Direct Elicitation of Parametric Belief Distributions: An application to inflation expectations*

Pedro Gonzalez-Fernandez[†] ⓡ Ciril Bosch-Rosa[‡] ⓡ Thomas Meissner[§]

August 31, 2024

Abstract

We introduce a novel method to elicit belief distributions and apply it to elicit inflation expectations in a representative US sample through a pre-registered survey experiment. Our approach elicits beta belief distributions directly in a two-step process. First, participants specify their minimum and maximum expected inflation. They then use a graphical interface with two sliders to adjust the mean and variance of their inflation belief distribution. We benchmark our method against the "Bins" method, popularized by the New York Fed's Survey of Consumer Expectations (SCE). Our findings reveal significant variations in elicited belief distributions depending on the method used. Specifically, our approach yields higher mean inflation estimates and substantially reduces the standard deviations of the distributions. Respondents report that our method is easier to use, more engaging, and better allows them to express their beliefs. Furthermore, the resulting distributions more accurately reflect participants' beliefs across several dimensions and show stronger correlations with their point predictions.

Keywords Belief Elicitation · Inflation Expectations · Macroeconomic Survey

*Corresponding author: Pedro Gonzalez-Fernandez (pedrofide@gmail.com). The author order was randomized using the AEA author randomization tool. All authors contributed equally to the paper. We would like to thank Paolo Crosetto, Frank Heinemann, Arno Riedl, Stefan Trautmann, and Elias Tsakas for their valuable comments. Bosch-Rosa acknowledges financial support from the Deutsche Forschungsgemeinschaft (DFG) through the CRC TRR 190 Rationality and Competition. Gonzalez-Fernandez acknowledges financial support from the Graduate School of Business and Economics (GSBE) of Maastricht University. This study is pre-registered at https://aspredicted.org/3NT_YSM

[†]Maastricht University

[‡]Technische Universität Berlin

[§]Maastricht University

1 Introduction

Understanding macroeconomic beliefs, particularly inflation expectations, is crucial for policymakers and economic agents. These expectations influence various economic decisions, including consumption, savings, investment, and wage bargaining, which in turn affect macroeconomic outcomes such as inflation or aggregate demand (Bernanke et al., 2007). For central banks, particularly those operating under inflation-targeting regimes, managing inflation expectations is a key tool for maintaining price stability and guiding economic behavior. The importance of inflation expectations is evident in the recent increase in academic research and policy discussions on this topic (Armantier et al., 2021; Coibion et al., 2022; D'Acunto et al., 2023; Stantcheva, 2023).

However, much of the analysis of household inflation expectations still relies on point estimates, which overlook the uncertainty in their beliefs — a crucial factor that influences consumption and investment behaviors (Coibion et al., 2022) and becomes especially important during periods of economic uncertainty, as it can lead to substantial deviations from central bank targets (D'Acunto et al., 2022). Given the importance of uncertainty in household beliefs, a recent strand in the literature underscores the need to elicit full belief distributions for inflation expectations (Armantier et al., 2013, 2015; Hartzmark and Sussman, 2024).

One of the most popular methods for eliciting full belief distributions is the "Bins" method. This approach, based on the work of Manski (2004) and popularized by the New York Fed's Survey of Consumer Expectations (SCE) (Armantier et al., 2013; D'Acunto et al., 2023), asks participants to allocate percentage chances to ten predefined inflation intervals (or bins). While this method allows respondents to express a wide range of beliefs, the responses often require *ex post* fitting of a parametric distribution to facilitate the data analysis.

Engelberg et al. (2009) introduced a method for fitting beta distributions to data collected through the Bins method, which has since been adopted by the New York Fed, the Bundesbank, and the European Central Bank, among others (Manski, 2018). However, fitting respondents' beliefs to a beta distribution *ex post* imposes parametric constraints that respondents are unaware of when expressing their beliefs. This might result in elicited distributions that do not accurately reflect respondents' true beliefs, as they might have preferred a different beta distribution if given the choice.

This paper introduces a new method for eliciting belief distributions that allows respondents to directly select the beta distribution closest to their beliefs, eliminating the need for any *ex post* parametric fitting. Our method follows a two-step procedure: first, participants specify the minimum and maximum expected inflation values; second, they use two interactive sliders to form their belief distribution. The first slider adjusts the mean, while the second controls the variance. This simple design allows respondents to intuitively and dynamically shape their belief distributions, aided by graphical and tabular representations that update in real-time as the sliders are adjusted.

In addition to enabling the direct selection of a beta distribution, a key advantage of our direct method is that it does not require participants to partition their beliefs into predefined bins, thereby avoiding the systematic biases that this approach introduces (Becker et al., 2023; Hartzmark and Sussman, 2024).[1] Our survey tool is programmed in JavaScript, making it easily implementable on survey platforms such as Qualtrics, and it can be made incentive-compatible in laboratory experiments (Gonzalez-Fernandez, 2024) or in contexts where a true realization of the belief distribution is available. Finally, while our interface can be adapted to other parametric functional forms, this paper focuses exclusively on the beta distribution due to its widespread use in eliciting inflation expectations.

To validate our new method, we conducted a large pre-registered online survey eliciting twelve-month-ahead inflation expectations from a U.S. representative sample. All respondents participated in our new *Direct* method and were randomized into either the *Original* version of the Bins method or the *Adaptive* version. The *Original* version, based on Armantier et al. (2013) and employed by the New York Fed's Survey of Consumer Expectations (SCE), utilizes predefined bin intervals centered around zero. In contrast, the new *Adaptive* version adjusts the bin range according to the maximum and minimum values provided by participants, allowing for distributions that may not be centered around zero.

Because the "true" belief distribution of respondents is unobservable, it is challenging to compare the validity of the different belief elicitation methods. To address this issue, we exploited the within-subject design of our survey and developed a series of measures to directly compare the various methods. The first set of measures asked respondents which

[1]See Tversky and Koehler (1994); Benjamin (2019) for a review on partition-dependence bias.

method they participated in allowed them to express their beliefs most accurately, which they found easiest to use, and which was most engaging. The second set of measures asked participants to rate how well several inflation scenarios aligned with their beliefs. These scenarios were tailored to each respondent's elicited belief distributions and included the median inflation, the probability of deflation, and the probability of inflation exceeding 5%. The third set of measures, following Goldfayn-Frank et al. (2024), tested how well the means of the elicited belief distributions correlated with stated point beliefs and how well they predicted future planned consumption.

The results show that the characteristics of the belief distributions critically depend on the elicitation method. Specifically, the new *Direct* method consistently produces belief distributions with larger mean inflation and smaller standard deviations than the *Original* bins method. The *Direct* method also reduces the presence of mean deflationary expectations, which are common in the *Original* method and likely an artifact of the question's format rather than the respondents' true beliefs (Hartzmark and Sussman, 2024; Goldfayn-Frank et al., 2024). Additionally, while the *Adaptive* bins method yields mean inflation beliefs similar to those of the *Direct* method, it produces wider standard deviations.

Crucially, the comparative measures reveal a strong preference among respondents for the *Direct* method over both the *Original* and *Adaptive* bins methods. Respondents not only find the *Direct* method easier to use and more engaging, but they also consider it a significantly better tool for expressing their beliefs. These preferences are corroborated by our finding that the tailored inflation scenarios calculated using the *Direct* method are closer to respondents' beliefs than those calculated using the *Original* method. However, we do not detect any differences between the scenarios calculated using the *Direct* and *Adaptive* methods, suggesting that part of the increased accuracy of the *Direct* method comes from allowing respondents to set the maximum and minimum boundaries of their belief distribution.[2] Furthermore, we find that the means of the inflation expectation distributions in the *Direct* method significantly better predict respondents' point inflation beliefs. Regarding the correlation between inflation expectations and planned consump-

[2]This aligns with the findings of Becker et al. (2023), who show that inflation expectations are highly susceptible to the pre-specified ranges used in the bins method.

tion, we observe considerable heterogeneity both across methods and across different consumption categories.

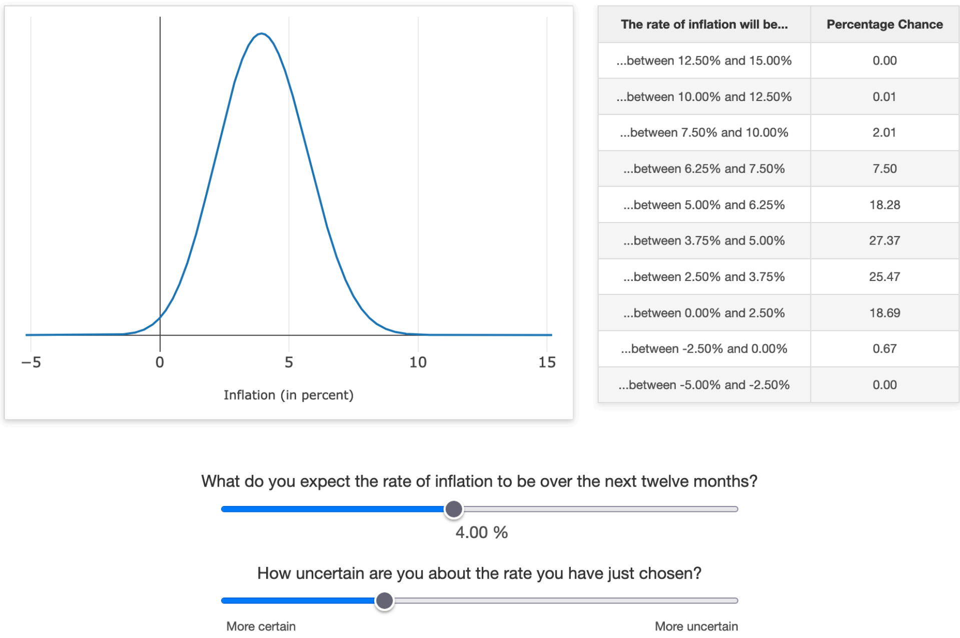
This paper contributes to three strands of the existing literature. First, it adds to the literature on survey design by reaffirming the impact of question framing and design on elicited beliefs, a concern previously highlighted by Coibion et al. (2022, 2023); Becker et al. (2023). In particular, we show that belief distributions can vary substantially within individuals depending on the elicitation method. Second, we contribute to the literature on inflation expectations by introducing the new *Direct* method and comparing it to two versions of the bins method. Despite the popularity of the bins method, recent studies have scrutinized its reliance on predefined bins, which can prime participants toward certain values and increase uncertainty, especially when nudged toward extreme outcomes (Becker et al., 2023; Boctor et al., 2024; Goldfayn-Frank et al., 2024). In contrast, our *Direct* method avoids these limitations and performs better across a range of subjective and objective comparative measures. Third, we contribute to the broader literature on belief distribution elicitation tools. Previous approaches have explored graphical elicitation interfaces that allow respondents to directly construct their belief distributions, such as the early "Distribution Builder" by Sharpe et al. (2000) or the "Click-and-Drag" tool by Crosetto and De Haan (2023). While our approach is more restrictive in terms of functional form compared to these tools, it offers a simple and intuitive way to elicit beliefs, requiring only the manipulation of two sliders. Moreover, our approach is especially advantageous in cases where the goal is to elicit distributions that conform to a specific functional or continuous distribution, such as the beta distribution commonly used in the inflation expectations literature.

The remainder of the paper is organized as follows: Section 2 describes the elicitation methods we test. Section 3 explains the survey design and the measures we use to compare the different elicitation methods. Section 4 discusses our results, and section 5 concludes.

2 Belief distribution elicitation methods

In line with Boctor et al. (2024) and Goldfayn-Frank et al. (2024) we benchmark our new *Direct* method of eliciting beta belief distributions against the current gold standard popularized by the New York Fed's Survey of Consumer Expectations (SCE, Armantier

Figure 1: Screenshot *Direct* elicitation interface



et al., 2017), henceforth the *Original* bins method, and a variation of this method which we call the *Adaptive* bins method.

2.1 The *Direct* elicitation method

The *Direct* method elicits belief distributions through a two-step procedure. In the first step, respondents set the maximum and minimum inflation rates they consider feasible. In the second, they manipulate the mean and variance of the beta distribution using two sliders, as shown in Figure 1. For the mean, we asked respondents for their expected inflation over the next twelve months. For the variance, we asked them to indicate their uncertainty about the rate they had just chosen. We avoided using technical terms like 'mean' and 'variance' in the interface to make the tool accessible to all respondents.[3] The resulting beta distribution was graphically displayed in the interface and dynamically updated with any slider adjustments.[4] In addition to the graph, participants also saw a table that showed what their input implied across ten discrete inflation intervals. These

[3]We carefully explained what distributions are and how to interpret them in the instructions. Screenshots of these can be found in the Appendix Section A.2.

[4]To enhance participants' understanding of the tool, and in line with the recommendations of Engelberg et al. (2009), the shape of the beta distribution was restricted to be unimodal.

intervals mirrored those used in the *Adaptive* bins method introduced below. The table was displayed by default, but participants could choose to hide it at any time.[5]

2.2 The *Original* bins method

The *Original* bins method, as described in Armantier et al. (2013), is the most common approach to measure full inflation expectation distribution by central banks and researchers (Becker et al., 2023). In it, participants were presented with ten bins covering all possible inflation rates (ranging from "-12% or more" to "12% or more") and were asked to assign the probability that inflation over the next 12 months will fall within each bin (see Figure 9 for a screenshot of the implementation in our survey). A beta distribution is then fitted to participants' responses using the method of Engelberg et al. (2009).[6]

2.3 The *Adaptive* bins method

The *Adaptive* bins method is a two-step variation of the *Original* bins method. In the first step, participants specified the maximum and minimum inflation rates they considered feasible. This range was dynamically partitioned into ten bins centered around the mean of the specified range. In the second step, respondents assigned probabilities to each of the ten bins, as in the *Original* method (see Figure 10 for an example). The advantage of the *Adaptive* method is that the bins are closed at the extremes and centered at the midpoint of the specified range. Again, we used the method of Engelberg et al. (2009) to fit a beta distribution to participants' responses.

2.4 Discussion of the methods

The main difference between the *Direct* method and the bins methods is that respondents choose their preferred beta distribution, rather than having it fitted to their beliefs ex-

[5]A minimal working example of the *Direct* eliciation tool implemented in Qualtrics can be found via https://maastrichtuniversity.eu.qualtrics.com/jfe/form/SV_9ZVeYQQZvUjGmJ8.

[6]As the first and last intervals are open, we follow Armantier et al. (2017) by choosing -38% and 38% as bounds of the fitted beta distribution, in case these bins are filled in. Also, because Engelberg et al. (2009) recommend fitting an isosceles triangle distribution when only one or two bins are filled in, we pre-registered as a robustness check to only analyze responses with at least three or more bins filled in. See section B in the Appendix.

post. Moreover, we believe the *Direct* method is simpler to use, as respondents only need to manipulate two sliders instead of inputting different percentage values that must sum to 100 across ten bins. Another important difference is that the *Original* method uses pre-defined intervals centered around zero, which are open at the extremes. In contrast, our *Direct* method allows participants to choose their own minimum and maximum inflation values. To explore the impact of this design feature, we introduce the *Adaptive* method, which uses the same interface as the *Original* method but allows custom minimum and maximum values. We believe this adapted version of the bins method could be especially useful for surveys in high-inflation countries (e.g., Turkey or Argentina) where the pre-defined bounds may not contain the most likely inflation rates. While the *Original* bins method has been adapted by shifting the midpoint to the point estimate of inflation (Goldfayn-Frank et al., 2024), our method offers greater flexibility by adjusting the bins based on the respondents' chosen minimum and maximum. To the best of our knowledge, such an adaptive bins method has never been tested before.

3 Survey

We pre-registered to collect 1,000 observations and received 968 complete responses from a U.S. representative sample using Prolific and Qualtrics. As specified in our pre-registration, we excluded responses where the difference between the maximum and minimum expected inflation was less than 0.5 percentage points, and responses with an inflation point belief below -20 percent or above 50 percent. After these exclusions, we retained a total of 939 observations. As additional robustness checks, we conducted all analyses excluding respondents who i) filled in only one or two bins in the bins method, ii) were in the bottom 2% of time spent on either belief distribution elicitation method,[7] and iii) had inflation point beliefs below 0% or above 15%. For all regressions, we pre-registered additional Huber robust specifications to account for the impact of outliers. All results, unless specified otherwise, remain robust to these checks and can be found in Appendix B.

The median completion time was 10:38 minutes, and participants were paid a flat fee of £1.38, resulting in an average hourly reward of £7.91. Informed consent was obtained from

[7] In the pre-registration, *speeders* were mistakenly defined as those in the top 2% of completion time when it should have been those in the *bottom* 2%. We corrected this for the analysis in this paper.

all participants before they took part in the survey. The study received ethical approval from the Ethics Review Committee Inner City faculties (ERCIC) at Maastricht University, ensuring that all procedures complied with ethical guidelines for research involving human subjects.

3.1 Survey design

The survey consisted of four different blocks.[8] In the first block, we asked respondents about their estimate of the inflation rate over the next 12 months, along with their lowest and highest possible estimates. These responses represent each respondent's point prediction, as well as their minimum and maximum expectations for the inflation rate. In the second block, we elicited the inflation belief distributions. In two treatment arms, respondents either completed the *Direct* and *Original* bins method ($N = 469$) or the *Direct* and *Adaptive* bins method ($N = 470$). Assignment to treatment and order of method within treatment were randomized. In the third block, we validated our method with three types of questions: i) direct comparison questions, where respondents ranked the two methods they participated in along several dimensions, ii) questions about how well their beliefs align with different tailored scenarios based on their elicited beta distributions, and iii) questions about future planned consumption. See Section 3.2 for a detailed description of these three types of questions. In the fourth block, we collected a series of socio-demographic variables such as gender, age, education level, income, and financial literacy (using the "Big 3" questions suggested in Lusardi and Mitchell, 2011).

3.2 Comparative Measures

3.2.1 Direct comparison of methods

In our direct comparison measures, respondents compared the *Direct* method to the bins method, using a Likert scale ranging from 1 (Method A much better) to 7 (Method B much better). Method A always corresponded to the *Direct* method, and Method B corresponded to the bins method the respondent participated in. We asked respondents which method was easier to use (*Ease*), which method is more engaging (*Engagement*),

[8] These blocks are not apparent to respondents. We use this terminology to organize the different data collection sections of the survey.

Table 1: Direct Comparison of Methods

Question: Which of these methods...	
Label	Text of the question
Ease	... was easier to use?
Engagement	.. was more "fun" to use?
Express	... allowed you to better express your expectations?

Note: Participants rated questions on a Likert scale from 1 (Method A much better) to 7 (Method B much better), where method A always corresponds to the *Direct* method and method B always corresponds to either the *Original* or *Adaptive* bins method.

and which method allows them to express their beliefs better (*Express*).[9] Since these direct comparison measures explicitly asked respondents which method is better, we interpret these questions as the key validation measures for our method. A summary of the measures and their exact wording is provided in Table 1. A screenshot of the survey implementation is depicted in Figure 11.

3.2.2 Inflation scenarios

We asked respondents to rate how well three inflation scenarios align with their beliefs using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). For example, respondents rated how well their beliefs align with a probability X that there will be deflation in the next twelve months. Importantly, the value of X is tailored to each respondent and is derived twice, once for each of the respondent's elicited beta distributions. Therefore, respondents answered six different scenarios in total — two on the probability of deflation, two on the chance of inflation exceeding 5% in the next twelve months, and two on the medians of the elicited beta distributions. This setup allows us to compare how well either the directly elicited or fitted beta distributions describe respondents' beliefs. Importantly, and in contrast to the direct comparison measures, this is a "blind" comparison across methods, as respondents were unaware that the values shown across the different scenarios are based on their earlier inflation estimates. Thus,

[9] We conjecture that a more engaging method is more likely to yield higher-quality responses, especially when the elicitation method is not incentivized (as it is usually the case in such surveys).

Table 2: Alignment with inflation scenarios

Question: How much do your inflation expectations align with the following statements:	
Label	Text of the question
Deflation	The chance of deflation (i.e. negative inflation rates) in the next 12 months is $X\%$.
Larger-than-5	The chance of inflation larger than 5% in the next 12 months is $Y\%$.
Median	It is equally likely that inflation over the next 12 months will be above or below $Z\%$.

Note: Participants rated questions on a Likert scale from 1 (strongly disagree) to 7 (strongly agree). X, Y, and Z represent the respective values computed from the participants' responses. All questions were repeated twice: once for the *Direct* method, and once for either the *Original* or *Adaptive* bins method.

these ratings should be unaffected by how engaging or cumbersome the method was to use. See Table 2 for the exact wording of each question and Figure 12 for a screenshot of the implementation.

3.2.3 Planned Consumption

In line with Goldfayn-Frank et al. (2024), we evaluate the external validity of the different belief elicitation methods by examining their ability to predict planned consumption. To collect this data, we used the consumption questions included in the Bundesbank Household Survey on Consumer Expectations, where participants assess whether their consumption across nine different categories will decrease, increase, or remain the same over the next year. These categories include durable goods, essential goods like food and cleaning products, clothing, entertainment, transportation, services, traveling, housing costs, and financial reserves. See Figure 13 for a screenshot and exact wording.

4 Results

In this section, we follow our pre-registered analysis to validate the *Direct* belief elicitation method and compare it to the bins methods. In Section 4.1, we demonstrate that the

Table 3: Summary of mean, standard deviation, and completion time

		Method			p-value		
Measure	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Mean Inflation	mean	5.547	3.751	5.928	< 0.001	0.033	< 0.001
	SD	6.3	4.172	6.42			
SD Inflation	mean	0.775	3.486	1.356	< 0.001	< 0.001	< 0.001
	SD	1.357	3.214	2.86			
Time	mean	52.459	87.424	105.908	< 0.001	< 0.001	< 0.001
	SD	57.446	73.699	113.861			
Observations		939	469	470			

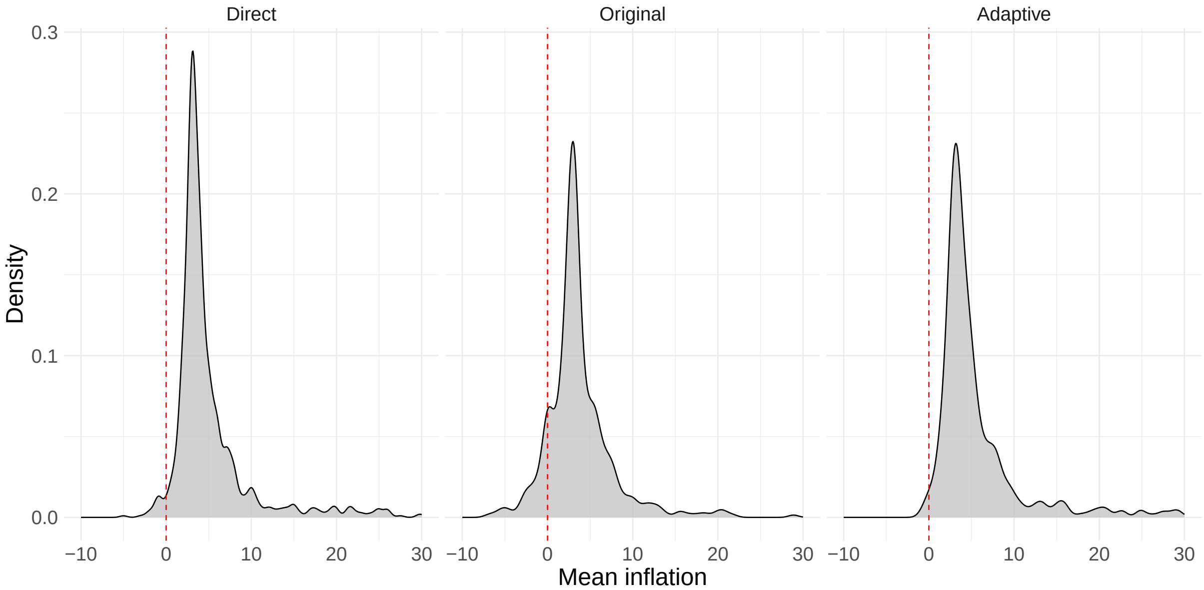
Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

different belief elicitation methods result in substantially different beta distributions. In Section 4.2, we show that in all direct comparisons, respondents prefer the *Direct* method over either of the bins methods. Section 4.3 illustrates that respondents' beliefs align more closely with the beta distribution elicited using the *Direct* method compared to those elicited using the *Original* bins method. Section 4.4 highlights that the mean inflation values derived from the *Direct* method better predict the point inflation beliefs of respondents than either bins method. Lastly, in Section 4.5, we examine the correlation between expected inflation and planned consumption.

4.1 Distribution differences

To compare the distributions across methods, we calculate the mean and standard deviation of the resulting beta distributions (see Table 3). As illustrated in Figure 2 the means from the *Direct* and *Adaptive* methods are larger, on average, than those from the *Original* method ($p < 0.001$, Wilcoxon Signed rank test and Mann-Whintey U test, respectively). Additionally, the *Adaptive* method yields larger means than the *Direct* method ($p = 0.033$, Wilcoxon Signed rank test). Notably, the *Original* method produces a substantial share

Figure 2: Densities of the mean of the elicited beta distributions



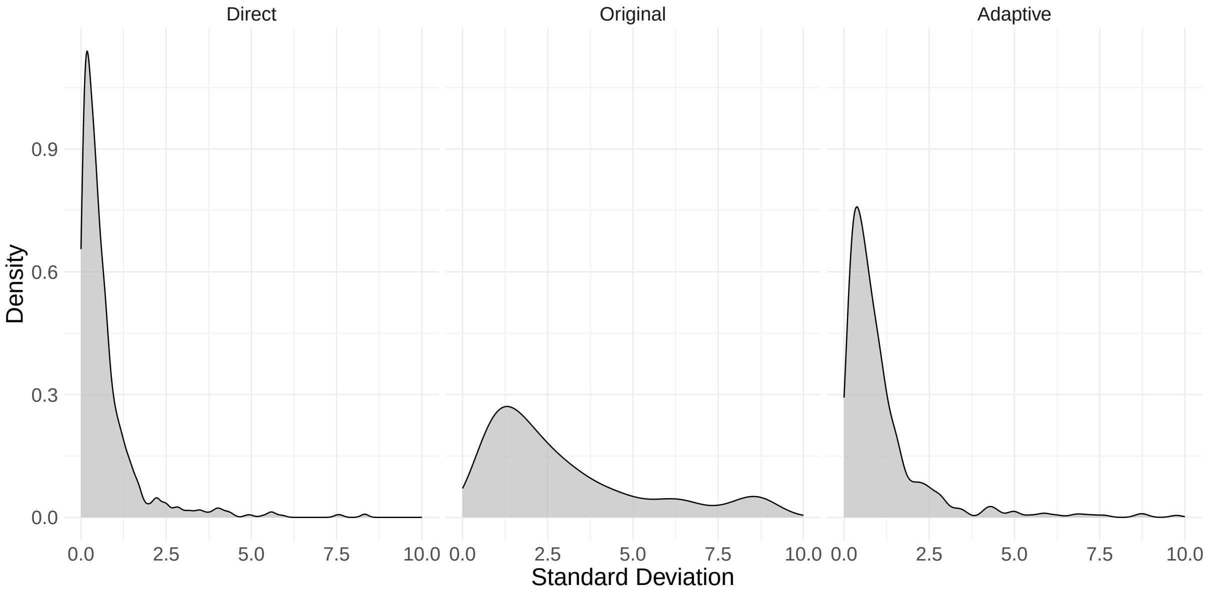
Note: The horizontal axis is limited to the range of -10 to 30 for clarity. For a full view of all values, refer to Figure 18.

of mean inflation estimates below zero, a prediction that appears unrealistic and is likely driven by the elicitation method, as suggested by Becker et al. (2023).

Figure 3 plots the standard deviations of the beta distributions across methods. The *Direct* method yields "tighter" beta distributions compared to the *Adaptive* and *Original* methods ($p < 0.001$, Wilcoxon Signed rank test in both cases). Moreover, the *Adaptive* method, on average, produces beta distributions with lower standard deviations than the *Original* method ($p < 0.001$, Mann-Whitney U test). This suggests that part of the larger standard deviations in the *Original* method may stem from the imposition of *ad hoc* boundaries. Finally, as Table 3 illustrates, the *Direct* method is faster than both bins methods ($p < 0.001$, Wilcoxon Signed rank test in both cases). However, this time does not include the additional 77 seconds the median participant spent reading the extra instructions screen for the *Direct* method.

Summing up, this section illustrates that the choice of the elicitation method matters. Using the *Direct* method yields higher mean inflation and less uncertainty about inflation predictions than both bins methods. We summarize these results as follows:

Figure 3: Densities of the standard deviation of the elicited beta distributions



Note: The horizontal axis is truncated at a value of 10 for clarity. Figure 19 in the appendix shows all values.

Result 1: The *Direct* belief elicitation method results in both a larger mean and a lower standard deviation of the elicited belief distributions compared to the *Original* bins method.

Result 2: The *Direct* belief elicitation method results in a comparable mean but a lower standard deviation of the elicited belief distributions compared to the *Adaptive* bins method.

4.2 Direct comparison of methods

In this section, we compare which method is easier and more engaging to use, and which method allows respondents to better express their beliefs. Because these questions allowed respondents to directly rank their preferred method, we consider these the most relevant results. In these comparisons, each respondent evaluated the *Direct* method against the bins method they used in terms of ease of use (*Ease*), engagement (*Engagement*), and ability to express beliefs (*Express*). In all cases, respondents used a Likert scale from 1 (Direct Method Much Better) to 7 (Bins Method Much Better) to rank the methods.

Table 4: Direct comparison of methods

		Comparison		p-value	
Measure	Statistic	Direct vs Original	Direct vs Adaptive	D–O	D–A
Ease	mean	3.452	2.93	< 0.001	< 0.001
	SD	2.311	2.252		
Express	mean	3.795	3.374	0.062	< 0.001
	SD	2.314	2.279		
Engagement	mean	2.45	2.345	< 0.001	< 0.001
	SD	1.867	1.875		
Observations		469	470		

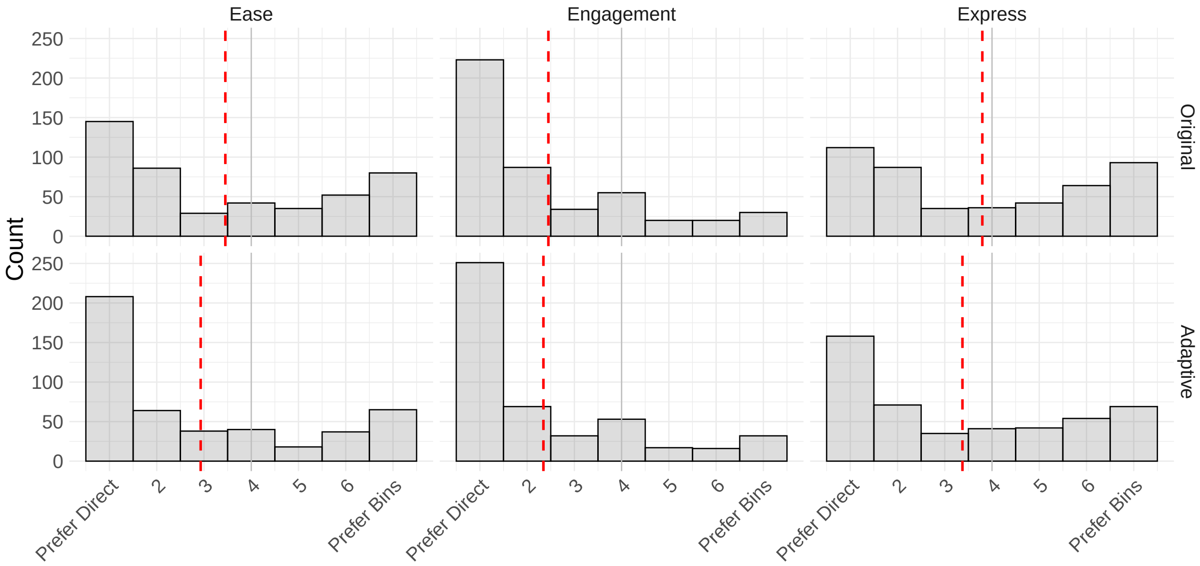
Note: All p-values come from Wilcoxon signed rank tests testing against the null that the value is equal to four.

The results are presented in Figure 4. The first row shows the comparison between the *Direct* and *Original* methods, while the second row compares the *Direct* and *Adaptive* methods. The red dashed line represents the mean value across subjects, and the vertical gray line at four represents the indifference point between the two methods. Across all dimensions and comparisons, the red dashed line falls to the left of four, indicating that participants find the *Direct* method easier to use, more engaging, and better for expressing their beliefs than either bins method. A matched pairs test shows that the values are statistically different from the indifference value of 4, except for the comparison of belief expression between the *Original* and *Direct* methods, where $p = 0.062$.

Overall, these results suggest that respondents prefer the *Direct* method over either of the two bins methods. On average, respondents find the *Direct* method easier to use, more engaging, and more effective for expressing their beliefs. The latter result is surprising given that the bins methods do not impose a functional form on subjects' belief distributions *ex ante*, and thus, in principle, allow for more flexibility when responding to the question.[10] The results of this section can be summarized as follows:

[10] Researchers typically fit beta distributions to the responses *ex post*, but respondents are not aware of this.

Figure 4: Histograms comparing ease of use, engagement, and ability to express beliefs



Result 3: The *Direct* belief elicitation method is easier to use, more engaging, and allows respondents to better express their beliefs than the *Original* bins method.

Result 4: The *Direct* method is easier to use and more engaging than the *Adaptive* bins method.

Result 5: Neither the *Original* nor the *Adaptive* bins method allows respondents to express their beliefs better than the *Direct* method.

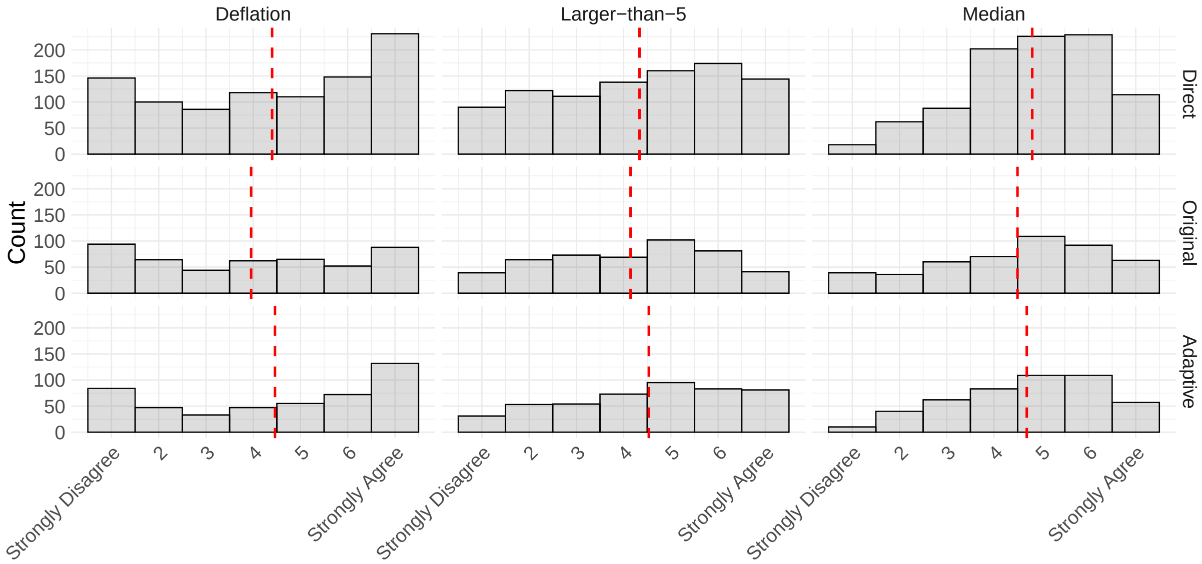
4.3 Alignment with inflation scenarios

To capture how well respondents' beliefs aligned with their elicited distributions, we presented them with three tailored scenarios derived from the elicited beta distributions. For each scenario, respondents indicated how much they agreed using a Likert scale from 1 (Strongly disagree) to 7 (Strongly agree). The scenarios included the computed probability of deflation (*Deflation*), the computed median of the distribution (*Median*), and the probability that inflation exceeds 5% (*Larger-than-5*).[11]

The results for each scenario are shown in Figure 5. In all cases, the average reported value for the *Direct* method is larger than that for the *Original* method. The differences

[11] See Table 2 for the exact wording of the scenarios.

Figure 5: Histograms comparing alignment with inflation scenarios



Note: Each column of the figure measures agreement with the statements i) ii) and iii) of table 2 for each method. Note that the *Direct* method has twice as many observations as the *Original* and *Adaptive*.

are significant for *Deflation* and *Median* ($p < 0.001$, Wilcoxon signed rank test), but not for *Larger-than-5* ($p = 0.761$), indicating that the beta distributions elicited by the *Direct* method align more closely with respondents' beliefs than those elicited by the *Original* method in two out of three comparisons. We find no statistical differences between the *Direct* and *Adaptive* methods. However, when comparing the two bins methods for the *Deflation* and *Larger-than-5* questions, participants are statistically more favorable to the *Adaptive* method than the *Original* one. These findings are summarized in Table 5 and the following results:

Result 6: Participants' beliefs align better with the *Direct* method compared to the *Original* bins method in the *Deflation* and *Median* scenarios.

Result 7: There is no significant difference in the alignment of beliefs across the different scenarios between the *Direct* and *Adaptive* methods.

Result 8: Participants' beliefs align better with the *Adaptive* method than with the *Original* method in the *Deflation* and *Larger-than-5* scenarios.

Table 5: Alignment with inflation scenarios

		Method			p-value		
Scenario	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Deflation	mean	4.399	3.955	4.46	< 0.001	0.643	< 0.001
	SD	2.174	2.173	2.266			
Larger-than-5	mean	4.335	4.147	4.534	0.761	0.152	< 0.001
	SD	1.919	1.769	1.828			
Median	mean	4.809	4.497	4.694	< 0.001	0.446	0.191
	SD	1.479	1.78	1.562			
Observations		939	469	470			

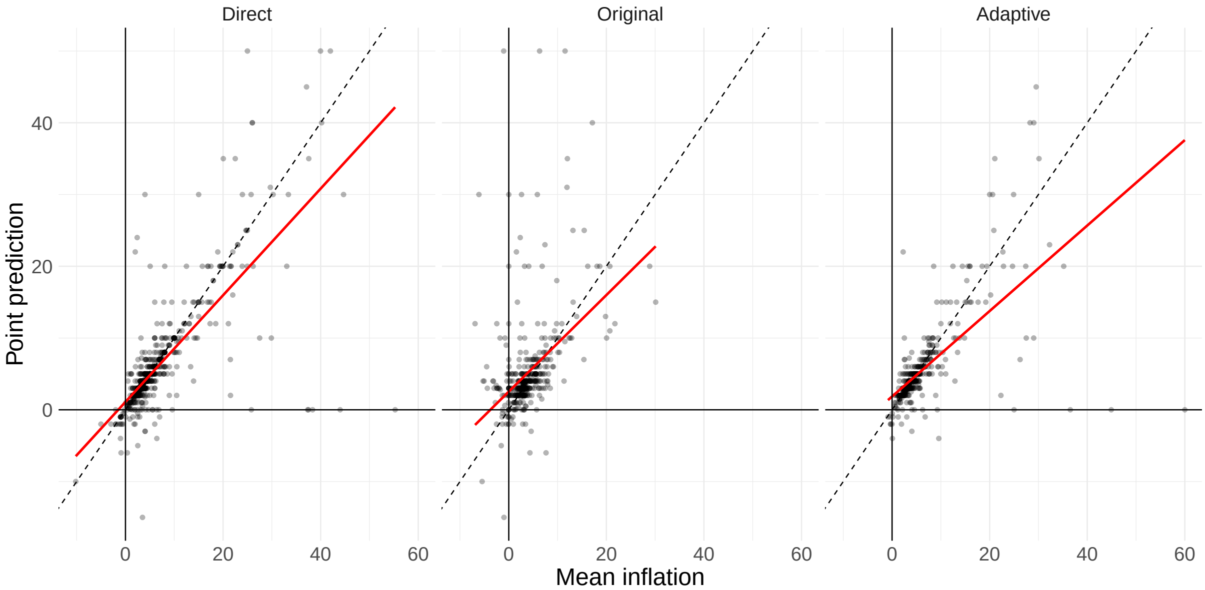
Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

4.4 Correlation with point beliefs

In this section, we follow Goldfayn-Frank et al. (2024) and correlate the means of the elicited beta distributions with the point predictions. Table 6 shows the results of OLS regressions, while Figure 6 provides a scatter plot for each method. There is a stronger correlation between the mean of the *Direct* method and participants' point predictions, as indicated by a larger coefficient, a higher R^2, and a fitted line closer to the 45-degree line. This result is robust to all the pre-registered robustness checks (see Appendix B.5 for details).

Result 9: The mean inflation expectation elicited using the *Direct* method better predicts point inflation beliefs compared to the *Original* and *Adaptive* bins methods.

Figure 6: Scatter Plot of Inflation Point Prediction and Beta Distribution Mean



Note: The dotted line represents the 45-degree line, while the red line is a fitted line based on the data.

Table 6: OLS Regression Results for Point Inflation

	Direct	Original	Adaptive
	(1)	(2)	(3)
mean inflation	0.743***	0.673***	0.596***
	(0.021)	(0.066)	(0.032)
Constant	1.097***	2.535***	1.849***
	(0.179)	(0.371)	(0.279)
Observations	939	469	470
R^2	0.564	0.181	0.426
Note:	*p<0.1; **p<0.05; ***p<0.01		

4.5 Inflation expectations and planned consumption

The correlation between inflation expectations and consumption is central to macroeconomic theory and policy design. However, the empirical evidence for this correlation is mixed. For example, Bachmann et al. (2015) find a negative correlation between durable consumption and inflation expectations, while Coibion et al. (2023) provide causal evidence of a positive effect of inflation expectations on such consumption. Additionally, Dräger and Nghiem (2021) find a positive correlation between inflation expectations and the consumption of a wide variety of goods.

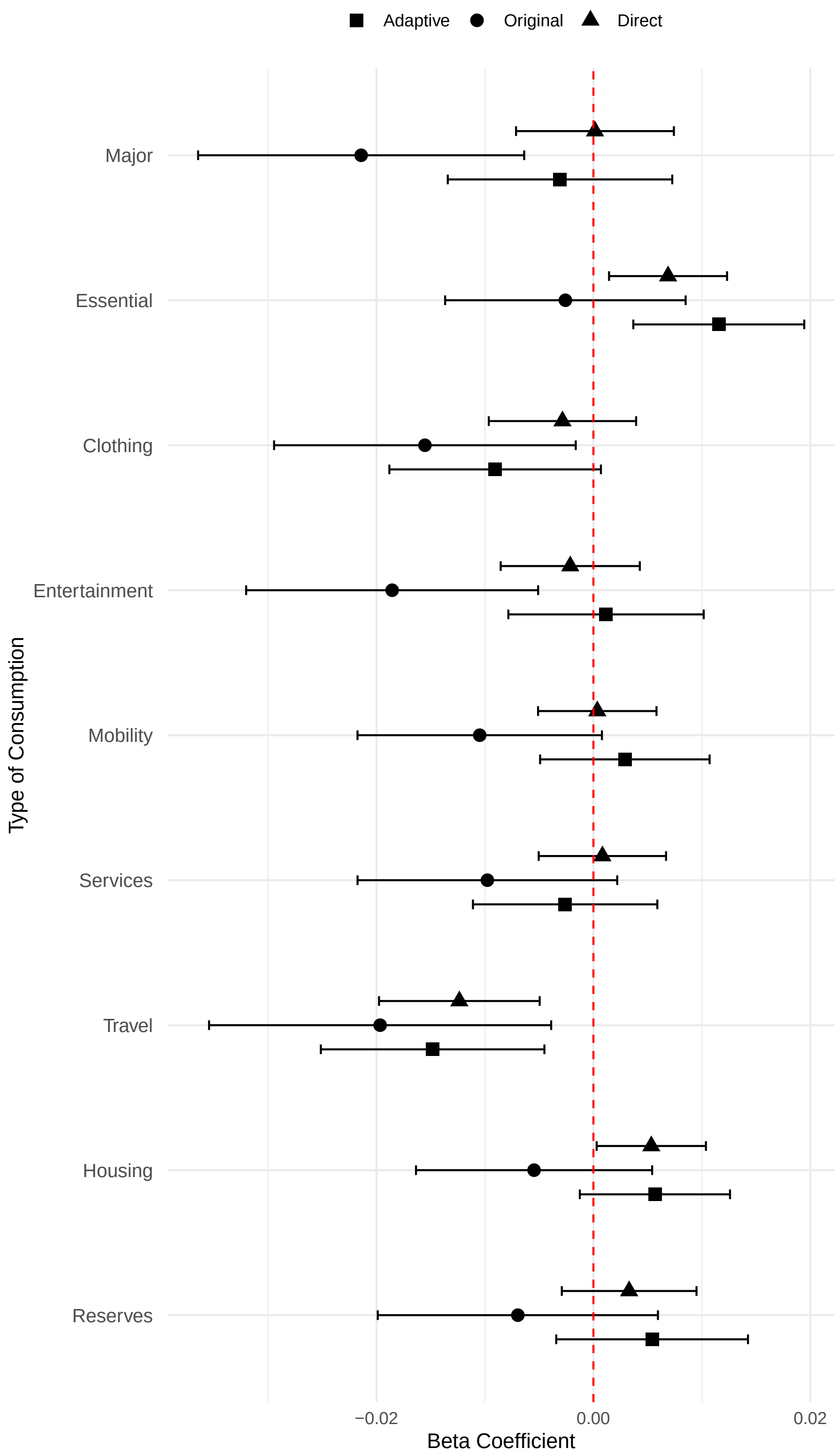
We analyze the correlation between respondents' self-reported planned consumption and the inflation expectations elicited through each method. Specifically, we ask respondents about their planned changes in consumption over the next 12 months for durable goods, essential goods (e.g., food and cleaning products), clothing, entertainment, transportation, services, traveling, housing costs, and financial reserves. Using OLS, we regress the mean inflation expectation of each elicitation method on planned consumption, coding planned increases as 1, decreases as -1, and constant consumption as 0.

Figure 7 illustrates the estimated coefficients of mean expected inflation on planned consumption using a forest plot.[12] Overall, the results are mixed, showing both positive and negative correlations depending on the elicitation method and type of consumption. Inflation means elicited with the *Original* bins method produce exclusively negative correlations, which are statistically different from zero for the categories *Major*, *Clothing*, *Entertainment*, and *Travel*. In contrast, the *Direct* and *Adaptive* methods produce positive correlations that are significantly different from zero for the categories *Essential* and (weakly) for *Housing*. The only consumption category where all methods agree on a significant negative correlation is *Travel*. Finally, the coefficients resulting from the *Original* bins method have consistently larger error bars, suggesting worse predictive power than either the *Direct* or *Adaptive* elicitation methods.

In summary, the *Original* method tends to favor negative correlations and predicts consumption plans more noisily, while the *Direct* and *Adaptive* methods predict consumption plans more precisely but often yield statistically insignificant results. While most results remain robust after applying the pre-registered robustness checks, restricting the sample to participants with inflation point predictions in the range $(0, 0.15)$ eliminates most of the significantly negative correlations between the *Original* method's means and consumption plans (see Figure 25 in the Appendix for more details). This suggests that the negative correlations are largely driven by participants with relatively extreme inflation expectations. We interpret this as warranting caution when interpreting correlations

[12]Ordered logit regressions produce very similar results; see Appendix B.5.

Figure 7: Inflation expectations and planned consumption



between consumption and elicited inflation expectations using the *Original* bins method, as it may generate spurious negative correlations.

Result 10: The sign of the correlation between expected inflation and planned consumption varies substantially across elicitation methods and consumption categories.

5 Discussion and conclusion

We introduce a new method for eliciting belief distributions from survey respondents. Unlike existing methods, our tool allows respondents to directly choose their preferred belief distribution using two sliders. This approach eliminates the need for *ex post* fitting of beliefs and avoids the systematic biases common in current methods (Becker et al., 2023).

To validate our method, we conducted a survey using a representative sample of the U.S. population. The within-subject design of our survey allows us to compare the new *Direct* method to the current standard in the literature (the *Original* method) and a variation of it (the *Adaptive* method). Our findings show that the elicitation format has a considerable impact on the resulting belief distributions, highlighting the importance of using adequate elicitation methods for measuring beliefs. Comparing our new *Direct* elicitation method with the *Original* method, we demonstrate that our method results in both larger means and smaller standard deviations of the belief distributions, on average. Notably, the occurrence of distribution means suggesting deflation is markedly reduced compared to the *Original* bins method, potentially explaining the overall larger mean inflation beliefs in our method.

Respondents rated our new *Direct* method as more engaging, easier to use, and better to express their beliefs than the alternative methods. We also show that the beliefs resulting from the *Direct* method correlate more closely with point beliefs of respondents but do not find any clear pattern between inflation expectations and planned consumption across the different methods and categories of consumption.

There are two main takeaways from our results. First, we contribute to recent evidence showing that expectations are highly dependent on the elicitation format (see, e.g., Becker et al., 2023; Boctor et al., 2024; Goldfayn-Frank et al., 2024). The different methods we test not only yield different estimates for the mean and standard deviation of respondents' belief distributions, but also result in vastly different correlations between inflation and planned consumption across consumption categories. This suggests that the choice of method could lead to spurious correlations between inflation and consumption. Second, and most importantly, we show that our new *Direct* method compares favorably to the current gold standard in the literature. Not only do the elicited results of our method align more closely with respondents' beliefs, but they also rate the *Direct* method as better for expressing their beliefs, easier to use, and more engaging. This is particularly important for online surveys, where time and respondent attention are crucial determinants of data quality.

While our results are promising, we acknowledge certain limitations of the *Direct* method. One potential drawback is that the *Direct* method requires specifying the functional form of beliefs in advance, which may limit respondents' ability to fully express their true belief distribution. However, this approach allows participants to choose the distribution that is closest to their beliefs, rather than having one fitted *ex post*, as is typically done with current methods. Another limitation is that, although our method has the fastest response time, it requires additional instructions, making the overall process more time-intensive. It is important to note, however, that we did not prioritize time efficiency, so future research could further refine the method.

In light of our findings, we believe that central banks and researchers currently using the bins method would benefit from adopting our new *Direct* method. This recommendation is particularly relevant in cases where a distribution is fitted to participants' responses *ex post*. Some advantages of our new method could be achieved by adopting the *Adaptive* method, a variation of the current *Original* elicitation method, where respondents set the maximum and minimum values of their belief distribution. This approach could be a promising alternative for surveys in regions where inflation far exceeds the fixed bins of the *Original* method (e.g., Argentina or Turkey) and where other considerations prevent the use of the *Direct* method.

References

Armantier, O., W. Bruine de Bruin, S. Potter, G. Topa, W. Van Der Klaauw, and B. Zafar (2013): "Measuring inflation expectations," *Annu. Rev. Econ.*, 5, 273–301. Cited on pages 1, 2, and 6.

Armantier, O., W. Bruine de Bruin, G. Topa, W. Van Der Klaauw, and B. Zafar (2015): "Inflation expectations and behavior: Do survey respondents act on their beliefs?" *International Economic Review*, 56, 505–536. Cited on page 1.

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. Cited on pages 1 and 27.

Armantier, O., G. Topa, W. Van der Klaauw, and B. Zafar (2017): "An overview of the survey of consumer expectations," *Economic Policy Review*, 51–72. Cited on pages 4 and 6.

Bachmann, R., T. O. Berg, and E. R. Sims (2015): "Inflation expectations and readiness to spend: Cross-sectional evidence," *American Economic Journal: Economic Policy*, 7, 1–35. Cited on page 18.

Becker, C. K., P. Duersch, and T. Eife (2023): "Measuring inflation expectations: How the response scale shapes density forecasts," *Available at SSRN 4323706*. Cited on pages 2, 3, 4, 6, 12, and 21.

Ben-David, I., E. Fermand, C. M. Kuhnen, and G. Li (2018): "Expectations uncertainty and household economic behavior," Tech. rep., National Bureau of Economic Research. Cited on page 27.

Benjamin, D. J. (2019): "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2, 69–186. Cited on page 2.

Bernanke, B. S. et al. (2007): "Inflation expectations and inflation forecasting," in *Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts*, vol. 10, 11. Cited on page 1.

Boctor, V., O. Coibion, Y. Gorodnichenko, and M. Weber (2024): "On Eliciting Subjective Probability Distributions of Expectations," *Kilts Center at Chicago Booth Marketing Data Center Paper No. Forthcoming.* Cited on pages 4, 21, and 27.

Burke, M. A. and A. Ozdagli (2023): "Household inflation expectations and consumer spending: evidence from panel data," *Review of Economics and Statistics*, 105, 948–961. Cited on page 27.

Caldwell, S., S. Nelson, and D. Waldinger (2023): "Tax refund uncertainty: evidence and welfare implications," *American Economic Journal: Applied Economics*, 15, 352–376. Cited on page 27.

Coibion, O., D. Georgarakos, Y. Gorodnichenko, and M. Van Rooij (2023): "How does consumption respond to news about inflation? Field evidence from a randomized control trial," *American Economic Journal: Macroeconomics*, 15, 109–152. Cited on pages 4, 19, and 27.

Coibion, O., Y. Gorodnichenko, and M. Weber (2022): "Monetary policy communications and their effects on household inflation expectations," *Journal of Political Economy*, 130, 1537–1584. Cited on pages 1, 4, and 27.

Crosetto, P. and T. De Haan (2023): "Comparing input interfaces to elicit belief distributions," *Judgment and Decision Making*, 18, e27. Cited on page 4.

Crump, R. K., S. Eusepi, A. Tambalotti, and G. Topa (2022): "Subjective intertemporal substitution," *Journal of Monetary Economics*, 126, 118–133. Cited on page 27.

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. Cited on pages 1 and 27.

Dräger, L. and G. Nghiem (2021): "Are Consumers' Spending Decisions in Line with A Euler Equation?" *Review of Economics and Statistics*, 103, 580–596. Cited on page 19.

D'Acunto, F., D. Hoang, and M. Weber (2022): "Managing households' expectations with unconventional policies," *The Review of Financial Studies*, 35, 1597–1642. Cited on page 1.

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. Cited on pages 1, 5, and 6.

Goldfayn-Frank, O., P. Kieren, and S. T. Trautmann (2024): "A choice-based approach to the measurement of inflation expectations," Tech. rep., AWI Discussion Paper Series. Cited on pages 3, 4, 7, 10, 17, and 21.

Gonzalez-Fernandez, P. (2024): "Belief Bias Identification," *arXiv preprint arXiv:2404.09297*. Cited on page 2.

Hartzmark, S. M. and A. B. Sussman (2024): "Eliciting Expectations," *Available at SSRN 4780506*. Cited on pages 1, 2, and 3.

Hoffmann, M., E. Moench, L. Pavlova, and G. Schultefrankenfeld (2023): "Forceful or persistent: Wow the ECB's new inflation target affects households' inflation expectations," . Cited on page 27.

Kim, G. and C. Binder (2023): "Learning-through-survey in inflation expectations," *American Economic Journal: Macroeconomics*, 15, 254–278. Cited on page 27.

Lusardi, A. and O. S. Mitchell (2011): "Financial literacy around the world: an overview," *Journal of pension economics & finance*, 10, 497–508. Cited on page 8.

Manski, C. F. (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376. Cited on page 1.

——— (2018): "Survey measurement of probabilistic macroeconomic expectations: progress and promise," *NBER Macroeconomics Annual*, 32, 411–471. Cited on page 1.

Reiche, L. (2023): "That's what she said: an empirical investigation on the gender gap in inflation expectations," . Cited on page 27.

Sharpe, W. F., D. G. Goldstein, and P. W. Blythe (2000): "The distribution builder: A tool for inferring investor preferences," *preprint*. Cited on page 4.

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. Cited on page 1.

Tversky, A. and D. J. Koehler (1994): "Support theory: A nonextensional representation of subjective probability." *Psychological review*, 101, 547. Cited on page 2.

Appendix

A Extra Graphs and Tables

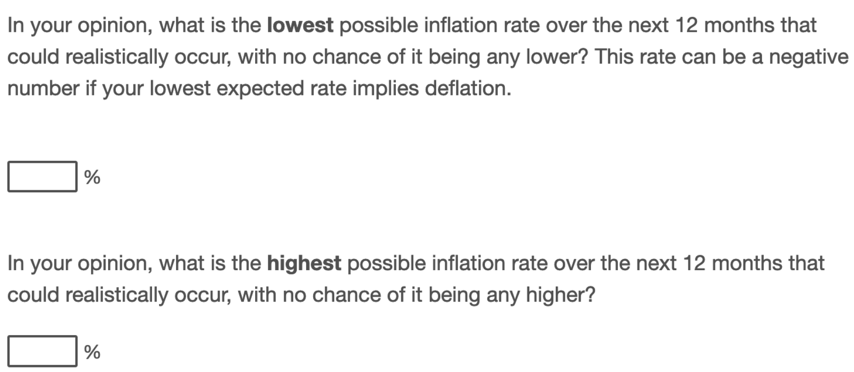
A.1 Literature

Table 7: Papers that fit beta distributions to macroeconomic expectations data

reference	journal	belief data
Armantier et al. (2021)	JEBO	SCE (NY Fed)
Ben-David et al. (2018)	WP	SCE (NY Fed)
Boctor et al. (2024)	WP	Nielsen Homescan Panel
Burke and Ozdagli (2023)	ReStat	SCE (NY Fed)
Caldwell et al. (2023)	AEJ:Applied	own survey + admin data
Coibion et al. (2023)	AEJ:Macro	CentER Internet panel
Coibion et al. (2022)	JPE	Nielsen & SCE (NY Fed)
Crump et al. (2022)	JME	SCE (NY Fed)
D'Acunto et al. (2023)	Handbook Ec. Exp.	SCE (NY Fed)
Hoffmann et al. (2023)	WP	BOP-HH (Bundesbank)
Kim and Binder (2023)	AEJ:Macro	SCE (NY Fed)
Reiche (2023)	WP	BOP-HH & SCE

A.2 Experiment screenshots

Figure 8: Screenshot minimum and maximum inflation

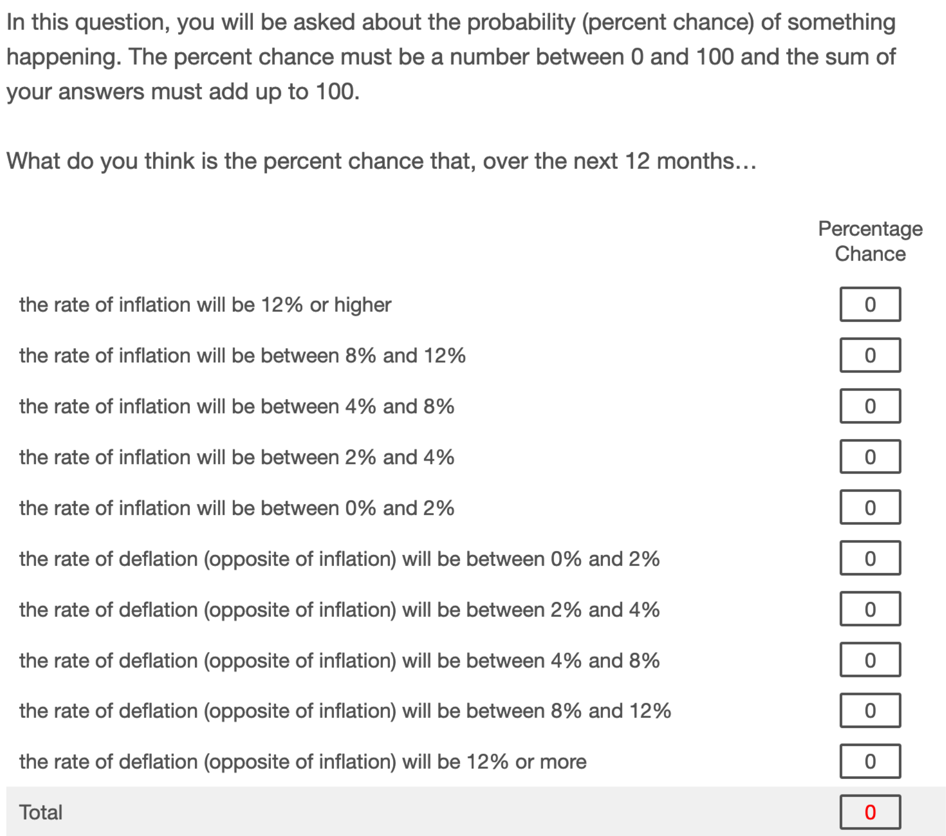

In your opinion, what is the **lowest** possible inflation rate over the next 12 months that could realistically occur, with no chance of it being any lower? This rate can be a negative number if your lowest expected rate implies deflation.

%

In your opinion, what is the **highest** possible inflation rate over the next 12 months that could realistically occur, with no chance of it being any higher?

%

Figure 9: Screenshot *Original* bins method

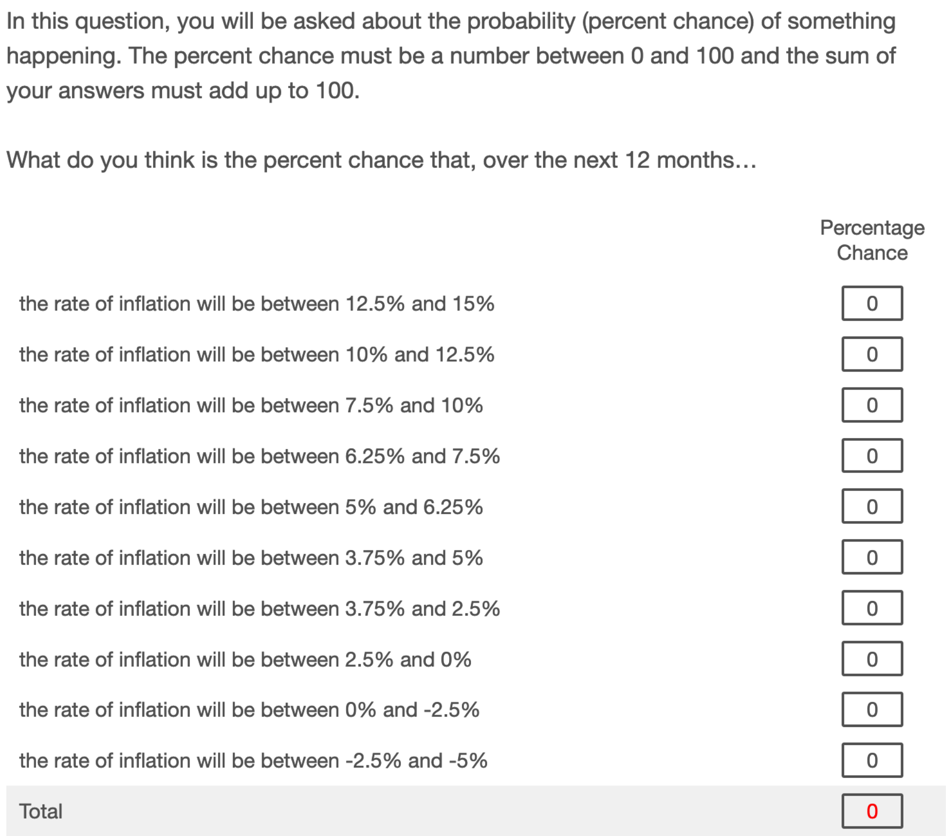


In this question, you will be asked about the probability (percent chance) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, over the next 12 months…

	Percentage Chance
the rate of inflation will be 12% or higher	0
the rate of inflation will be between 8% and 12%	0
the rate of inflation will be between 4% and 8%	0
the rate of inflation will be between 2% and 4%	0
the rate of inflation will be between 0% and 2%	0
the rate of deflation (opposite of inflation) will be between 0% and 2%	0
the rate of deflation (opposite of inflation) will be between 2% and 4%	0
the rate of deflation (opposite of inflation) will be between 4% and 8%	0
the rate of deflation (opposite of inflation) will be between 8% and 12%	0
the rate of deflation (opposite of inflation) will be 12% or more	0
Total	0

Figure 10: Screenshot *Adaptive* bins method



In this question, you will be asked about the probability (percent chance) of something happening. The percent chance must be a number between 0 and 100 and the sum of your answers must add up to 100.

What do you think is the percent chance that, over the next 12 months…

	Percentage Chance
the rate of inflation will be between 12.5% and 15%	0
the rate of inflation will be between 10% and 12.5%	0
the rate of inflation will be between 7.5% and 10%	0
the rate of inflation will be between 6.25% and 7.5%	0
the rate of inflation will be between 5% and 6.25%	0
the rate of inflation will be between 3.75% and 5%	0
the rate of inflation will be between 3.75% and 2.5%	0
the rate of inflation will be between 2.5% and 0%	0
the rate of inflation will be between 0% and -2.5%	0
the rate of inflation will be between -2.5% and -5%	0
Total	0

Note: *Adaptive* bins displayed for an exemplary participant who chose minimum $= -5\%$ and maximum $= 15\%$ inflation

Figure 11: Screenshot direct comparison questions

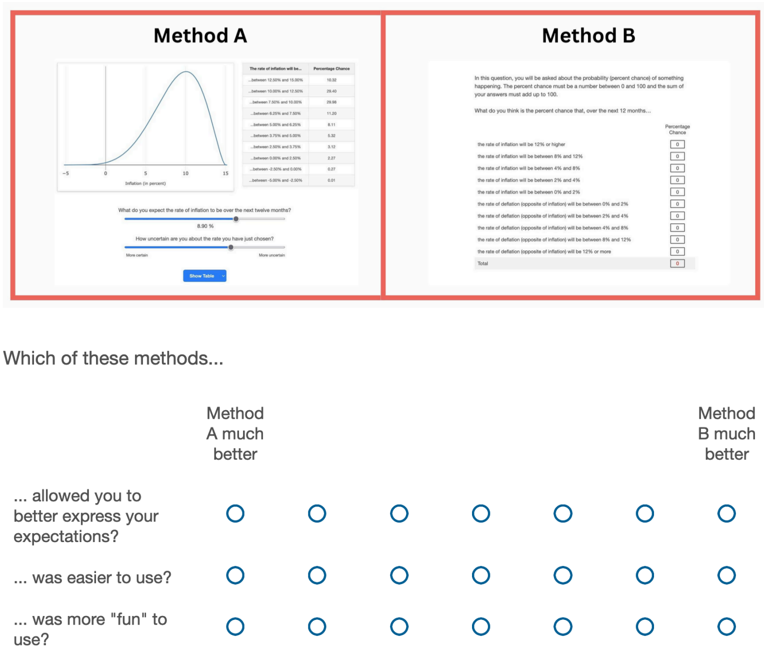
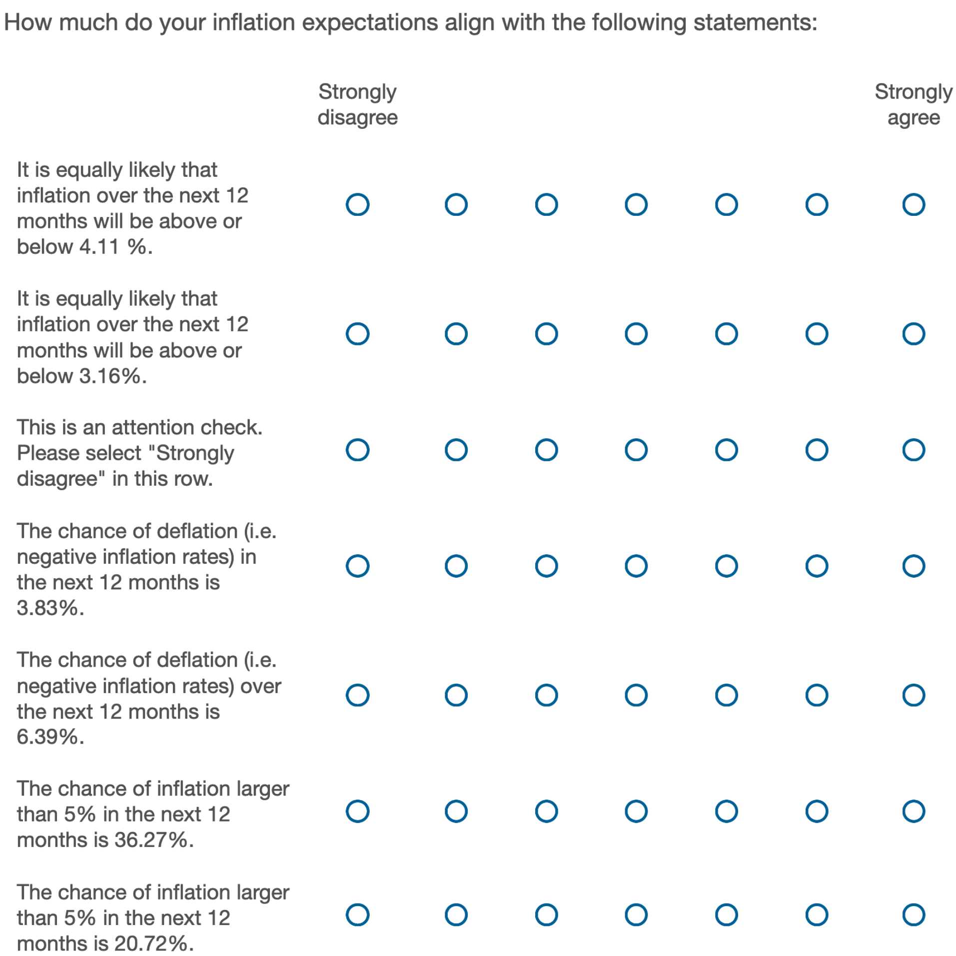


Figure 12: Screenshot inflation scenario questions



How much do your inflation expectations align with the following statements:

	Strongly disagree						Strongly agree
It is equally likely that inflation over the next 12 months will be above or below 4.11 %.							
It is equally likely that inflation over the next 12 months will be above or below 3.16%.							
This is an attention check. Please select "Strongly disagree" in this row.							
The chance of deflation (i.e. negative inflation rates) in the next 12 months is 3.83%.							
The chance of deflation (i.e. negative inflation rates) over the next 12 months is 6.39%.							
The chance of inflation larger than 5% in the next 12 months is 36.27%.							
The chance of inflation larger than 5% in the next 12 months is 20.72%.							

Figure 13: Screenshot planned consumption

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

	I plan to spend less	I plan to spend about the same	I plan to spend more
Major Purchases (e.g. car, furniture, electrical appliances, etc.)	○	○	○
Essential goods (e.g. food and beverages, non-food items such as cleaning products or similar)	○	○	○
Clothing & footwear	○	○	○
Entertainment & recreation (e.g. restaurant visits, cultural events, gym)	○	○	○
Mobility (e.g. fuel, car loans and running costs, bus and train tickets)	○	○	○
Services (e.g. hairdresser, childcare, medical costs)	○	○	○
Travel & holidays	○	○	○
Housing costs (e.g. rent, mortgage, ancillary costs)	○	○	○
Financial reserves	○	○	○

Figure 14: Screenshot *Direct* instructions - 1

Instructions for indicating your inflation expectations

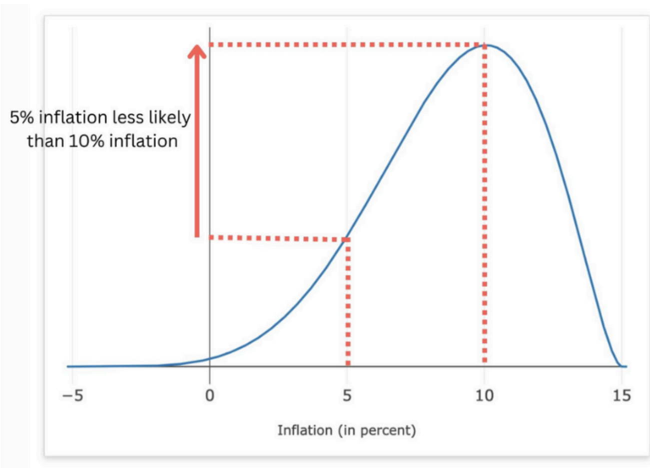
In the next part, we have developed a simple graphical tool that allows you to express your beliefs about future inflation as a probability distribution. Don't worry if you are not familiar with distributions—this guide will explain everything you need to know.

What is a distribution?

A distribution is a way to express the likelihoods of different outcomes. For example, you predicted earlier that inflation over the next 12 moths will be 5%. With what probability will that happen? What about 5% + 2%? And 5% – 2%? Because it is likely that you think that there are values other than 5% that could potentially happen, our tool will allow you to easily assign a probability to all events you believe are feasible by adjusting two sliders. Below we will explain how.

Understanding the distribution graph

The main part of the tool is a graphical representation of a probability distribution. You can see one example in the graph below.



Possible inflation values are represented on the horizontal axis. The blue curve illustrates the likelihood of each inflation rate. The higher the blue curve, the more likely it is that the corresponding inflation rate will occur. For instance, in the graph above, an inflation of 5% is less likely than 10% as the blue curve is higher at 10%. In fact, 10% is the most likely inflation rate, as no other point on the blue curve is higher than at 10%. For inflation values below -5% or above 15%, the curve is at zero, indicating that these inflation rates are impossible.

Figure 15: Screenshot *Direct* instructions - 2

How to use our tool to express your beliefs

To express the distribution that best expresses your own beliefs about inflation, you can use our tool that provides two sliders to modify the distribution graph.

Here is how that works:

1. **Initial graph**: When you start, you will see a bell-shaped curve on the graph. This is the starting probability distribution.
2. **Adjust the first slider**: the first slider indicates your average expected inflation rate. Move the slider left or right to set your expected rate. The peak of the curve will move accordingly. **Example:** If you believe inflation will be around 3% on average, you would move this slider to 3%.

3. **Adjust the second slider:** The second slider adjusts how how certain you are about the value you picked using the first slider. A wider curve indicates more uncertainty, while a narrower curve indicates more confidence. **Example:** If you are very certain that inflation will be 3% and think that it is very unlikely that inflation will be below 2% or above 4%, then the curve should be tight around 3%. You can do this by moving the slider to the left. On the other hand, if you think inflation might be around 3%, but you are fairly uncertain about it, you should move the slider to the right. As you do this, the curve widens, increasing the probability of values further away from 3%.

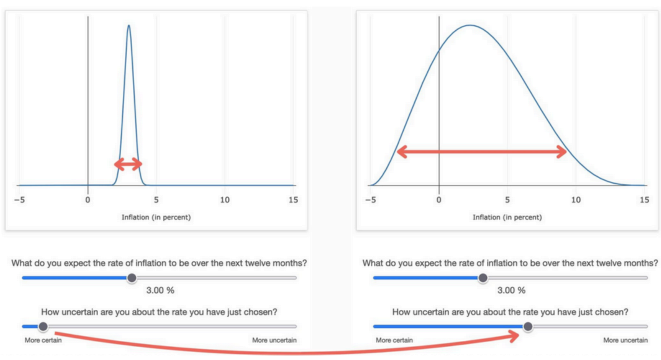
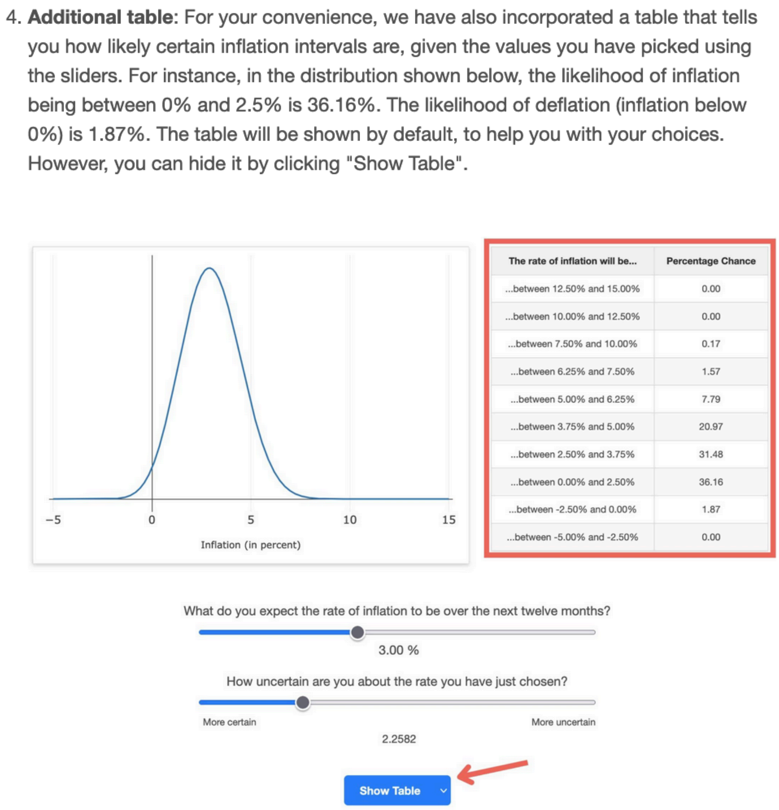


Figure 16: Screenshot *Direct* instructions - 3



4. **Additional table**: For your convenience, we have also incorporated a table that tells you how likely certain inflation intervals are, given the values you have picked using the sliders. For instance, in the distribution shown below, the likelihood of inflation being between 0% and 2.5% is 36.16%. The likelihood of deflation (inflation below 0%) is 1.87%. The table will be shown by default, to help you with your choices. However, you can hide it by clicking "Show Table".

The rate of inflation will be...	Percentage Chance
...between 12.50% and 15.00%	0.00
...between 10.00% and 12.50%	0.00
...between 7.50% and 10.00%	0.17
...between 6.25% and 7.50%	1.57
...between 5.00% and 6.25%	7.79
...between 3.75% and 5.00%	20.97
...between 2.50% and 3.75%	31.48
...between 0.00% and 2.50%	36.16
...between -2.50% and 0.00%	1.87
...between -5.00% and -2.50%	0.00

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

3.00 %

How uncertain are you about the rate you have just chosen?

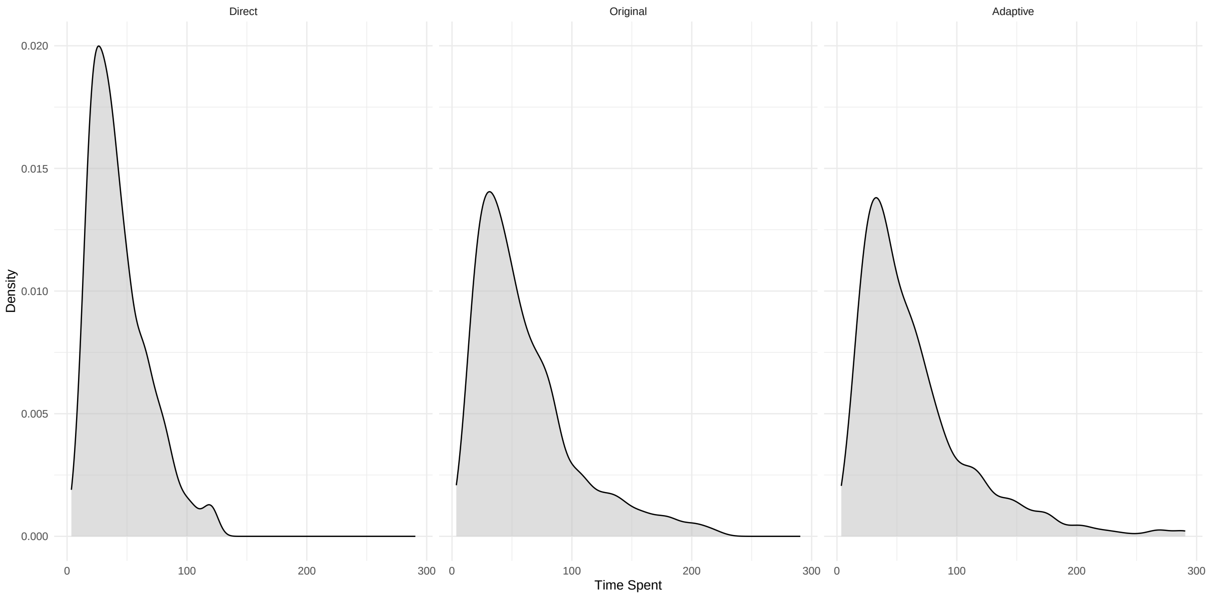
More certain More uncertain

2.2582

Show Table

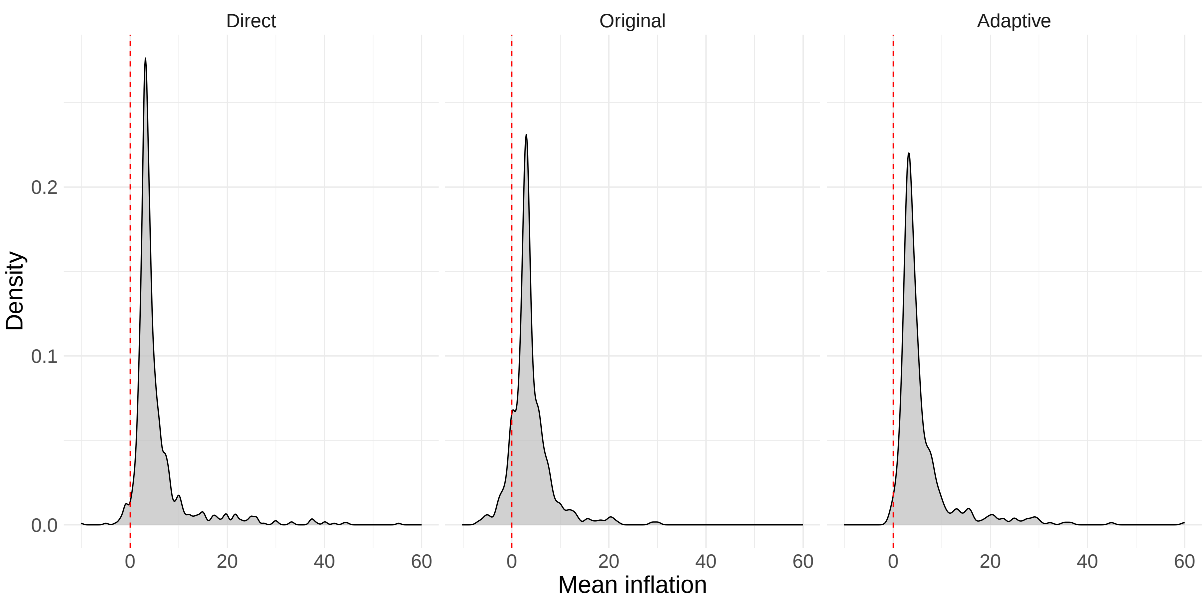
A.3 Mean inflation, standard deviation and time

Figure 17: Density of the Time Spent for each of the three elicitation methods



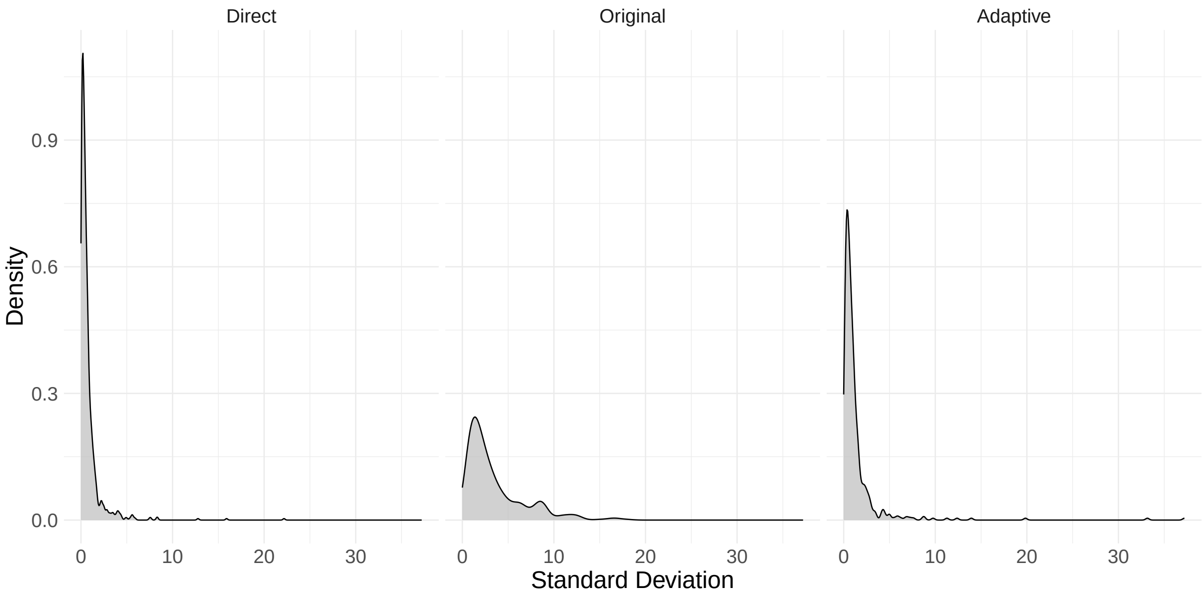
Note: Time spent, measured in seconds for each treatment. This graph contains 95% of the observations, densities including outliers can be found in Figure 20 in the appendix reproduces all values

Figure 18: Density of the means for each of the Beta distribution for each of the three elicitation methods



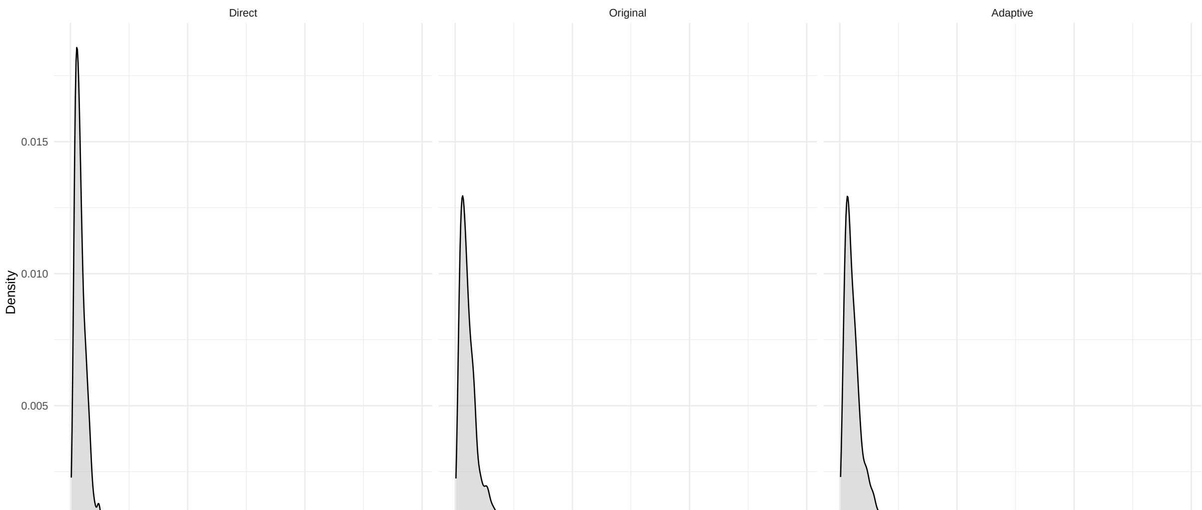
Note: Mean of the Beta distribution for each elicitation method.

Figure 19: Density of the Standard Deviation of the Beta distribution for each of the three elicitation methods



Note: Standard Deviation of the Beta distribution for each elicitation method.

Figure 20: Density of the Time for each of the three elicitation methods



Note: Time spent in seconds for each elicitation method, not truncated.

Table 8: Summary statistics - no speeders

		Method			p-value		
Measure	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Mean Inflation	mean	5.485	3.774	5.767	< 0.001	0.034	< 0.001
	SD	6.182	4.202	5.883			
SD Inflation	mean	0.753	3.498	1.332	< 0.001	< 0.001	< 0.001
	SD	1.242	3.217	2.757			
Time	mean	53.026	89.018	107.832	< 0.001	< 0.001	< 0.001
	SD	57.884	73.693	114.334			
Observations		919	459	460			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

B Robustness Checks

This section shows our results after applying our pre-registered robustness checks namely: i) Excluding those subjects who only fill in one or two bins in any of the two Bins methods (*more than two bins*), ii) Excluding respondents who are in the bottom 2% of the time used by all respondents in both elicitation tasks (*no speeders*), and iii) Keeping only subjects whose point beliefs for inflation are within 0% and 15% (*reasonable*).

B.1 Summary of mean, standard deviation and completion time

Table 9: Summary statistics - reasonable

		Method			p-value		
Measure	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Mean Inflation	mean	4.366	3.586	4.557	< 0.001	0.053	< 0.001
	SD	2.895	3.333	3.197			
SD Inflation	mean	0.619	3.071	0.97	< 0.001	< 0.001	< 0.001
	SD	0.817	2.823	1.856			
Time	mean	52.521	87.251	105.206	< 0.001	< 0.001	< 0.001
	SD	59.169	76.89	115.459			
Observations		816	407	409			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

Table 10: Summary statistics - more than two bins

		Method			p-value		
Measure	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Mean Inflation	mean	5.557	3.751	5.951	< 0.001	0.037	< 0.001
	SD	6.315	4.172	6.449			
SD Inflation	mean	0.776	3.486	1.361	< 0.001	< 0.001	< 0.001
	SD	1.36	3.214	2.875			
Time	mean	52.479	87.424	106.297	< 0.001	< 0.001	< 0.001
	SD	57.502	73.699	114.21			
Observations		934	469	465			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

Table 11: Direct comparison of methods - no speeders

		Comparison		p-value	
Measure	Statistic	Direct vs Original	Direct vs Adaptive	D–O	D–A
Ease	mean	3.438	2.913	< 0.001	< 0.001
	SD	2.308	2.247		
Express	mean	3.776	3.37	0.042	< 0.001
	SD	2.308	2.281		
Engagement	mean	2.42	2.322	< 0.001	< 0.001
	SD	1.849	1.859		
Observations		459	460		

Note: All p-values come from Wilcoxon signed rank tests testing against the null that the value is equal to four.

Table 12: Direct comparison of methods - reasonable

		Comparison		p-value	
Measure	Statistic	Direct vs Original	Direct vs Adaptive	D–O	D–A
Ease	mean	3.396	2.856	< 0.001	< 0.001
	SD	2.289	2.231		
Express	mean	3.796	3.308	0.09	< 0.001
	SD	2.315	2.248		
Engagement	mean	2.405	2.252	< 0.001	< 0.001
	SD	1.809	1.813		
Observations		407	409		

Note: All p-values come from Wilcoxon signed rank tests testing against the null that the value is equal to four.

B.2 Direct comparison of methods

Table 13: Direct comparison of methods - more than two bins

		Comparison		p-value	
Measure	Statistic	Direct vs Original	Direct vs Adaptive	D–O	D–A
Ease	mean	3.452	2.923	< 0.001	< 0.001
	SD	2.311	2.249		
Express	mean	3.795	3.378	0.062	< 0.001
	SD	2.314	2.281		
Engagement	mean	2.45	2.346	< 0.001	< 0.001
	SD	1.867	1.878		
Observations		469	465		

Note: All p-values come from Wilcoxon signed rank tests testing against the null that the value is equal to four.

Table 14: Alignment with inflation scenarios - no speeders

		Method			p-value		
Scenario	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Deflation	mean	4.402	3.943	4.465	< 0.001	0.632	< 0.001
	SD	2.177	2.168	2.273			
Larger-than-5	mean	4.339	4.15	4.535	0.801	0.199	< 0.001
	SD	1.921	1.764	1.832			
Median	mean	4.812	4.486	4.685	< 0.001	0.382	0.178
	SD	1.47	1.765	1.563			
Observations		919	459	460			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

B.3 Alignment with inflation scenarios

Table 15: Alignment with inflation scenarios - reasonable

		Method			p-value		
Scenario	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Deflation	mean	4.453	3.983	4.421	< 0.001	0.251	0.005
	SD	2.179	2.189	2.298			
Larger-than-5	mean	4.261	4.047	4.484	0.572	0.115	< 0.001
	SD	1.91	1.728	1.82			
Median	mean	4.82	4.428	4.694	< 0.001	0.386	0.065
	SD	1.461	1.773	1.55			
Observations		816	407	409			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

Table 16: Alignment with inflation scenarios - more than two bins

		Method			p-value		
Scenario	Statistic	Direct	Original	Adaptive	D–O	D–A	O–A
Deflation	mean	4.409	3.955	4.482	< 0.001	0.672	< 0.001
	SD	2.171	2.173	2.26			
Larger-than-5	mean	4.339	4.147	4.538	0.761	0.175	< 0.001
	SD	1.919	1.769	1.833			
Median	mean	4.816	4.497	4.703	< 0.001	0.427	0.163
	SD	1.477	1.78	1.563			
Observations		934	469	465			

Note: The first two p-values, comparing Direct to both Original and Adaptive methods, come from Wilcoxon signed rank tests. The last p-value, comparing Original and Adaptive methods, comes from a Mann-Whitney U test.

Table 17: OLS Regression Results for Point Inflation - no speeders

	Direct	Original	Adaptive
	(1)	(2)	(3)
mean inflation	0.790***	0.671***	0.734***
	(0.021)	(0.067)	(0.032)
Constant	0.911***	2.581***	1.144***
	(0.174)	(0.379)	(0.263)
Observations	919	459	460
R^2	0.604	0.180	0.536
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 18: OLS Regression Results for Point Inflation - more than two bins

	Direct	Original	Adaptive
	(1)	(2)	(3)
mean inflation	0.743***	0.673***	0.595***
	(0.021)	(0.066)	(0.032)
Constant	1.103***	2.535***	1.865***
	(0.180)	(0.371)	(0.281)
Observations	934	469	465
R^2	0.564	0.181	0.426
Note:	*p<0.1; **p<0.05; ***p<0.01		

B.4 Correlation with point beliefs

Table 19: OLS Regression Results for Point Inflation - reasonable

	Direct	Original	Adaptive
	(1)	(2)	(3)
mean inflation	0.622***	0.384***	0.458***
	(0.017)	(0.031)	(0.024)
Constant	1.499***	2.806***	2.161***
	(0.091)	(0.153)	(0.136)
Observations	816	407	409
R^2	0.610	0.272	0.464
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 20: OLS Regression Results for Point Inflation - robust

	Direct	Original	Adaptive
	(1)	(2)	(3)
Mean Inflation	0.743***	0.673***	0.596***
	(0.092)	(0.102)	(0.142)
Constant	1.097***	2.535***	1.849***
	(0.414)	(0.446)	(0.695)
Observations	939	469	470
R^2	0.564	0.181	0.426
Note:	*p<0.1; **p<0.05; ***p<0.01		

B.5 Inflation expectations and planned consumption

Figure 21: Inflation expectations and planned consumption - ordered logit

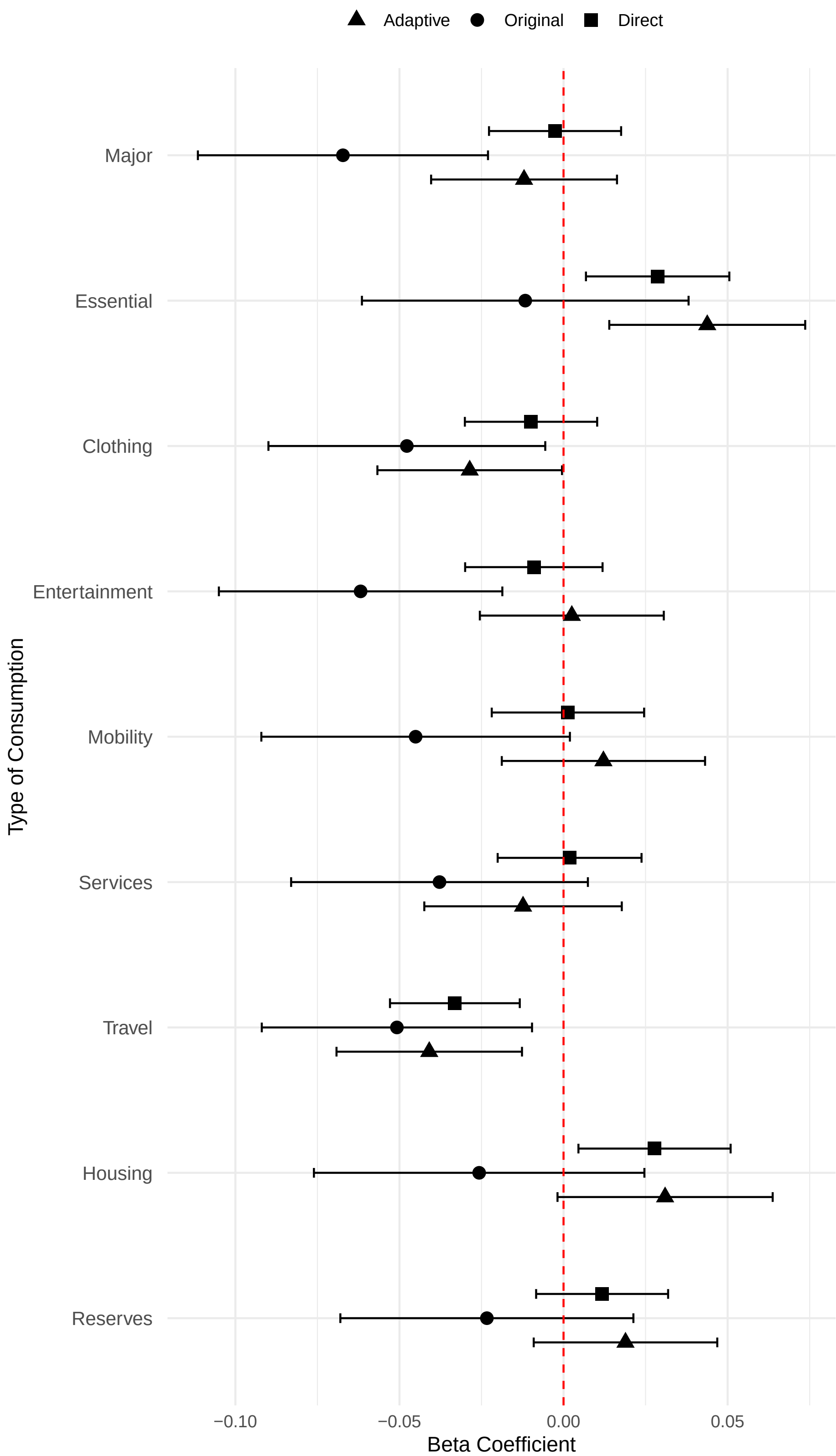


Figure 22: Inflation expectations and planned consumption - robust standard errors

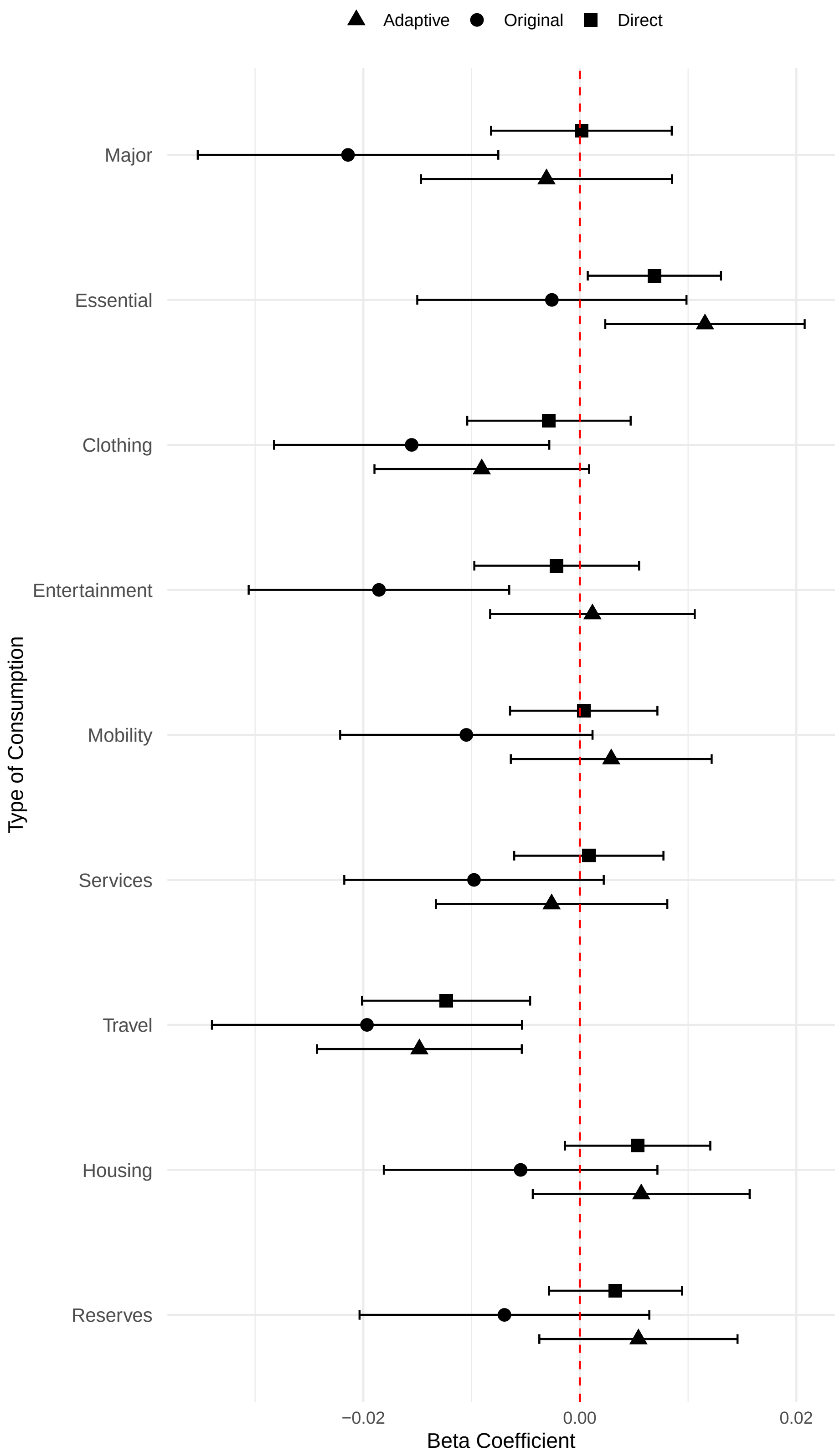


Figure 23: Inflation expectations and planned consumption - no speeders

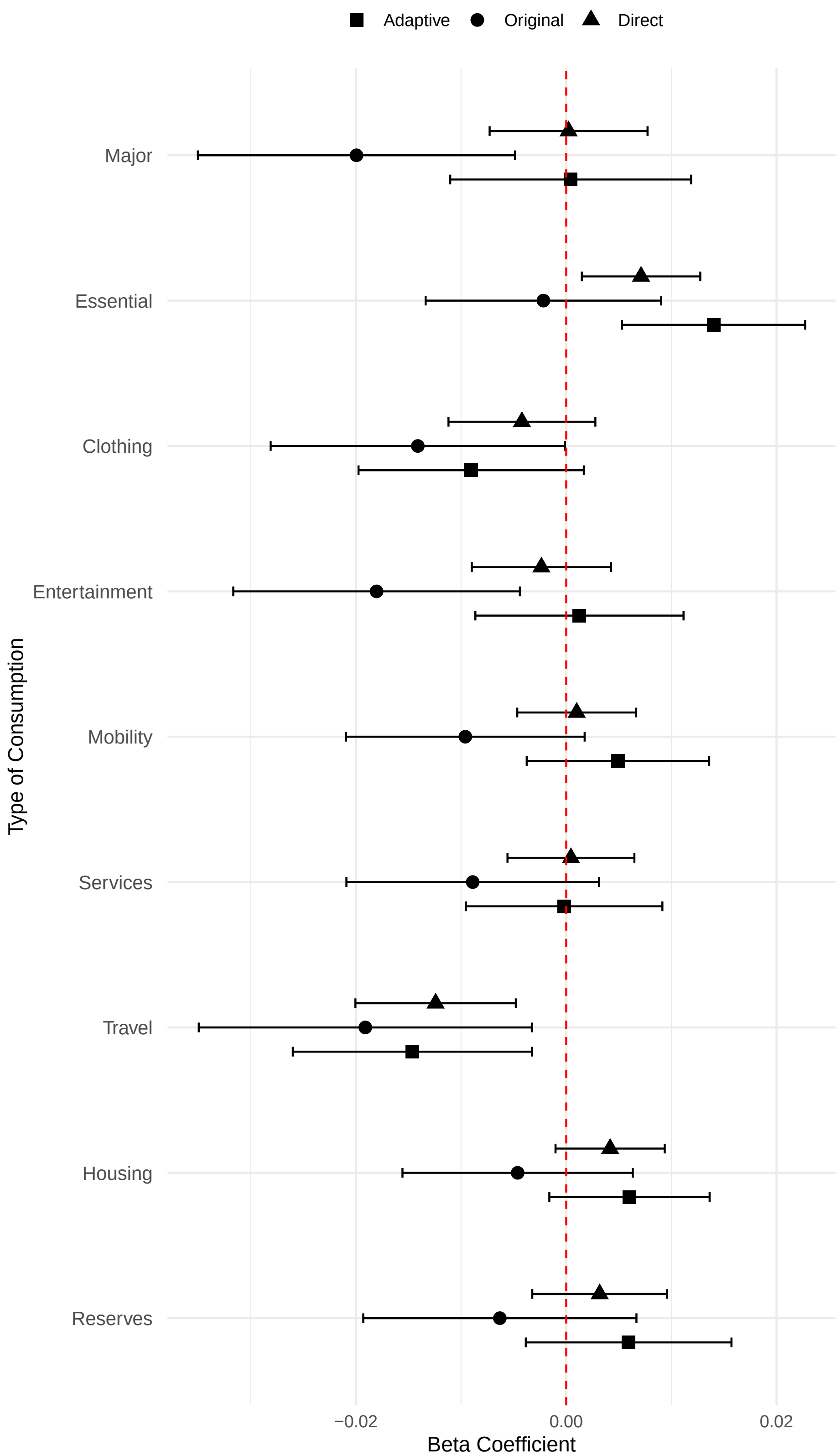


Figure 24: Inflation expectations and planned consumption - more than two bins.

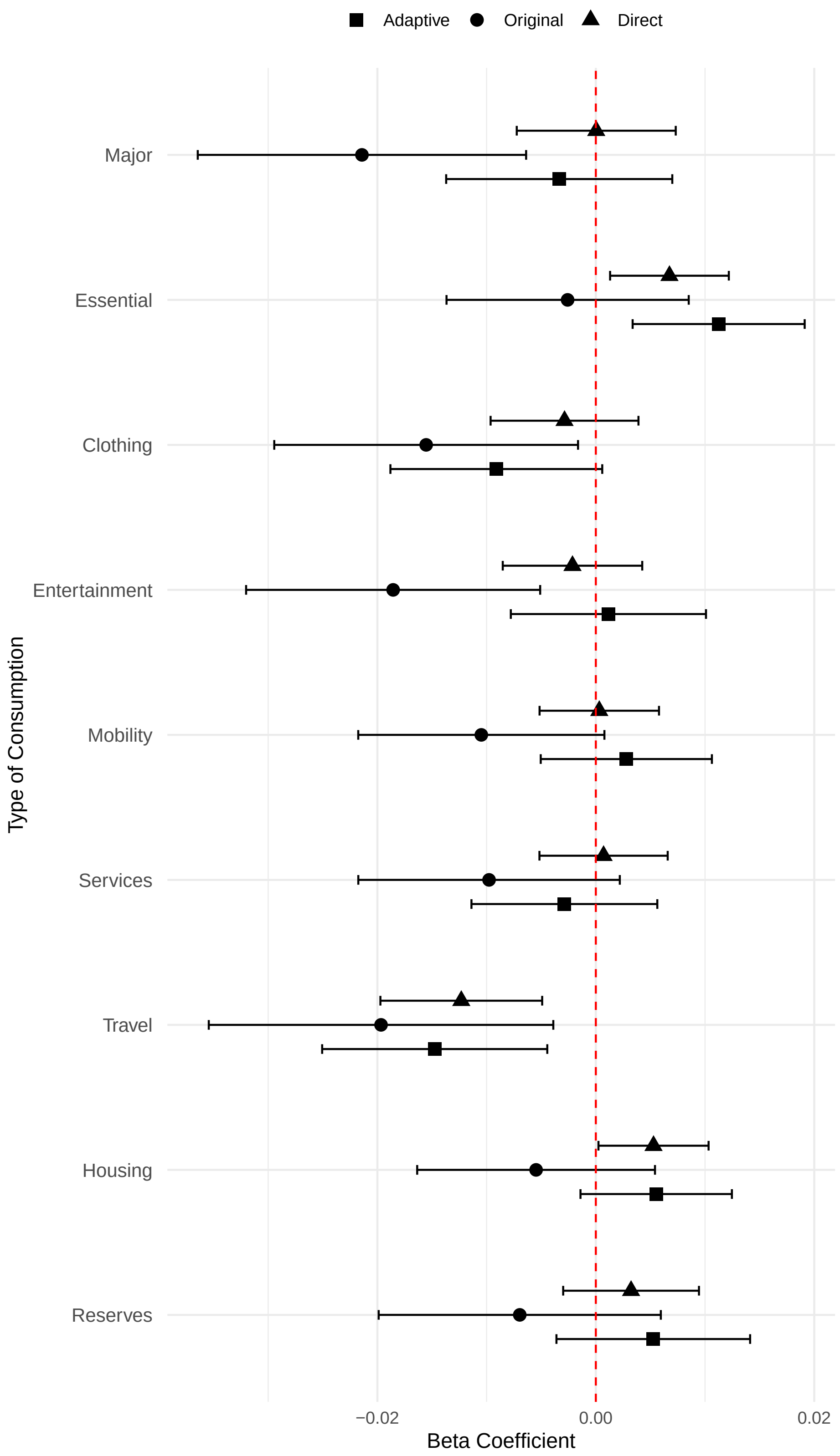


Figure 25: Inflation expectations and planned consumption - reasonable