly smaller in the mean-shift group. The effect size of almost 0.6 percentage point is economically relevant and similar in magnitude to the observed difference in the histogram means across groups.

In sum, the findings in columns (2), (4), (8) and (9) suggest that participants in the mean-shift group are able to more adequately communicate their higher probabilistic beliefs about future inflation. This, in turn, leads to a higher degree of consistency between the point forecasts and the probabilistic expectations reported by those individuals.

3.3 Heterogeneity in treatment effects

In this section, we analyze potential heterogeneity in the estimated treatment effects by including interaction terms between the treatment indicator and several characteristics of BOP-HH participants.

In a first step, we consider interactions of treatment status with socioeconomic characteristics. If households with different socioeconomic background react differently when presented with an alternative bin design, it may be recommendable to stick to the baseline design in order not to introduce additional distortions to the histogram characteristics. Tables A.4-A.10 in the Appendix present the results. Overall, we find no evidence that the treatment effects significantly vary in the cross-section of households.

In a recent paper, Weber et al. (2022) notes that repeated participation may induce individuals to learn about a specific topic or details of the survey questionnaire. This effect is known as 'panel conditioning' and can also apply to the probabilistic expectations. Of the 2,733 households in our sample for Wave 30, 196 (7%) participated in the BOP-HH for the first time. 106 of these individuals are assigned to the baseline group and the other 90 to the mean-shift group. The remaining 2,537 individuals participated at least once before. The impact of our treatment on the probabilistic expectations may be stronger for more experienced survey participants. New entrants could simply assume that the alternative bin design represents the standard approach. On the other hand, it may be argued that participants with previous experience in the BOP-HH are somewhat 'anchored' around the original bin design. To explore these issues, we consider interactions between the *meanshift*-dummy and an indicator variable for first-time participants (*firsttimer*). Table 4 presents the results.

We find that, on average, first-time participants in Wave 30 assign nonzero probability to 0.78 more bins relative to more experienced respondents and assign over four percentage points of additional probability mass to the rightmost bin. As a result, new panelists provide more dispersed histograms that also tend to be more left-skewed than those of more households with more survey experience. They also have a higher probability of reporting multi-peaked probability distributions. In contrast, the point forecasts, histogram means and kurtosis of new entrants do not differ significantly from those of other participants. Importantly for our analysis, the interaction between *meanshift* and *firsttimer* is

Table 4: Differences in inflation expectations: interaction with *firsttimer*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.31***	-5.04***	0.01	0.68***	-0.17**	0.05	-0.12	-0.00	-0.63***
	(0.08)	(0.65)	(0.01)	(0.14)	(0.07)	(0.05)	(0.13)	(0.13)	(0.12)
firsttimer	0.78***	4.19*	0.11***	0.28	0.82***	-0.18*	-0.35	0.27	-0.07
	(0.23)	(2.17)	(0.04)	(0.33)	(0.22)	(0.09)	(0.22)	(0.35)	(0.29)
meanshift × *firsttimer*	-0.01	-0.96	-0.09*	-0.36	-0.37	-0.07	0.58	0.26	0.66
	(0.34)	(3.09)	(0.05)	(0.54)	(0.28)	(0.15)	(0.37)	(0.60)	(0.45)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.03	0.02	0.02	0.02	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, a dummy variable for first-time participants, an interaction with treatment status and socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table 5: Differences in inflation expectations: interaction with *ninterest*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.27***	-4.94***	-0.00	0.72***	-0.23***	0.05	-0.05	0.01	-0.64***
	(0.08)	(0.63)	(0.01)	(0.14)	(0.07)	(0.04)	(0.12)	(0.13)	(0.12)
ninterest	-0.28	7.33	-0.05*	0.85	-0.55**	-0.23	0.80	-0.21	-0.78***
	(0.31)	(4.84)	(0.02)	(0.51)	(0.25)	(0.24)	(0.58)	(0.57)	(0.24)
meanshift × *ninterest*	0.94*	-6.47	0.04	-2.08**	0.75*	0.01	-0.59	-0.01	1.46**
	(0.55)	(5.45)	(0.05)	(0.87)	(0.40)	(0.30)	(0.73)	(0.95)	(0.71)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.03	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, a dummy variable that indicates whether respondents found the BOP-HH questionnaire uninteresting, an interaction with treatment status and socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

insignificant for all dependent variables except *multipeak*. This suggests that experienced participants do not react differently when presented with the alternative designs compared to new entrants who are confronted with questions about their probabilistic expectations for the first time.

In the last step, we consider interactions between treatment status and characteristics that capture the engagement of respondents with the survey. Table 5 shows the estimates of interacting treatment status with an indicator variable that states whether the respondent found the survey not interesting (*ninterest*), which is the case for 88 out of the 2,733 participants (3%) in Wave 30.

We find a significantly negative interaction between *meanshift* and *ninterest* in Column (4). This implies that while interested individuals in the mean-shift group report higher histogram means than the baseline group, the mean of uninterested individuals is, on average, more than one percentage point lower. In other words, those respondents tend to report lower inflation expectations than the baseline group despite the bins being moved towards higher inflation rates. As a result, the mismatch between point forecasts and histogram means tends to be *higher* for those individuals relative to the baseline group, whereas the opposite is the case for interested individuals in the treatment group. We also find that uninterested individuals in the baseline group express considerably lower inflation uncertainty than interested respondents in the baseline group. In contrast, uninterested individuals in the mean-shift group tend to report higher inflation uncertainty than individuals in the baseline group, whereas interested individuals in the mean-shift group tend to report lower standard deviations. In light of these findings, it may be recommendable to discard uninterested individuals from the sample altogether.[7]

We ran similar regressions with dummy variables that indicate whether the respondent found the survey too difficult (*difficult*, 8% of respondents) or too long (*toolong*, 22%). Tables A.12-A.13 present the estimates. While the results point in a similar direction as those for *ninterest*, the estimates are insignificant in most cases. However, we note that individuals that assign a high degree of difficulty to the survey questionnaire tend to report significantly different histogram moments than those who consider the survey as rather easy to answer. Moreover, for those individuals we also observe significant differences in the estimated treatment effects for higher moments such as skewness and kurtosis.

3.4 Revisions of histogram moments

The rotating panel structure of the BOP-HH allows us to not only analyze differences in the point forecasts and probabilistic expectations in the cross-section of households, but also changes in revisions of such variables over time. In particular, we analyze i) how individuals who participated in Waves 29 to 31 updated their probabilistic expectations across time and ii) whether such revisions differ for those in the baseline group relative to those in the mean-shift group.

Of the 2,733 households in our sample for Wave 30, 738 also participated in Wave 29 and Wave 31. 368 of these respondents are in the baseline group and 370 in the mean-shift group. For those individuals we can compute revisions in point forecasts and histogram moments. For example, the revision of the histogram mean between Wave 29 and Wave 30 is defined as $\Delta\mu_i = \mu_{i,June} - \mu_{i,May}$. Similarly, $\Delta\mu_i = \mu_{i,July} - \mu_{i,June}$ is the corresponding revision between Wave 30 and Wave 31. The calculation for revisions of other variables proceeds analoguously. Table A.14 in the Appendix replicates Table 3 for the subset of respondents that participated in Waves 29 through 31. The estimates are very similar to our main results, although the magnitude of the effects tends to be slightly higher.

3.4.1 Updating from Wave 29 to Wave 30

While it is expected that some participants update their expectations from one period to the next, the magnitude of these changes can differ between treatment groups. In particular, if the differences between baseline and mean-shift groups in Table 3 can truly be ascribed to the treatment, the differences in revisions of histogram characteristics between Waves 29 and 30 should be similar in size to the estimated treatment effects. Table 6

[7] Table A.11 shows that our main results are very similar when focusing only on interested respondents.

Table 6: Differences in revisions of inflation expectations between Wave 29 and Wave 30

	(1) $\Delta bins$	(2) $\Delta phigh$	(3) $\Delta multipeak$	(4) $\Delta\mu_i$	(5) $\Delta\sigma_i$	(6) $\Delta\gamma_i$	(7) $\Delta\kappa_i$	(8) $\Delta\hat{\pi}_i^E$	(9) $\Delta\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.02	-3.94***	-0.03	0.63**	-0.42***	0.10	0.07	0.29	-0.54**
	(0.12)	(1.41)	(0.02)	(0.29)	(0.11)	(0.11)	(0.23)	(0.23)	(0.27)
Observations	738	738	738	738	738	386	386	713	713
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.00	0.03	0.01	0.00	0.03	-0.01	0.01	0.00	0.00

Notes: This table presents the estimates from linear regressions of revisions of histogram characteristics and point forecasts between Wave 29 and 30 on treatment status and socioeconomic characteristics for the subset of individuals that participated in Waves 29, 30 and 31. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

presents the results from linear regressions of revisions of point forecasts and histogram moments on the treatment indicator variables and socioeconomic characteristics.

We find significant differences in revisions between the baseline group and the mean-shift group for the probability mass assigned to the rightmost bin, the histogram mean and the standard deviation. As expected, the differences in revisions between baseline and mean-shift group are closely associated with the size of the estimated treatment effects in Table 3. This further reinforces the notion that the observed differences can indeed be ascribed to the alternative bin design. Similarly, the coefficient on $\Delta|\hat{\pi}_i^E - \mu_i|$ is significant and the effect size is close to the corresponding estimate in Table 3.

3.4.2 Updating from Wave 30 to Wave 31

Next, we assess differences in revisions between Waves 30 and 31, i.e., immediately after we conducted our experiment. Individuals who were assigned to the mean-shift group in Wave 30 are now again presented with the baseline bin definitions. We are interested in the question of whether those individuals now revise their probabilistic expections as strongly in the opposite direction as they did when they were originally presented with the alternative bin design. Table 7 presents the estimates we obtain when replacing the revisions between Waves 29 and 30 with the revisions between 30 and 31. The socioeconomic characteristics are now drawn from Wave 31 instead of Wave 30.

When presented again with the original bin design, the individuals in the mean-shift group react by significantly reducing the number of bins, assigning considerably higher probability mass to the rightmost bin and reporting significantly lower histogram means and higher standard deviation. The significant estimates have the opposite sign as those in Table 6 and are similar in size. For the probability assigned to the rightmost bin and the misalignment between point forecasts and histogram means, the difference in revisions is even larger, which suggests that participants do not completely revert back to their pre-treatment expectations. Instead, they seem to partially retain their higher distribution from the mean-shift setting.

Table 7: Differences in revisions of inflation expectations between Wave 30 and Wave 31

	(1) $\Delta bins$	(2) $\Delta phigh$	(3) $\Delta multipeak$	(4) $\Delta\mu_i$	(5) $\Delta\sigma_i$	(6) $\Delta\gamma_i$	(7) $\Delta\kappa_i$	(8) $\Delta\hat{\pi}_i^E$	(9) $\Delta\|\hat{\pi}_i^E - \mu_i\|$
meanshift	-0.26**	7.11***	-0.01	-0.57**	0.25**	-0.09	0.10	-0.22	0.72***
	(0.10)	(1.39)	(0.02)	(0.28)	(0.10)	(0.12)	(0.31)	(0.22)	(0.26)
Observations	738	738	738	738	738	354	354	716	716
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.01	0.04	0.00	0.00	0.02	0.02	-0.01	0.01	0.01

Notes: This table presents the estimates from linear regressions of revisions of histogram characteristics and point forecasts between Wave 30 and 31 on treatment status and socioeconomic characteristics for the subset of individuals that participated in Waves 29, 30 and 31. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

4 Discussion

We show that the mean-shift setting affects the probabilistic expectations of BOP-HH participants by allowing them to communicate more clearly their true beliefs at times when inflation is unusually high. However, other factors may also contribute to the observed deviations between treatment groups as discussed below.

One alternative explanation is that the differences in responses are driven by a central tendency bias, i.e., some respondents may believe that values close to the center of the distribution—zero for the baseline group, four for the mean-shift group—are deemed more likely by the Bundesbank (see Becker et al., 2023). Table 1 shows that the histograms in the baseline (mean-shift) group are centered around an average inflation rate of 6.57 (7.24) percent. These values are far away from the center of the respective distribution. Moreover, the difference in average histogram means of 0.67 percentage points is much smaller than the shift in the bins of four percentage points for the treatment group. These findings are more consistent with the interpretation that households are able to better state their true beliefs in the alternative setting rather than them using the center of the distribution as a focal point for their probabilistic expectations.

A second explanation is that at least some of the differences can be ascribed to different discretizations of the scale across treatments which affects histogram moments even under the assumption of stable beliefs. To explore the magnitude of such 'technical errors', we consider a hypothetical setting where a household with fixed probabilistic expectations is confronted with the two bin designs. The expectations of this household are normally distributed with known mean μ_0 and variance σ_0^2, i.e., $\mathcal{N}(\mu_0, \sigma_0^2)$. While it is unrealistic to assume that all households have normally distributed expectations, this may be an appropriate assumption for highly educated respondents. Also, Table 1 shows that the average skewness and kurtosis of BOP-HH participants are close to values expected under normality. For the mean, we choose $\mu_0 \in \{0, 4, 8\}$, where a value of zero corresponds to the center of the bin definitions for the baseline group, four corresponds to the center of the definitions for the mean-shift group and eight is close to the actual inflation rate in May 2022 (7.9%). For the variance, we consider $\sigma_0^2 \in \{4, 9\}$ to capture settings with low and high inflation uncertainty. For each combination of μ_0, σ_0^2 and the bin definitions, we

Table 8: Histogram moments under stable expectations

	Baseline group	Mean-shift group	Baseline group	Mean-shift group
	$\mathcal{N}(0,4)$		$\mathcal{N}(0,9)$	
μ	0.00	−0.23	0.00	−0.14
σ^2	4.77	4.85	10.46	9.87
γ	0.00	0.16	0.00	0.05
κ	3.61	2.80	3.03	3.07
	$\mathcal{N}(4,4)$		$\mathcal{N}(4,9)$	
μ	4.23	4.00	4.14	4.00
σ^2	4.85	4.77	9.87	10.46
γ	−0.16	0.00	−0.05	0.00
κ	2.80	3.61	3.07	3.03
	$\mathcal{N}(8,4)$		$\mathcal{N}(8,9)$	
μ	8.02	8.23	8.04	8.14
σ^2	5.24	4.85	9.52	9.87
γ	0.12	−0.16	0.02	−0.05
κ	2.24	2.80	2.74	3.07

Notes: For both bin definitions, this table presents the empirical histogram moments derived under the assumption that respondents have normally distributed inflation expectations.

calculate the probability mass assigned to each bin and compute the histogram moments using Eqn. (1)-(4). Table 8 presents the results. To faciliate the comparison between true and empirical moments, we report variances instead of standard deviations.

While the empirical histogram moments clearly deviate across settings, they are usually fairly close to the true values. The absolute difference between the empirical histogram means across treatments is at most 0.23 percentage point in case of the setting with low uncertainty and small values of μ_0. This is much smaller than the estimated difference of 0.65 percentage point between baseline and mean-shift group in Column (4) of Table 3. Turning to the variances, we observe that the empirical variances exceed their true value in all settings. The largest difference between empirical variances across bin definitions—0.59 percentage point in absolute terms—is observed for the high-uncertainty scenario and small values of μ_0. This corresponds to an absolute difference in standard deviations of 0.09 percentage point. In contrast, Table 3 Column (5) shows that the estimated difference in standard deviations between treatment groups is more than twice as large. We conclude that our estimated treatment effects are too large to merely be the result of different discretizations across bin definitions.

5 Conclusion

For the current high-inflation environment, we find evidence that the moments of households' probabilistic inflation expectations vary with the response scale used to elicit them.

In our sample, this is particularly the case for the histogram mean. As a result, the wedge between point forecast and histogram mean depends on the setup used for the probabilistic expectations. We show that the histogram variance is also affected. These findings do not appear to be the result of a central tendency bias or due to the use of different discretizations under the assumption of constant expectations. Rather, our results suggest that the inflation beliefs of German households have shifted upwards on average. Using the original scale to elicit expectations under these new beliefs tends to distort histogram moments as respondents have to allocate more probability mass to the higher, half-open interval in order to state their expectations. While we find the mean and variance to be affected, higher moments such as skewness and kurtosis appear to be relatively robust.

Our results have important implications for survey operators because they suggest that the interval design in household surveys could, and indeed should, be adjusted to the current macroeconomic environment as it is commonly done in surveys of professional forecasters. A more fine-grained interval design might also be advisable to accurately capture inflation expectations once inflation surges. However, such adjustments come at the cost of the comparibility across different household surveys. Another alternative would be to use sample splits such as the one used in this paper at times when inflation is unusually low or high. While some of the participants receive the original design to retain consistency with previous waves, the remaining panelists are confronted with an alternative design where the center of the bin is closer to the actual inflation rate.

References

Armantier, O., G. Kosar, 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. V. der Klaauw, and B. Zafar (2017). An overview of the Survey of Consumer Expectations. Staff Reports 800, Federal Reserve Bank of New York.

Becker, C., P. Duersch, and T. Eife (2023). Measuring inflation expectations: How stable are density forecasts? Mimeo.

Becker, C., P. Duersch, T. Eife, and A. Glas (2022). Extending the procedure of Engelberg et al. (2009) to surveys with varying interval-widths. FAU Discussion Papers in Economics No. 05/2022.

Beckmann, E. and T. Schmidt (2020). Bundesbank Online Pilot Survey on Consumer Expectations. Technical Paper.

Bruine de Bruin, W., C. F. Manski, G. Topa, and W. van der Klaauw (2011). Measuring consumer uncertainty about future inflation. *Journal of Applied Econometrics 26*(3), 454–478.

Bruine de Bruin, W., W. van der Klaauw, J. 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*(2), 381–402.

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

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

D'Acunto, F., D. Hoang, M. Paloviita, and M. Weber (2021). Human frictions in the transmission of economic policies. NBER Working Paper 29279.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021). Exposure to grocery prices and inflation expectations. *Journal of Political Economy 129*(5), 1615–1639.

Das, S., C. M. Kuhnen, and S. Nagel (2020). Socioeconomic status and macroeconomic expectations. *The Review of Financial Studies 33*(1), 395–432.

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*(1), 30–41.

Glas, A. (2020). Five dimensions of the uncertainty-disagreement linkage. *International Journal of Forecasting 36*(2), 607–627.

Glas, A. and M. Hartmann (2022). Uncertainty measures from partially rounded probabilistic forecast surveys. *Quantitative Economics 13*(3), 979–1022.

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.

Jonung, L. (1981). Perceived and expected rates of inflation in Sweden. *The American Economic Review 71*(5), 961–968.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? *The Quarterly Journal of Economics 126*(1), 373–416.

Malmendier, U. and S. Nagel (2016). Learning from inflation experiences. *The Quarterly Journal of Economics 131*(1), 53–87.

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

Schwarz, N. (2010). *Measurement as Cooperative Communication: What Research Participants Learn from Questionnaires*, pp. 43–61. Los Angeles: Sage.

SOEP-IS Group (2017). SOEP-IS 2015.1-PGEN: Person-related status and generated variables. SOEP Survey Papers 462: Series D - Variable Descriptions and Coding. Berlin: DIW Berlin/SOEP.

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*(3), 157–84.

Zhao, Y. (2022). Internal consistency of household inflation expectations: Point forecasts vs. density forecasts. *International Journal of Forecasting*.

Appendix

Table A.1: Variable construction

Variable	BOP-HH Questionnaire	Description
Probabilistic inflation expectations		
meanshift	*drandom2*	Equals one if the respondent belongs to the mean shift group (*drandom2*= 2), and zero for those in the baseline group (*drandom2*= 1).
bins	*infexprob_[a-j]* (CM004), *infexprob_rct1_[a-j]* (P3001A)	Number of bins to which the respondent assigns nonzero probability.
phigh	*infexprob_j* (CM004), *infexprob_rct1_j* (P3001A)	Probability mass assigned by the respondent to the highest available bin.
multipeak	same as *bins*	Equals one if the respondent provides a histogram with multiple peaks, and zero otherwise.
μ_i	same as *bins*	Mean of the histogram forecast for the German inflation rate over the next twelve months. We assume that the exterior bins have a width of four percentage points and that the probability mass in each bin is located at the midpoint.
σ_i	same as *bins*	Standard deviation of the histogram forecast.
γ_i	same as *bins*	Skewness of the histogram forecast.
κ_i	same as *bins*	Kurtosis of the histogram forecast.
Point forecasts		
$\hat{\pi}_i^P$	*devinfpoint* (CQ002)	Perceived German inflation rate over the previous twelve months in percent. This question was only asked to approximately one third of the participants in Wave 30.
$\hat{\pi}_i^E$	*infdef* (CM002) and *inflexppoint* (CM003)	Expected German inflation rate over the next twelve months in percent. Equals *inflexppoint* if *infdef* equals 'Inflation' and $(-1)\cdot$ *inflexppoint* if *infdef* equals 'Deflation'.
$\lvert\hat{\pi}_i^E - \mu_i\rvert$	same as $\hat{\pi}_i^E$ and μ_i	Absolute difference between the point forecast and the histogram mean.
Socioeconomic characteristics		
age	*age*	Age of individual. Set to 80 if *age* equals '80 years or older'.
east	*region*	Equals one if *region* equals 'east', and zero otherwise.
female	*gender*	Equals one if *gender* equals 'female', and zero otherwise.
fullemploy	*employ* (CS003)	Equals one if *employ* equals 'employed, full-time', and zero otherwise.
hhsize	*hhsize* (CS006)	Household size. Set to 6 if *hhsize* equals '6 or more'.
income	*hhinc* (CS008)	Monthly household income in €1,000 (using bin midpoints): = 0.25 if *hhinc* equals 'Less than €500', = 0.75 if *hhinc* equals '€500 to €999', = 1.25 if *hhinc* equals '€1,000 to €1,499', = 1.75 if *hhinc* equals '€1,500 to €1,999', = 2.25 if *hhinc* equals '€2,000 to €2,499', = 2.75 if *hhinc* equals '€2,500 to €2,999', = 3.25 if *hhinc* equals '€3,000 to €3,499', = 3.75 if *hhinc* equals '€3,500 to €3,999', = 4.50 if *hhinc* equals '€4,000 to €4,999', = 5.50 if *hhinc* equals '€5,000 to €5,999', = 7.00 if *hhinc* equals '€6,000 to €7,999', = 9.00 if *hhinc* equals '€8,000 to €9,999', = 11.00 if *hhinc* equals '€10,000 or more'.

Notes: This table describes the construction of the variables used in the empirical analysis. In the middle column, we refer to the names of the original variables as listed in the questionnaire for Wave 30 (June 2022) of the BOP-HH.

Table A.1: Variable construction (cont.)

Variable	BOP-HH Questionnaire	Description
yoe	*eduschool* (CS001)	Years of education of individual following SOEP-IS Group (2017): = 7 if *eduschool* equals 'No school-leaving certificate', = 9 if *eduschool* equals 'Secondary school-leaving certificate', = 10 if *eduschool* equals 'Other school-leaving certificate', = 10 if *eduschool* equals 'Intermediate secondary school certificate', = 10 if *eduschool* equals 'Polytechnical secondary school certificate (8th/10th grade)', = 13 if *eduschool* equals 'University of applied sciences entrance diploma / completed technical school', = 13 if *eduschool* equals 'Senior school-leaving certificate/ general or subject-specific university entrance diploma', = 18 if *eduschool* equals 'College / university degree'.
	Aditional characteristics	
firsttimer	*id*	Equals one if the respondent participated in the BOP-HH for the first time in Wave 30, and zero otherwise.
ninterest	*qinterest*	Equals one if the respondent found the BOP-HH 'not so interesting' or 'not interesting at all', and zero otherwise.
difficult	*qeasy*	Equals one if the respondent found the BOP-HH 'somewhat difficult' or 'very difficult', and zero otherwise.
toolong	*qlong*	Equals one if the respondent found the BOP-HH 'a little too long' or 'far too long', and zero otherwise.

Notes: This table describes the construction of the variables used in the empirical analysis. In the middle column, we refer to the names of the original variables as listed in the questionnaire for Wave 30 (June 2022) of the BOP-HH.

Table A.2: Treatment assignment and socioeconomic characteristics

	meanshift
age	-0.06
	(0.08)
east	-0.74
	(2.57)
female	2.64
	(2.03)
fullemploy	-2.02
	(2.36)
hhsize	-0.17
	(1.09)
ln(*income*)	0.51
	(2.12)
yoe	-0.35
	(0.60)
Constant	53.91***
	(16.52)
Observations	2,733
$\bar{R}^2$	0.00

Notes: This table presents the estimates from a linear regression of treatment status on socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. Heteroskedasticity-consistent standard errors are reported in parentheses. The reported coefficients and standard errors are the estimated ones times 100. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.3: Inflation expectations and socioeconomic characteristics: mean-shift group

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
age	-0.02***	0.00	0.00***	0.01	-0.01**	-0.00	0.02***	-0.00	-0.00
	(0.01)	(0.04)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
east	-0.10	0.87	0.02	0.25	0.02	0.08	0.06	0.53*	0.11
	(0.17)	(1.07)	(0.02)	(0.25)	(0.14)	(0.08)	(0.19)	(0.28)	(0.20)
female	-0.16	2.19***	0.03**	0.76***	0.16	0.01	-0.15	0.99***	0.43**
	(0.13)	(0.77)	(0.02)	(0.20)	(0.11)	(0.06)	(0.17)	(0.21)	(0.17)
fullemploy	0.13	2.53**	0.01	0.32	0.09	-0.06	-0.01	0.17	0.04
	(0.15)	(1.01)	(0.01)	(0.24)	(0.11)	(0.06)	(0.15)	(0.23)	(0.19)
hhsize	-0.00	0.92**	0.02**	0.19	0.03	-0.07**	0.09	0.35***	0.25**
	(0.07)	(0.43)	(0.01)	(0.12)	(0.05)	(0.03)	(0.07)	(0.11)	(0.12)
ln(*income*)	-0.29*	-2.64***	-0.05***	-0.76***	-0.25*	0.07	-0.50**	-0.93***	-0.46*
	(0.16)	(0.96)	(0.02)	(0.21)	(0.13)	(0.07)	(0.21)	(0.25)	(0.26)
yoe	0.07**	0.03	-0.01**	-0.06	0.01	0.02	0.08	-0.05	-0.02
	(0.04)	(0.23)	(0.00)	(0.06)	(0.03)	(0.02)	(0.05)	(0.06)	(0.05)
Constant	6.18***	19.68***	0.39***	12.87***	4.09***	-0.46	5.32***	15.14***	4.80**
	(1.24)	(6.47)	(0.14)	(1.53)	(1.01)	(0.55)	(1.54)	(2.20)	(2.33)
Observations	1,377	1,377	1,377	1,377	1,377	778	778	1,350	1,350
$\bar{R}^2$	0.03	0.02	0.02	0.02	0.01	0.00	0.02	0.05	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on socioeconomic characteristics. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.4: Differences in inflation expectations: interaction with *age*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.44	-7.60***	-0.02	1.28**	-0.34	0.04	-0.16	0.65	-0.80
	(0.33)	(2.59)	(0.04)	(0.57)	(0.25)	(0.15)	(0.42)	(0.55)	(0.51)
meanshift × *age*	-0.00	0.04	0.00	-0.01	0.00	0.00	0.00	-0.01	0.00
	(0.01)	(0.04)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.03	0.01	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and age. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.5: Differences in inflation expectations: interaction with *east*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.29***	-4.63***	-0.01	0.73***	-0.23***	0.04	-0.07	0.04	-0.63***
	(0.09)	(0.65)	(0.01)	(0.15)	(0.07)	(0.05)	(0.13)	(0.14)	(0.13)
meanshift × *east*	0.05	-3.06	0.04	-0.45	0.17	0.05	-0.01	-0.15	0.21
	(0.22)	(2.04)	(0.03)	(0.36)	(0.19)	(0.12)	(0.39)	(0.38)	(0.29)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.03	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and a dummy variable for East Germans. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.6: Differences in inflation expectations: interaction with *female*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.31***	-3.93***	0.01	0.54***	-0.20***	-0.02	-0.04	-0.03	-0.53***
	(0.10)	(0.68)	(0.01)	(0.16)	(0.07)	(0.05)	(0.15)	(0.15)	(0.13)
meanshift × *female*	-0.03	-3.25**	-0.03	0.31	-0.01	0.20**	-0.10	0.11	-0.17
	(0.17)	(1.43)	(0.02)	(0.30)	(0.15)	(0.09)	(0.25)	(0.28)	(0.25)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.01	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and a dummy variable for female respondents. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.7: Differences in inflation expectations: interaction with *fullemploy*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.18*	-5.71***	-0.00	0.62***	-0.28***	0.10	0.04	0.06	-0.51***
	(0.11)	(0.80)	(0.01)	(0.18)	(0.10)	(0.06)	(0.19)	(0.15)	(0.15)
meanshift × *fullemploy*	0.27*	1.29	0.01	0.07	0.18	-0.11	-0.24	-0.12	-0.19
	(0.16)	(1.31)	(0.02)	(0.28)	(0.13)	(0.09)	(0.24)	(0.27)	(0.23)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and a dummy variable for full-time employed individuals. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.8: Differences in inflation expectations: interaction with *hhsize*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.47**	-4.84***	-0.03	0.64*	-0.17	0.05	-0.03	-0.43	-0.91***
	(0.19)	(1.59)	(0.02)	(0.34)	(0.16)	(0.10)	(0.29)	(0.32)	(0.29)
meanshift × *hhsize*	-0.08	-0.15	0.01	0.01	-0.02	-0.00	-0.02	0.20	0.15
	(0.08)	(0.70)	(0.01)	(0.15)	(0.06)	(0.04)	(0.10)	(0.14)	(0.13)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.01	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and household size. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.9: Differences in inflation expectations: interaction with ln(*income*)

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	2.25	-14.91	0.08	2.72	1.00	0.64	4.11**	-1.08	-1.54
	(1.38)	(11.08)	(0.15)	(2.25)	(1.16)	(0.72)	(2.09)	(2.32)	(2.24)
meanshift × ln(*income*)	-0.24	1.20	-0.01	-0.25	-0.15	-0.07	-0.51**	0.13	0.12
	(0.17)	(1.34)	(0.02)	(0.27)	(0.14)	(0.09)	(0.25)	(0.28)	(0.27)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.02	0.01	0.02	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and household income. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.10: Differences in inflation expectations: interaction with *yoe*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	-0.13	-14.20***	-0.00	0.86	-0.44	0.52*	0.63	-1.24	-1.49*
	(0.59)	(4.91)	(0.08)	(0.99)	(0.52)	(0.31)	(0.88)	(0.95)	(0.84)
meanshift × *yoe*	0.04	0.78*	0.00	-0.02	0.02	-0.04	-0.06	0.11	0.08
	(0.05)	(0.42)	(0.01)	(0.08)	(0.04)	(0.03)	(0.07)	(0.08)	(0.07)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.02	0.01	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, socioeconomic characteristics and an interaction between treatment status and years of education. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.11: Differences in inflation expectations: interested participants only

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.27***	-4.94***	-0.00	0.72***	-0.23***	0.05	-0.05	0.01	-0.63***
	(0.08)	(0.63)	(0.01)	(0.14)	(0.07)	(0.04)	(0.12)	(0.13)	(0.12)
Observations	2,645	2,645	2,645	2,645	2,645	1,425	1,425	2,592	2,592
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.03	0.05	0.02	0.03	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status and socioeconomic characteristics when focusing only on participants who find the BOP-HH interesting. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.12: Differences in inflation expectations: interaction with *difficult*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.30***	-5.00***	-0.00	0.70***	-0.21***	0.08*	-0.13	0.00	-0.65***
	(0.08)	(0.67)	(0.01)	(0.14)	(0.07)	(0.05)	(0.13)	(0.13)	(0.12)
difficult	0.61***	-0.37	0.04	-0.67*	0.50**	0.14*	-0.37*	-0.61*	-0.18
	(0.24)	(1.83)	(0.03)	(0.36)	(0.20)	(0.09)	(0.20)	(0.33)	(0.28)
meanshift × *difficult*	0.06	-2.25	0.04	-0.77	0.13	-0.32**	0.66**	0.05	0.82*
	(0.35)	(1.94)	(0.05)	(0.50)	(0.30)	(0.14)	(0.33)	(0.52)	(0.49)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.03	0.03	0.02	0.01	0.01	0.05	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, a dummy variable that indicates whether respondents found the BOP-HH questionnaire too difficult, an interaction with treatment status and socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.13: Differences in inflation expectations: interaction with *toolong*

	(1) *bins*	(2) *phigh*	(3) *multipeak*	(4) μ_i	(5) σ_i	(6) γ_i	(7) κ_i	(8) $\hat{\pi}_i^E$	(9) $\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.32***	-5.06***	-0.00	0.79***	-0.22***	0.05	-0.11	0.03	-0.66***
	(0.09)	(0.70)	(0.01)	(0.15)	(0.07)	(0.05)	(0.13)	(0.15)	(0.13)
toolong	0.22*	1.59	0.01	0.20	0.10	-0.07	0.16	-0.03	-0.16
	(0.13)	(1.31)	(0.02)	(0.24)	(0.11)	(0.08)	(0.22)	(0.22)	(0.21)
meanshift × *toolong*	-0.05	-0.31	-0.00	-0.63*	0.08	-0.03	0.22	-0.11	0.35
	(0.20)	(1.61)	(0.02)	(0.34)	(0.17)	(0.12)	(0.32)	(0.32)	(0.28)
Observations	2,733	2,733	2,733	2,733	2,733	1,473	1,473	2,678	2,678
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.05	0.02	0.03	0.02	0.01	0.01	0.04	0.02

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status, a dummy variable that indicates whether respondents found the BOP-HH questionnaire too long, an interaction with treatment status and socioeconomic characteristics. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.

Table A.14: Differences in inflation expectations: Wave 29 to Wave 31 participants

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	bins	*phigh*	*multipeak*	μ_i	σ_i	γ_i	κ_i	$\hat{\pi}_i^E$	$\lvert\hat{\pi}_i^E - \mu_i\rvert$
meanshift	0.09	-5.22***	-0.01	0.83***	-0.37***	0.09	0.21	-0.13	-0.97***
	(0.14)	(1.08)	(0.02)	(0.28)	(0.12)	(0.08)	(0.21)	(0.25)	(0.24)
Observations	738	738	738	738	738	405	405	722	722
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{R}^2$	0.04	0.06	0.01	0.04	0.04	0.03	0.02	0.06	0.03

Notes: This table presents the estimates from linear regressions of histogram characteristics and point forecasts on treatment status and socioeconomic characteristics for the subset of individuals that participated in Waves 29, 30 and 31. The baseline group consists of the individuals that were presented with the original bin design. In columns (6) and (7), we consider only the responses of participants who use at least three bins. For columns (8) and (9), we trim $\hat{\pi}_i^E$ by 1% from top and bottom. Heteroskedasticity-consistent standard errors are reported in parentheses. Asterisks '*', '**', and '***' indicate significance at the 10%, 5%, and 1% critical level, respectively.