WORKING PAPER · NO. 2023-143

Greater Than the Sum of Its Parts: Aggregate vs. Aggregated Inflation Expectations

Alexander Dietrich, Edward S. Knotek II, Kristian O. Myrseth, Robert W. Rich, Raphael Schoenle, and Michael Weber

NOVEMBER 2023

Becker Friedman Institute
FOR ECONOMICS AT UCHICAGO

5757 S. University Ave.
Chicago, IL 60637
Main: 773.702.5599
bfi.uchicago.edu

GREATER THAN THE SUM OF ITS PARTS: AGGREGATE VS. AGGREGATED INFLATION EXPECTATIONS

Alexander Dietrich
Edward S. Knotek II
Kristian O. Myrseth
Robert W. Rich
Raphael Schoenle
Michael Weber

Dietrich gratefully acknowledges support from the Deutscher Akademischer Austauschdienst (DAAD, German Academic Exchange Service). We thank Hassan Afrouzi, Philippe Andrade, Nicola Gennaioli, Mark Freeman, Charles Manski, Andrei Shleifer, Wilbert van der Klaauw, Gabriel Zuellig, as well as seminar and conference participants at the BAFA FMI NAG, Bundesbank, Bank of England, Bank of Finland, Bank of Japan 2023 IMES, Beijing University, Bocconi University, CEBRA Annual Meeting, EEA-ESEM Annual Congress, ICEA, NBER Summer Institute, SBS Summer Conference, SEA, Swiss National Bank, Tokyo University, UCLA Anderson and Verein für Socialpolitik for useful comments and discussions. Dietrich gratefully acknowledges support from the Deutscher Akademischer Austauschdienst (DAAD, German Academic Exchange Service). The views stated in this paper are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland, the Board of Governors of the Federal Reserve System, Danmarks Nationalbank, the European System of Central Banks, or the National Bureau of Economic Research.

Greater Than the Sum of Its Parts: Aggregate vs. Aggregated Inflation Expectations
Alexander Dietrich, Edward S. Knotek II, Kristian O. Myrseth, Robert W. Rich, Raphael Schoenle, and Michael Weber
NBER Working Paper No. 31822
November 2023
JEL No. C83,E31,E52

ABSTRACT

This paper introduces a novel measure of consumer inflation expectations: We elicit and combine inflation forecasts across categories of personal consumption expenditure to form an aggregated measure of inflation expectations. Drawing on nearly 60,000 respondents, our data comprise the early low-inflation environment of the COVID pandemic and the 2021 inflation surge. Conventionally elicited inflation expectations consistently exceed aggregated measures constructed under plausible weighting schemes. Aggregated measures display less disagreement and volatility and are stronger predictors of consumers' spending plans. The relative informational value of aggregated measures rises with the individual-level gap between conventional and aggregated inflation expectations. Our results chart a new course for designing measurement of inflation expectations.

Alexander Dietrich
Danish National Bank
amdi@nationalbanken.dk

Edward S. Knotek II
Federal Reserve Bank of Cleveland
Edward.Knotek@clev.frb.org

Kristian O. Myrseth
University of York
kristian.myrseth@outlook.com

Robert W. Rich
P.O. Box 6387
Cleveland, OH 44101-1387
robert.rich@clev.frb.org

Raphael Schoenle
Brandeis University
Department of Economics
415 South Street
Waltham, MA 02454
and CEPR and Federal Reserve Bank of Cleveland
schoenle@brandeis.edu

Michael Weber
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
michael.weber@chicagobooth.edu

1 Introduction

Inflation expectations are a decisive factor for charting the course of monetary policy, and central banks "spend a lot of time watching them" (Powell, 2021). Policymakers pay special attention to consumer expectations because they affect households' consumption-saving decisions and wage-price spirals.[1] Consumer inflation expectations, however, are notoriously difficult to capture accurately with prevailing survey methods. Many consumers struggle to grasp the concept of inflation; they rely on salient cues when reporting forecasts; and survey responses are vulnerable to a host of cognitive biases.[2]

These challenges inherent to the measurement of inflation expectations resonate with the canonical work by Tversky and Kahneman (1974) on heuristics and biases in judgment under uncertainty and the literature on the psychology of human judgment (Fischhoff and Broomell, 2020), and they point to a potential mismatch between human cognition and the demands of conventional elicitation techniques. Conventional techniques aim to elicit aggregate inflation, which is a concept comprising price changes across categories of personal consumption expenditures (PCE). Because the cognitive complexity of decisions is likely to increase in the number of relevant components (Gabaix and Graeber, 2023), one might surmise that it would be harder for an individual to produce an expectation for aggregate inflation than for its category-specific constituent parts.

Our paper, therefore, proposes an alternative approach to measuring inflation expectations, by decomposing aggregate inflation into its more tangible components—price changes for disaggregated categories of goods and services. Accordingly, we develop a survey that elicits inflation forecasts for the full range of PCE categories. This granular elicitation allows us to construct a novel measure of *aggregated* consumer inflation expectations, by combining category-specific forecasts. We find that aggregated inflation expectations are not only less noisy than the conventional measure of aggregate inflation expectations, but also predict planned consumer spending better. These properties hold consistently across respondents in the population. Arguably, therefore, *aggregated* inflation expectations yield a more informative representation of consumers' effective beliefs about future inflation—that is, the beliefs on which they make future consumption plans—highlighting the appeal of *aggregated* inflation expectations for policymakers who aim to elicit decision-relevant beliefs across the population.

We collect these expectations from almost 60,000 US consumers in a nationally representative

[1]For consumption-savings decisions, see Bachmann et al. (2015), Crump et al. (2022), and Ryngaert (2022); for wage-price spirals, see Blanchard (1986) and Lorenzoni and Werning (2023).

[2]D'Acunto et al. (2022) find that low-IQ respondents, especially, seem to have limited understanding of the concept of aggregate inflation; low-IQ respondents report that they associate inflation with price changes of specific, salient goods rather than macroeconomic variables, and they have difficulty with probabilistic terms. Potential salient cues include extreme price movements (Bruine de Bruin et al., 2011), changes in grocery and gas prices (Binder, 2018; Cavallo et al., 2017; D'Acunto et al., 2021), weighting consumption categories with the frequency of purchases (Georganas et al., 2014), and, under conditions with rapid price increases in specific categories, the category-specific inflation rates (Niu and Harvey, 2023).

survey, at a daily frequency between July 2020 and August 2022, as part of the Federal Reserve Bank of Cleveland's Daily Survey of Consumers (Knotek et al., 2020). The survey measures 12-month-ahead inflation expectations in two distinct ways. First, it elicits aggregate inflation expectations following the conventional point-estimate approach from the New York Fed's Survey of Consumer Expectations (SCE). Second, the survey asks consumers about inflation expectations for each of 11 consumption categories, covering the entire range of PCE. In doing so, we match closely the conventional question format of the SCE for aggregate inflation expectations. While the SCE also elicits inflation expectations for several salient products—such as gasoline, housing, and groceries—to the best of our knowledge, our survey is the first to yield a dataset of comprehensive category-specific inflation expectations. In addition, we also ask survey participants about their own personal consumption expenditures and the relative importance of the consumption categories, along with their consumption plans.

We derive aggregated inflation expectations measures by combining the disaggregated category responses, and we compare them with the conventional, aggregate inflation expectations. In particular, we evaluate eight different procedures for aggregating category-specific inflation expectations, across two conceptual types: (i) *plausibly rational* and (ii) *behavioral* aggregations. The three aggregations within the first type use weights arguably reasonable for a rational agent: self-reported expenditure weights, self-reported importance weights, and the official PCE weights. In contrast, the behavioral aggregations capture weighting schemes that depart from plausibly rational procedures in favor of heuristic mechanisms known in the literature, such as reliance on salient categories (D'Acunto et al., 2021) or salient price changes (Bruine de Bruin et al., 2011) in forming aggregate inflation expectations: equally weighted categories, core and non-core inflation expectations, a max operator selecting the highest category expectation, and a second-max selecting the second highest. These aggregations thus allow for insensitivity to category weights, the heavy weighting of salient categories, and attention to salient price changes, respectively.

When we compare conventionally elicited aggregate inflation expectations to *aggregated* measures in the cross-section, we find the latter are generally lower and less dispersed than the former. The *aggregated* measures are on average also less volatile over time. Turning to a comparison at the respondent level, we obtain significant aggregation gaps between aggregate and *aggregated* inflation expectations, and the absolute gaps vary meaningfully with socioeconomic characteristics: Higher education, for example, yields a much closer alignment between aggregate and *aggregated* expectations.[3] Moreover, subjective uncertainty about aggregate inflation expectations correlates strongly with the absolute aggregation gap. One interpretation is that the more uncertain consumers are about their aggregate forecast, the more the complexity of the aggregation task bears on consumers'

[3]Several other papers have investigated inconsistencies in consumer surveys between responses to questions in different formats asking about aggregate inflation expectations (Stanisławska et al., 2021) or consumer spending (Winter, 2004). Similar to our results, inconsistencies increase with lower socioeconomic status. Professional forecasters, however, seem to be consistent in their forecasts for aggregate inflation, across different question formats (Engelberg et al., 2009).

ability to perform the associated computations. This interpretation resonates with the finding from psychology that individuals adapt the heuristics at play according to the demands of the task at hand (Payne et al., 1993). The aggregation gaps also reveal an inconsistency with rational expectations: Aggregation using the official PCE weights should align with the conventional aggregate forecast—but it does not.[4] This finding complements those of Coibion and Gorodnichenko (2012), who reject the hypothesis of full-information rational expectations (FIRE) by examining expectations and realizations. The internal inconsistency that we document between expectations at the individual level reveals another departure from FIRE.

Finally, we provide evidence for the superior predictive power of our measure of *aggregated* inflation expectations for consumer demand, compared to the conventional aggregate inflation expectations. Theoretically, the link between expected inflation and consumer spending is described by the consumer Euler equation (see for example, Galí, 2015). As the survey elicits planned changes in consumer spending for different goods and services, we can estimate the consumer Euler equation, following Crump et al. (2022). Our measures of *aggregated* inflation expectations *all* emerge as stronger predictors of planned consumer spending. Category-specific inflation expectations, therefore, appear more representative of the beliefs used in planning, and thus more informative for monetary policy, both during the COVID-19 pandemic and the 2021-inflation surge. Moreover, the relative benefit of using expenditure-weighted, aggregated inflation expectations over the more conventional, aggregate inflation expectations to predict spending plans increases with the individual-level gap between the two measures. This highlights the appeal of our measure of *aggregated* inflation expectations for policy makers who aim to elicit effective beliefs across respondent groups in the population.

Our paper builds on a growing literature addressing the formation of consumer inflation expectations, especially in the context of cognitive heuristics, such as reliance on salient cues. Bruine de Bruin et al. (2011), for example, provide evidence that households rely on salient, extreme prices to form their aggregate inflation expectations. In a similar spirit, but with different methods, D'Acunto et al. (2021) find that consumers rely on observed changes in grocery prices to form their aggregate inflation expectations and that the relative weights products receive depend on the frequency of purchase, rather than expenditure. Others have documented that consumers form aggregate inflation expectations by extrapolating from gasoline prices (e.g., Armantier et al., 2016; Binder and Makridis, 2022; Binder, 2018; Coibion and Gorodnichenko, 2015) or relying on recalled price changes (Cavallo et al., 2017; D'Acunto and Weber, 2022). Arora et al. (2013) and Trehan (2011) find that aggregate inflation expectations react excessively to non-core price changes, and Dietrich (2023) shows that consumers are relatively more attentive to their internal food and energy inflation forecasts. This literature highlights that particular pieces of information may carry

[4] Likewise, the gap resulting from aggregation by personal expenditure weights, as well as that by importance weights, may be interpreted as inconsistent with rational expectations, given a suitable model.

significant weight in the formation of inflation expectations—in line with our findings. Ironically, despite the evidence that they are relevant for inflation expectations, neither groceries nor gasoline are part of core inflation upon on which monetary policymakers often focus attention.

Beyond the domain of inflation expectations, specifically, past experiences may play a more general role in belief formation (Malmendier and Nagel, 2015), including expectations about future macroeconomic conditions or house prices (Kuchler and Zafar, 2019). The recent framework of selective recall and memory by Bordalo et al. (2022, 2023) posits a theory of belief formation in this context that provides a nuanced view of the role of experiences. To assess the probability of a scenario, individuals compare the similarity of a scenario at hand with the experiences they recall following a cue; such recall either facilitates the probability assessment or interferes with it. Our results accord with this class of models: Recall may prove easier in the context of concrete consumption categories than in the case of abstract aggregate inflation, which may offer more interference and thereby lead to noisier expectations.[5] Our finding that category-specific expectations are less dispersed and less volatile over time than aggregate inflation expectations is consistent with this interpretation. Likewise, that subjective uncertainty about aggregate inflation expectations correlates positively with the absolute aggregation gap—the difference between aggregate and *aggregated* inflation expectations—could reflect the increased difficulties in mapping recalled experiences to the aggregate inflation forecast. Appendix D presents a version of the Bordalo et al. (2023) framework that explicitly rationalizes our findings.

Our paper draws also on methodological insights from survey studies across areas of economics and related fields that find data-quality advantages from decomposing broad questions into subquestions. Menon (1997), for example, shows that the accuracy of frequency reports depends on the question format matching the cognitive processes employed by the respondent and that decomposed questions, therefore, improve frequency judgments of irregular events by easing the cognitive reporting burden. Consistent with these results, Winter (2004) finds that disaggregated questions yield improved data quality for nondurable consumption compared to questions asking about aggregates, and discrepancies vary with socioeconomic characteristics, similarly to what we find with the variation in the gaps between aggregate and *aggregated* inflation expectations. Along the same lines, but in the domain of development economics, Deaton (2019) argues that surveys of consumption spending with disaggregated questions are more reliable than those with questions about aggregates. Hurd and Rohwedder (2008, 2012), moreover, field surveys to ask households about past spending using disaggregated category questions. Taken together, this literature implies a possibility for improving measurement of consumer inflation expectations by eliciting expectations at the disaggregated, category-specific level.

Our paper proceeds as follows: Section 2 describes our novel survey data. Section 3 examines

[5]Several papers have documented a strong link between recall and inflation expectations (e.g., Jonung, 1981; Weber et al., 2022).

category-specific inflation expectations and compares them to aggregate inflation expectations. Section 4 investigates procedures for aggregating category-specific inflation expectations and explores the gap between aggregate and *aggregated* inflation expectations. Section 5 relates aggregate and *aggregated* expectations to household spending plans in Euler equation estimations. A final section concludes.

2 Survey

Our survey is conducted at a daily frequency, as a module within the Federal Reserve Bank of Cleveland's daily survey of consumer expectations, administered by Qualtrics Research Services. It includes a nationally representative sample of 59,920 responses, collected between July 9 2020 and August 9 2022, with a daily sampling size of at least 100 respondents. Qualtrics Research Services constructs a representative sample by drawing respondents from several actively managed, double-opt-in market-research panels, complemented with social media (Qualtrics, 2019). Dietrich et al. (2022) and Knotek et al. (2020) provide further information about other parts of the survey.

We require all respondents to be US residents and to speak English as their primary language. Respondents are representative of the US population according to several key demographic and socioeconomic characteristics; they have to be male or female with 50-percent probability; approximately one third is targeted to be between 18 and 34 years of age, another third between 35 and 55, and a final third older than age 55. We also require a distribution across US regions in proportion to population size, drawing 20 percent of our sample from the Midwest, 20 percent from the Northeast, 40 percent from the South and 20 percent from the West. The survey includes filters to eliminate respondents who enter nonsensical answers for at least one response, or who complete the survey in less (more) than five (30) minutes, and reCAPTCHA tests to reduce the likelihood that bots would interfere.[6]

Our sampling criteria generated a sample roughly representative of the US population along the targed dimensions, as seen in Table 1. To improve the fit further, we compute survey weights for each respondent; we apply iterative proportional fitting to create respondent weights following completion of the survey ("raking," see for example, Bishop et al., 1975; Idel, 2016). This allows us to calculate statistics that are *exactly* representative of the US population also according to age, gender, ethnicity, income, census region, and education—the variables in the right-hand column of Table 1.

Within the survey, we ask respondents first about their aggregate inflation expectations over the next 12 months (Q1 in Table 2), using point-forecast questions.[7] Our approach to eliciting aggregate inflation forecasts is methodologically similar to that of other influential household surveys, such

[6] Qualtrics Research Services provides the filtered data. The daily sample size refers to the number of respondents after filtering. Survey respondents receive monetary compensation for their time.

[7] On a subset of the data, we switched the ordering, asking about disaggregated category expectations first. We did not find a significant order effect.

Table 1: Survey Respondent Characteristics

	Survey	US population		Survey	US population
Age			**Race**		
18-34	33.1%	29.8%	non-Hispanic white	72.7%	60.1%
35-55	33.8%	32.4%	non-Hispanic black	9.3%	12.5%
>55	33.1%	37.8%	Hispanic	10.1%	18.5%
			Asian or other	7.9%	8.9%
Gender					
female	49.9%	50.8%	**Household Income**		
male	49.7%	49.2%	less than 50k$	47.8%	37.8%
other	0.4%	-%	50k$ - 100k$	29.5%	28.6%
			more than 100k$	22.7%	33.6%
Region					
Midwest	20.6%	20.7%	**Education**		
Northeast	21.9%	17.3%	some college or less	50.6%	58.3%
South	39.5%	38.3%	bachelor's degree or more	49.4%	41.7%
West	18.0%	23.7%			
			N=59,920		

Notes: The "Survey" column represents characteristics in our survey; the "US population" column gives the value for the US population, obtained from the US Census Bureau (Household income: CPS ASEC, 2021; gender, education: ACS, 2019, age, race, region: National Population Estimate, 2019).

as the University of Michigan's Surveys of Consumers (SoC) and the New York Fed's Survey of Consumer Expectations (SCE).[8] Subsequently, we elicit inflation expectations for 11 PCE categories (Q2 in Table 2) using a format closely aligned to that of the aggregate inflation question (Q1). For each category, however, we provide survey participants with at least one example—such as "Public transit tickets and airfare" for "Transportation services"—to reduce the risk of misinterpreting categories. Table 3 in Section 3 shows both the PCE categories used in the survey and some summary statistics. Our PCE-disaggregation is based on that of the US national income and product accounts (NIPA), with some small sectors combined in order to reduce the cognitive burden of completing the survey.[9]

While the SCE also elicits aggregate inflation expectations with a probability-distribution question, we choose to rely on point forecasts both for the aggregate and the category expectations.[10] The principal reason is that point forecasts prove more tractable in the present survey framework,

[8]The SoC has collected data on household inflation expectations since 1978; the SCE started in 2013. Both ask directly about aggregate inflation or the expected change in *aggregate* prices, at a monthly frequency, and they include some kind of panel structure; while the SoC asks a subset of participants to respond to the survey again, half a year later, the SCE has a rolling panel structure, with respondents answering 12 consecutive monthly surveys. Our survey does not feature a panel structure, but is conducted at a higher frequency (daily).

[9]We use what might be thought of as the third level of disaggregation of PCE-spending—the first would be by goods and services, and the second by durable and nondurable goods and expenditures on services, by households and nonprofit institutions serving households.

[10]We do, nonetheless, feature a probability-distribution question on aggregate inflation, but use it exclusively as a measure of subjective uncertainty.

Table 2: Survey Questions

Aggregate Inflation Question		
Q1	What do you expect the rate of inflation to be over the next 12 months? [...]	I expect [...] to be [positive/negative] ___ percent over the next 12 months.
Category Inflation Questions		
Q2	Twelve months from now, what do you think will have happened to the price of the following items?	I expect the price of [*category*] to [increase/decrease] by ___ percent.
Q3	In terms of consumption spending, how much money did you spend on each of the following broad consumption categories during the last month? [...]	Per category, participants enter an approximate amount in dollars in a bracket.
Q4	Which of the following broad consumption categories matter the most to you right now in your daily life? Please move the slider to indicate the importance for each of them [...]	Participants move a slider from 0 (no importance) to 100 (highest importance), per category.
Spending Questions		
Q5	Compared with your spending last month, how do you expect your total spending to change in the next twelve months?	[up/no change/down] by ___ percent.
Q6	Compared with your spending on services [...] last month, how do you expect your total spending to change in the next twelve months?	[up/no change/down] by ___ percent.
Q7	Compared with your spending on non-durable goods [...] last month, how do you expect your total spending to change in the next twelve months?	[up/no change/down] by ___ percent.

Notes: List of main questions asked in the survey. For other questions, please see Appendix E.

reducing the mental burden on participants who would otherwise have to indicate probability distributions for all 11 PCE categories. Moreover, Clements (2014) finds that point forecasts, relative to probability-distribution forecasts, offer superior data quality for the mean of expectations.

Besides inflation expectations within these categories, we also asked how much survey respondents spent on goods and services in the respective category during the last month (Q3 in Table 2) and how important they consider the category in their daily lives (Q4 in Table 2). Responses to these questions allow us to compute both expenditure shares per category (relative to total expenditure) and a measure of perceived relative importance.

Table 3: Summary Statistics

	Mean	Disagreement	Time-Series Volatility
Aggregate expectation	6.39	7.53	2.53
Category expectations			
Motor vehicles	5.49	5.95	1.78
Recreational goods	4.00	6.34	1.61
Other durable goods	4.12	6.14	1.69
Food and beverages	5.27	6.48	1.71
Gasoline	5.28	7.57	2.03
Other nondurable goods	4.15	6.02	1.41
Housing and utilities	4.93	6.46	1.50
Health care	3.96	6.52	1.58
Transportation services	4.82	6.19	1.53
Food services	4.78	6.46	1.54
Other services	4.32	5.64	1.29

Notes: This table presents summary statistics on the mean on expectations, the standard deviation in the cross-section, and the standard deviation in the (daily mean) time series. Mean expectation: Time-series mean of daily Huber-robust and survey-weighted mean expectations (see figure 1, upper row); Disagreement: Time-series mean of daily Huber-robust and survey-weighted standard deviation of expectations (see figure 1, lower row); Time-series volatility: Time-series standard deviation of daily Huber-robust and survey-weighted mean expectations.

Subsequently, we asked respondents to indicate their expected spending 12 months ahead, relative to that in the month prior to the survey. We also repeated the question for more narrowly defined spending categories, namely services spending and expenditures on nondurable consumption goods. Additionally, respondents reported their socioeconomic background and consumer habits. These questions, including demographic information and the exact layout of our inflation questions, are provided in Appendix E.

3 Category-Specific Inflation Expectations

This section presents the statistical properties of aggregate and category-specific inflation expectations (Q1 and Q2 in Table 2, respectively). We document that mean expectations about aggregate inflation in the cross-section exceed mean inflation expectations for every PCE category. In addition, aggregate expectations exhibit larger disagreement (except for gasoline), as well as higher volatility within the time series.

Between July 2020 and August 2022, survey participants expect on average aggregate inflation over the next 12 months to be 6.39 percent (see Table 3). The table reports the mean expectation and disagreement among households (cross-sectional standard deviation) in the first and second columns, with displayed statistics representing the average over the daily Huber-robust and survey-weighted mean and standard deviation, respectively.[11] Strikingly, every category-specific inflation rate is expected to be lower than the aggregate rate: from 3.96 percent for "Health care services" to 5.49 percent for "Motor vehicles."

An agent with views mirroring those of the cross-section, therefore, expected that aggregate inflation would exceed inflation expectations in *any* category. This pattern is driven by respondents reporting aggregate expectations outside the range of their own individual category-specific expectations: about 26 percent of respondents state an aggregate expectation greater their expectation for any category. For 12 percent of respondents, the opposite holds true; they report aggregate expectations below their smallest category-specific expectation. Consequently, only around 62 percent of respondents report their aggregate within the range of their category-specific expectations. Although such inconsistencies in isolation could be explained by random reporting errors on the part of the individual respondents (Bertrand and Mullainathan, 2001), white-noise reporting errors cannot account for the cross-sectional wedge between aggregate and category-specific inflation expectations.

The survey is designed such that all categories combined cover the entire range of US consumption expenditures, which— following the statistical methodology of the PCE price index reported by the BEA—form the basis for aggregate inflation. In theory, therefore, a linear combination of weights should exist for the representative agent, summing up to unity, such that the weighted category-specific expectations equate to the aggregate inflation expectation. This, however, is clearly not the case in the data.

A potential explanation for this gap might be that respondents interpret the aggregate inflation question as referring to a macroeconomic variable whereas they understand the category-specific questions as referring to subjective inflation rates, that is, based on the goods and services that they personally consume (within the specific category). However, the survey is designed to allow a commensurate comparison between aggregate and category-specific expectations, as both question types ask about inflation in general, as opposed to subjective, personal inflation rates. Following the New York Fed SCE, aggregate expectations ask about "inflation/deflation," while category-specific expectations refer to changes in "the price of" a category, with no suggestion that only one question type applies specifically to personal, subjective consumption. Thus, although we cannot rule out a subjective interpretation of every category-specific question, there is little reason to assume that a subjective interpretation, alone, would account for the asymmetric results obtained across aggregation levels.

As opposed to mere white noise, the pattern in Table 3 suggests that differential heuristics and expectations-formation processes could be at play when respondents report aggregate versus category-specific inflation expectations. That is, respondents might adapt the heuristics used ac-

[11] In order to minimize the sensitivity of reported statistics to outliers in the survey, the top and bottom 1% of responses for the aggregate and category inflation expectations are truncated. Huber-robust weights are applied to further reduce the sensitivity of reported statistics to outliers Huber (1964).

cording to the demands of the task at hand (i.e., Payne et al., 1993). We run a series of robustness checks that provide further insight into this pattern.

First, in a separate survey, we asked respondents about specific aggregate inflation indices, namely PCE- or CPI-price-index inflation (see Table A.4 in the appendix). Our main findings hold up qualitatively; aggregate inflation expectations exceed inflation expectations for any category. In fact, when survey participants are asked about CPI or PCE inflation, specifically, the gap to category-specific responses appears to widen.

Second, to ascertain that the relationship between aggregate and category-specific expectations is independent of a framing artifact, we randomized the order of aggregate and category-specific inflation questions for a subset of the sample (see Table A.5 in the appendix); the mean aggregate expectation exceeds any category-specific expectation, irrespective of whether we first elicit aggregate or category-specific expectations.

Third, to explore whether the patterns in the cross-section extend beyond the typical consumer, we administered an abridged version of our survey to a small sample of fund managers, who volunteered to participate in the lead-up to a practitioner's conference in November 2022. As seen in Appendix A.6, fund managers reported aggregate inflation expectations higher than category-specific expectations, with the exception of food and beverages. This is roughly consistent with the pattern we obtain for consumers (Table 3), but a notable contrast arises for the disagreement in aggregate inflation expectations. While smaller than that of any category for fund managers, it is larger for consumers. This result could be explained by the potentially differential memory base held by fund managers and according to the framework of selective recall and memory in belief-formation (Bordalo et al., 2022, 2023), which we discuss below.

Fourth, we gauge the stability of our results over time. The upper row of Figure 1 shows the time series, daily means, for aggregate and mean category-specific inflation expectations during the survey period. The left panel displays category-specific expectations for the durable (red lines) and nondurable (blue lines) consumption goods, whereas the right panel shows services categories (green lines). All time series are centered 11-day moving averages.[12] Aggregate inflation expectations, rising from around 4 percent in July 2020 to around 8 percent in July 2022, are higher than any category expectations for most of the sample period.[13] Consequently, for a representative agent, there is no possible linear combination of category-specific expectations with non-negative aggregation weights that maps category-specific expectations into aggregate expectations.

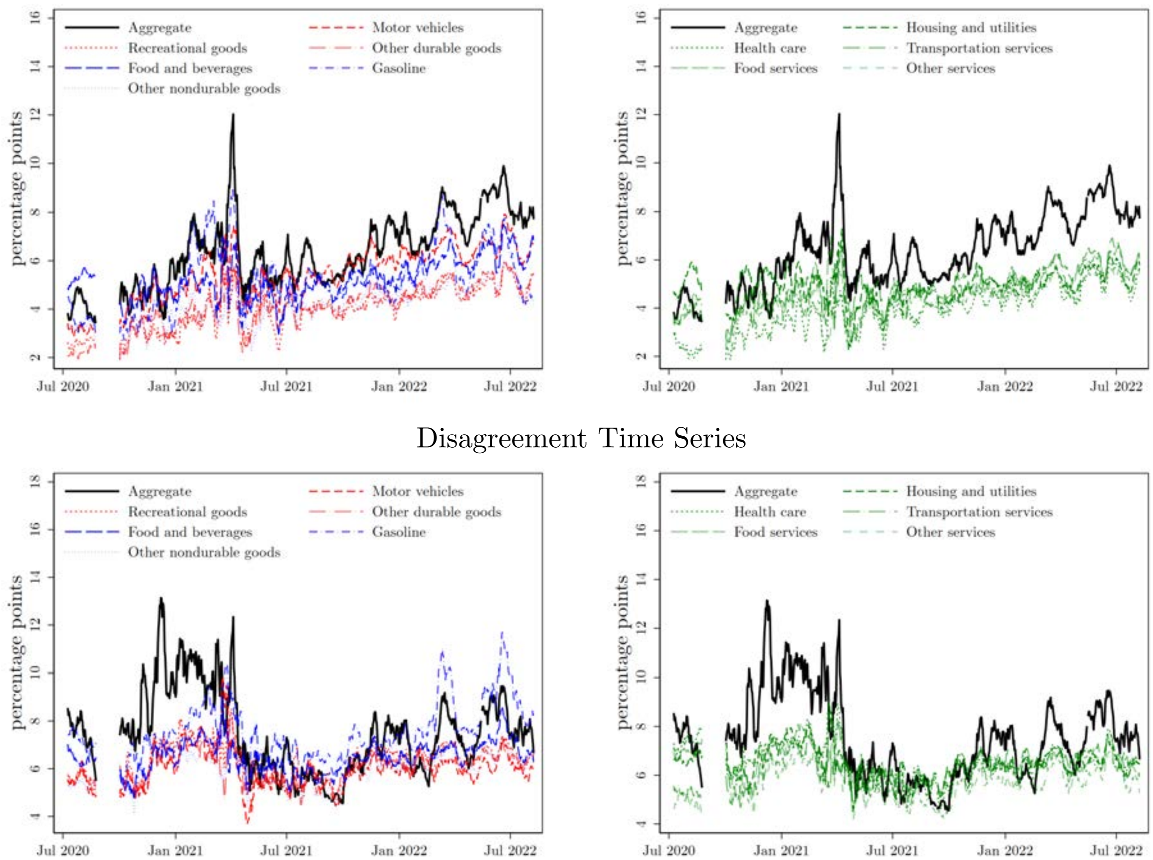
Time-series volatility (of expectations) is an important moment in economic analysis because it conveys information about cyclical dynamics, such as the role of rising economic uncertainty embodied in expectations (Dietrich et al., 2022). We find that volatility over time is higher for

[12]The centered moving average constructs for each day the average of the mean from the respective day and the five days before and after.

[13]The time series documents a temporary but pronounced increase in inflation expectations and disagreement in early 2021, with a spike around April 2021, coinciding with the surge in realized inflation and inflation news in the US.

Figure 1: Aggregate vs Category-Inflation Expectations

Mean Time Series



Notes: The top row shows mean aggregate inflation (black line) and category-inflation rates; the bottom row shows disagreement on aggregate inflation; left panels show durable and nondurable goods inflation by category; right panels show services inflation by category; the time series is an 11-day centered moving average. Underlying daily observations are Huber-robust and survey-weighted means. Questions on inflation expectations were not part of the survey during September 2020.

aggregate inflation expectations than it is for category-specific expectations (see Column 3 in Table 3). The bottom row of Figure 1 shows disagreement among respondents for aggregate inflation expectations (black line) and category-specific expectations, where we measure disagreement as the daily standard deviation across survey respondents. The figures display an 11-day moving average, with durable- and nondurable-goods sectors in the left panel and services in the right. For most of the sample period, disagreement is much higher for aggregate expectations than it is for category-specific expectations (see also Table 3). This pattern could be interpreted within the framework of selective recall and memory in belief-formation by Bordalo et al. (2022, 2023); recall may be facilitated in the context of concrete consumption categories relative to the case of abstract aggregate inflation, which may trigger more interference and thereby lead to noisier expectations.

This mechanism could explain why category-specific expectations are less dispersed and less volatile over time than aggregate inflation expectations.

Finally, we consider the pattern of demographic heterogeneity in inflation expectations (see Tables A.1 and A.2 in the Appendix). Lower income and less education are both associated with a substantially higher mean aggregate inflation expectation and higher cross-sectional disagreement. At the same time, category-specific expectations tend to be quite similar across education and income levels and, where they are not, do not diverge in a consistent fashion. Across almost all categories, women report higher inflation expectations and greater disagreement; this pattern holds also for aggregate inflation expectations, generally consistent with demographic patterns reported by Bruine de Bruin et al. (2010). A paradoxical pattern, however, arises with age (see Table A.3 in the Appendix). For the oldest age group in our sample (older than 55), aggregate inflation expectations are lower than those of younger respondents—but for expectations by category, the pattern is reversed: older respondents report higher expectations.

4 Aggregate vs. Aggregated Inflation Expectations

Next, we show that aggregate inflation expectations differ significantly from *aggregated* measures, which describe overall inflation expectations by aggregating category-specific expectations. *Aggregated* inflation expectations tend to be lower than aggregate expectations, and disagreement among survey participants is higher for the latter. A statistically significant, positive gap between aggregate inflation expectations and both expenditure- and PCE-weighted aggregations is especially noteworthy as it reflects internally inconsistent beliefs about inflation. The gap increases with uncertainty and varies in a meaningful way with socioeconomic and demographic characteristics. Section 4.1 introduces the aggregation methods, Section 4.2 the statistical properties of *aggregated* inflation expectations relative to aggregate expectations, and Section 4.3 the relation between aggregate inflation expectations and the *aggregated* measures.

4.1 Aggregated Inflation Expectations

We build multiple measures of *aggregated* inflation expectations, relying on category-specific expectations and sets of weights ω_k. For every set of weights, we assume the *aggregated* inflation expectation is a weighted average of categories in the sense that $\omega_k \geq k$ and that $\sum_{k=1}^{N} \omega_k = 1$. Thus,

$$\mathbb{E}_t^i \, \pi_{t+1}^{aggregated} = \sum_{k=1}^{N} \left[\omega_k^i \, \mathbb{E}_t^i \, \pi_{k,t+1} \right], \tag{1}$$

where $\mathbb{E}_t^i \, \pi_{t+1}^{aggregated}$ denotes the *aggregated* inflation expectation of respondent i and $\mathbb{E}_t^i \, \pi_{k,t+1}$ their expectations of category k. ω_k^i denotes the weight assigned to category k by respondent i.

Table 4: Aggregated Expectations - Weights

$\mathbb{E}_t^i \pi_{t+1}^{aggregated}$	Weights ω_k	Notes
Plausibly rational aggregation		
$\mathbb{E}_t^i \pi_{t+1}^{PCE}$	$\omega_k = \frac{C_{k,t}^{PCE}}{\sum_{k=1}^N C_{k,t}^{PCE}} \quad \forall k \forall i$	**PCE weights**; $C_{k,t}^{PCE}$ denotes monthly PCE expenditure from BEA.
$\mathbb{E}_t^i \pi_{t+1}^{exp}$	$\omega_k^i = \frac{C_{k,t}^i}{\sum_{k=1}^N C_{k,t}^i} \quad \forall k$	**Expenditure weights**; $C_{k,t}^i$ denotes average monthly expenditure of i on category k.
$\mathbb{E}_t^i \pi_{t+1}^{imp}$	$\omega_k^i = \frac{Imp_{k,t}^i}{\sum_{k=1}^N Imp_{k,t}^i} \quad \forall k$	**Importance weights**; $Imp_{k,t}^i \in [0, 100]$ denotes subjective importance to consumption of category k for i.
Behavioral aggregation		
$\mathbb{E}_t^i \pi_{t+1}^{equal}$	$\omega_k = \frac{1}{N} \quad \forall k \forall i$	**Equal weights**; each category receives the same weight.
$\mathbb{E}_t^i \pi_{t+1}^{core}$	$\omega_k^i = \frac{C_{k,t}^i}{\sum_{k=1}^N C_{k,t}^i} \quad \forall k \neq \{Gas, Food\}$ $\omega_k = 0 \quad \forall k = \{Gas, Food\}$	**Core-inflation weights**; relative average monthly expenditure of i on category k except for food and gasoline. Gas and food weights equal 0.
$\mathbb{E}_t^i \pi_{t+1}^{non-core}$	$\omega_k^i = \frac{C_{k,t}^i}{\sum_{k=1}^N C_{k,t}^i} \quad \forall k = \{Gas, Food\}$ $\omega_k = 0 \quad \forall k \neq \{Gas, Food\}$	**Non-core-inflation weights**; relative average monthly expenditure of i on food and gasoline. All other weights equal 0.
$\mathbb{E}_t^i \pi_{t+1}^{1^{st} max}$	$\omega_k^i = 1 \forall k = m; \omega_k^i = 0 \forall k \neq m$ $\mathbb{E}_t^i \pi_{m,t+1} = 1^{st} \max(\{\mathbb{E}_t^i \pi_{k,t+1}\})$	**Max**; aggregate expectation equal to highest category expectation.
$\mathbb{E}_t^i \pi_{t+1}^{2^{nd} max}$	$\omega_k^i = 1 \forall k = m; \omega_k^i = 0 \forall k \neq m$ $\mathbb{E}_t^i \pi_{m,t+1} = 2^{nd} \max(\{\mathbb{E}_t^i \pi_{k,t+1}\})$	**Second max**; aggregate expectation equal to second highest category expectation.

Notes: The table describes the construction of aggregated inflation expectations, based on the category-specific expectations as well as different sets of weights.

Our analysis considers two types of weights, summarized in Table 4. The first denotes weights that describe a plausibly rational agent and the second weights that describe a behavioral agent. Among the plausibly rational weights, a first set relies on the official monthly BEA nominal expenditure shares used to construct the official PCE-inflation statistics. In a FIRE general-equilibrium model, multiplying category-specific expectations with category-specific weights yields the aggregate economy-wide inflation expectation precisely up to the usual first-order log-linear approximation.[14] A second set of weights aggregates category inflation expectations with self-reported

[14] Up to second order, aggregate inflation and analogously, its expectation, is given by the (appropriately weighted) mean of category inflation $\bar{\pi}_{k,t}$ and a second-order variance term: $\pi_t \approx \bar{\pi}_{k,t} + \frac{1}{2} C \text{var}(\pi_{k,t})$, where C denotes a constant. This result follows directly from a Taylor approximation to common price aggregators. Appendix C.1 provides an example. Our analysis focuses on the first-order approximation. We do so due to the fact presented

expenditure shares. A third set uses weights derived from questions asking respondents to indicate the qualitative "importance" of each category for their consumption. The latter two sets of weights should be especially relevant for respondents who aggregate category-specific expectations according to personal consumption baskets. Together with the first set of weights, our analysis of consistency thus accounts for respondents having potentially different concepts of inflation in mind—either their personal or the official, published inflation rate.

The remaining five sets of weights, in contrast, represent forms of "behavioral" expectations formation. A first assigns equal weights, reflecting an agent who notices price changes but neglects expenditure shares. A second takes the self-reported expenditure weights discussed above, but sets food and gasoline weights to zero; this reflects an agent who pays attention to core inflation. A third is the opposite of the aforementioned, reflecting an agent who pays attention to non-core inflation. Non-core weights are motivated by earlier work, which demonstrates the salience of non-core prices for households, such as D'Acunto et al. (2021) for grocery prices or Trehan (2011), Coibion and Gorodnichenko (2015), Binder (2018), or Binder and Makridis (2022) for gas and energy prices. In particular, Arora et al. (2013) find that household inflation expectations react excessively to non-core price changes. A fourth and fifth set of weights take the highest and second-highest category expectation of each survey participant, respectively, as the aggregated inflation expectations, setting all other weights to 0. The choice of these measures is motivated by Bruine de Bruin et al. (2011), who find that extreme inflation rates play an important role in household expectations.

4.2 Statistical Properties of Aggregated Inflation Expectations

Based on these aggregation schemes, the following characteristics of aggregated inflation expectations emerge: First, mean aggregate inflation expectation exceeds those of all three plausibly rational aggregations, as well as equal-weighted expectations and non-core- and core-inflation expectations; it is lower than those of both max operators. Second, in the cross-section, the standard deviation of aggregate inflation expectations is higher than that of all aggregations, except the max operator. Similarly, in the time-series, aggregate inflation expectations and the max operator yield the two highest standard deviations. Table 5 provides summary statistics. Fund managers in our auxiliary survey (A.13) also report aggregate inflation expectations higher than most aggregations (all except the max operator), but disagreement for aggregate expectations is consistently lower.

For the cross section, we illustrate these differences between aggregate inflation expectations and *aggregated* inflation expectations by means of a bin-scatter plot in Figure 2.[15] Two features

above about bounds for the range of category expectations relative to the locus of reported aggregate expectations: A large fraction of respondents reports aggregate expectations either above or below the category range. We can thus rule out that a systematic one-sided approximation error drives our results.

[15] In a related exercise, in Table A.11 in the Appendix, we regress aggregate inflation expectations on *aggregated* expectations and a constant. For all measures of *aggregated* expectations, we find a positive, highly significant constant, as well as an *aggregated*-inflation-expectations coefficient smaller than one. The R^2 is largest for the

Table 5: Summary Statistics

	Mean	Disagreement	Time Series Volatility
Aggregate expectation	6.39	7.53	2.53
Aggregated expectations			
Plausibly rational aggregation			
Expenditure weights	4.95	5.28	1.30
Importance weights	4.59	4.73	1.33
PCE weights	4.46	4.62	1.24
Behavioral aggregation			
Equal weights	4.49	4.59	1.34
Core inflation	4.72	5.21	1.25
Non-core inflation	5.72	6.26	1.67
Max	11.29	8.50	2.81
Second max	6.96	6.44	1.80

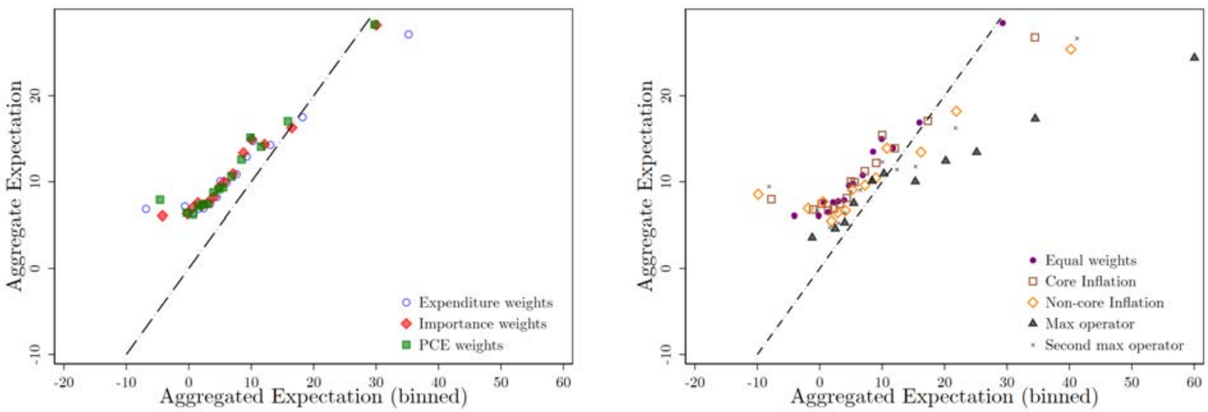
Notes: This table presents summary statistics on the mean on expectations, the standard deviation in the cross-section, and the standard deviation in the (daily mean) time series. Mean expectation: Time series mean of daily Huber-robust and survey-weighted mean expectations (see figure 3, upper row); Disagreement: Time series mean of daily Huber-robust and survey-weighted standard deviation of expectations (see figure 3, lower row); Time series volatility: Time series standard deviation of daily Huber-robust and survey-weighted mean expectations.

stand out: First, almost all observations lie above the 45°-line, indicating that aggregate inflation expectations tend to be higher than *aggregated* measures. This pattern, however, does not hold for the highest levels of *aggregated* expectations, above a cut-off of 18 percent inflation over the next 12 months. Second, the relationship is nonlinear; beyond a certain upper threshold, more extreme *aggregated* expectations correspond to only slightly more extreme aggregate expectations, whereas below a certain lower threshold, aggregate expectations diverge more. The same pattern holds if the conventional, aggregate inflation expectations are binned on the horizontal axis. While a strong ordinal relationship exists for moderate responses, more extreme responses within the conventional measure of inflation expectations do not necessarily correspond to equally extreme *aggregated* expectations beyond a certain upper threshold.

Several time-series patterns emerge, as well. Figure 3 shows that aggregate inflation expectations generally exceed the plausibly rational aggregations (top-left panel), the equal weights and core aggregations (top-right panel), but are exceeded by the max operator; the second-max and non-core aggregations cluster near aggregate inflation expectations. The bottom row of Figure 3 shows that disagreement in aggregate inflation expectations, measured as the daily cross-sectional standard deviation of expectations, consistently exceeds that in the plausibly rational aggregations (bottom-left panel), equal-weight aggregations (bottom-right panel), and, until about April

behavioral, equal-weights aggregation, showing that it explains the largest share of variation in reported aggregate expectations. The subsequent section shows which factors correlate with the gap between the two measures.

Figure 2: Aggregate vs. Aggregated Expectations



Notes: The figure divides aggregated expectations into 15 equal-sized bins and computes mean aggregate inflation expectations for each bin. Left panel: Blue circles: expectations aggregated using reported expenditure shares. Red diamonds: expectations aggregated using reported importance weights. Green squares: expectations aggregated using monthly PCE-weights. Right panel: Purple circles: expectations aggregated using equal weight. Brown squares: core-inflation expectations using reported expenditure shares. Orange diamonds: non-core-inflation expectations using reported expenditure shares. Dark grey triangles: max of category expectations. Light grey crosses: second max of category expectations.

2021, that in core, non-core, and second-max aggregations—after which it roughly coincides with disagreement in the latter three aggregations.

Two additional patterns time-series patterns are worth highlighting. First, the spike in aggregate inflation expectations around April 2021, contemporaneous with a surge in realized inflation and inflation news in the US, far exceeds those observed for plausibly rational aggregations (upper-left panel). Second, following the shift into a high-inflation regime in November 2021, the gap appears to widen between aggregate inflation expectations and plausibly rational aggregations.

4.3 Gap between Aggregate and Aggregated Inflation Expectations

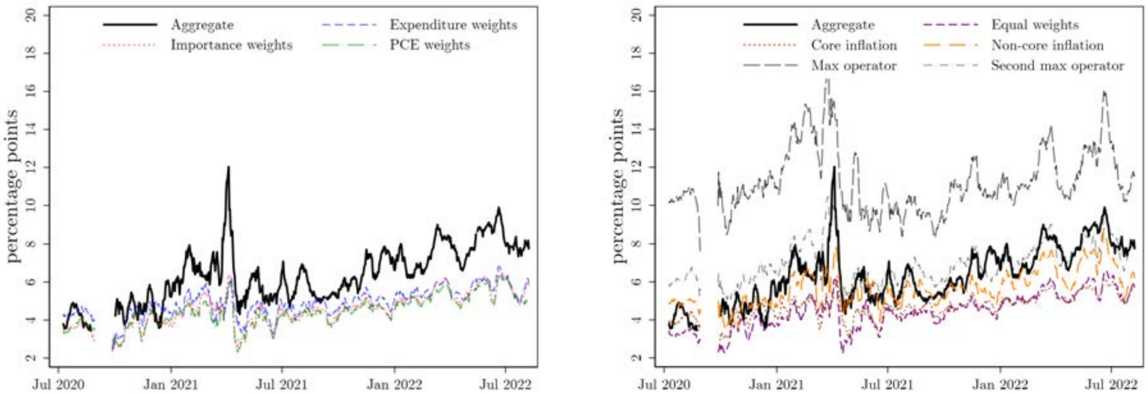
This section shows that the relationship between aggregate and *aggregated* inflation expectations, at the individual level, relates to socio-demographic characteristics and the uncertainty of aggregate inflation expectations. For this purpose, we define the *aggregation gap* as the difference between the aggregate expectation and any aggregator of category-specific inflation expectations:

$$\Lambda_i = \mathbb{E}_t^i \, \pi_{t+1} - \mathbb{E}_t^i \, \pi_{t+1}^{aggregated}.$$

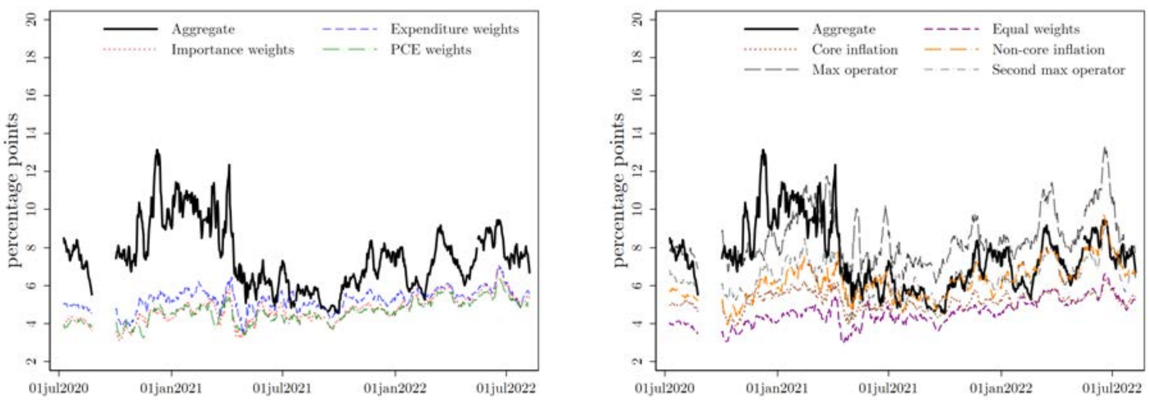
Λ_i defines the aggregation gap for survey participant i as the difference between the aggregate forecast $\mathbb{E}_t^i \, \pi_{t+1}$ and an *aggregated* expectation measure $\mathbb{E}_t^i \, \pi_{t+1}^{aggregated}$. Table 6 presents Huber-robust and survey-weighted estimates, across all individuals in our sample, for the absolute aggregation gap by *aggregated* measures. The absolute aggregation gap provides a measure of the discrepancy

Figure 3: Aggregate vs Aggregated Measures

Mean Time Series



Disagreement Time Series



Notes: The top row shows time-series for mean aggregate inflation expectations; the bottom the time-series for disagreement on aggregate inflation, as the daily cross-sectional standard deviation of expectations. The panels show an 11-day centered moving average of daily observations. Underlying daily observations are Huber-robust and survey-weighted means. In each panel, aggregate inflation expectations are given by a black line, measures of *aggregated* inflation expectations by colored lines. Questions on inflation expectations were not part of the survey during September 2020.

between aggregate and *aggregated* inflation expectations irrespective of sign. The max operator yields the largest gap, whereas the PCE-weights aggregation yields the smallest.

4.3.1 Demographics and the Aggregation Gap

When regressing the absolute aggregation gap on an array of demographic and socio-economic characteristics, we find that women tend to display a higher aggregation gap than men, as do younger relative to older respondents. Moreover, higher education is associated with a lower gap, consistent with the notion that responses to at least one of the two inflation-expectation measures—aggregate or *aggregated*—become noisier when the inflation questions are experienced as more complex or difficult to understand. Table A.9 in the Appendix summarizes these findings.

Table 6: Summary Statistics

	Absolute Aggregation Gap abs(Λ_i)
Plausibly rational aggregation	
Expenditure weights	5.64***
Importance weights	5.49***
PCE weights	5.33***
Behavioral aggregation	
Equal weights	5.34***
Core inflation	5.67***
Non-core inflation	6.06***
Max	9.09***
Second max	6.51***

Notes: This table presents Huber-robust and survey-weighted estimates for the mean absolute aggregation gap; Stars: significance level of a t-test that numbers are different from zero. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$;

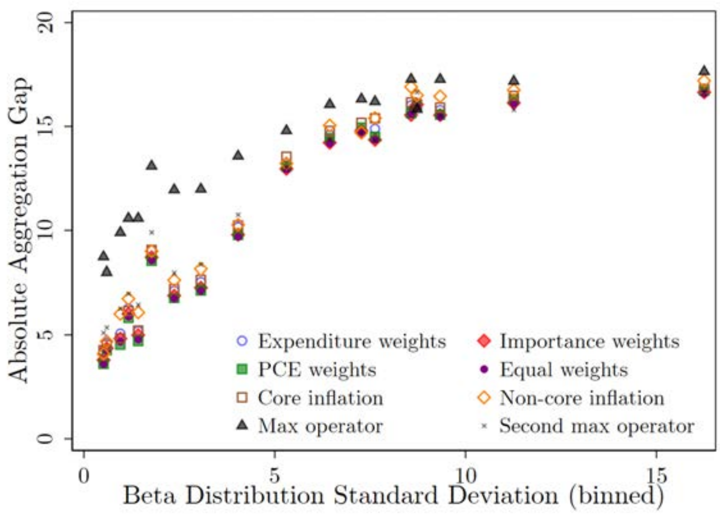
These results align with D'Acunto et al. (2019, 2022), who find that cognitive abilities play an important role in forecast accuracy. Moreover, D'Acunto et al. (2022) show that the responses of lower-IQ survey respondents, for which educational attainment might serve as a proxy, are more likely to be rounded, consistent with our interpretation that expectations become noisier for less-educated respondents, thereby yielding a higher aggregation gap. Stanisławska et al. (2021) find congruent demographic patterns for the probability of consistent responses to questions eliciting expected changes in inflation numerically and qualitatively.

4.3.2 Uncertainty and the Aggregation Gap

One way to probe the implications of forecast complexity is to consider the relationship between inflation uncertainty and the absolute aggregation gap. Presumably, elevated uncertainty about inflation expectations may indicate heightened perceived complexity. As a proxy for aggregate inflation-expectations uncertainty at the respondent-level, we take the standard deviation of aggregate inflation expectations reported in a density forecast (QDIST, Appendix E). To obtain this measure, we fit for each respondent an individual beta distribution over the probabilities reported of specific outcomes; the respondent-specific uncertainty is the standard deviation of the distribution fitted. This procedure follows the methodology of Armantier et al. (2017) developed for the SCE.

The absolute aggregation gap increases with respondents' uncertainty about aggregate inflation, as the left panel of Figure 4 shows. A plausible explanation for this pattern is that the cognitive processes underlying aggregate inflation expectations differ from the combination of cognitive processes and aggregation procedures constituting *aggregated* inflation expectations. This pattern could arise because individuals adapt the heuristics at play according to the demands of the task at hand (Payne et al., 1993), and those demands might become differentiated with greater

Figure 4: The Aggregation Gap and Uncertainty



Notes: The figure shows the correlation between the absolute aggregation gap abs(Λ_i) and the individual standard deviation of aggregate inflation expectations obtained via a beta distribution over a probabilistic question.

uncertainty about aggregate inflation. Alternatively, the absolute aggregation gap could be interpreted within the framework of Bordalo et al. (2022, 2023), given the role of selective recall and memory in belief-formation (see also Appendix D). If recall, given the question cue, proves easier in the context of specific consumption categories than in the case of aggregate inflation, for which there is more interference, then elevated uncertainty about aggregate inflation expectations should be a reflection of the relative difficulty. This, in turn, could render aggregate inflation expectations noisier, which would increase the absolute aggregation gap.

Our results are also consistent with those of Ben-David et al. (2018), who find within the SCE that uncertainty about aggregate inflation represents an effective measure of individual confidence in the forecast. Following new information over time, updates in mean expectations are larger for respondents with higher uncertainty. Our results show that lower personal confidence in forecasts, as measured by uncertainty, corresponds to higher gaps, conceivably because the inflation concept respondents have in mind is less clear.

4.3.3 The Directional Aggregation Gap

The directional aggregation gap is important as it can reveal whether discrepancies between the two measures of inflation expectations wash out *on average*; the absence of a statistically significant directional gap would suggest that the measures are internally consistent. This subsection demonstrates that the directional aggregation gap is positive and substantial, even when taking into account demographic and economic factors.

All plausibly rational aggregations, including expenditure and PCE weights, yield a positive gap, as Table A.7 in the Appendix shows. This result rejects the idea that the reported aggregate reduces to a mental process summing categories by either self-reported expenditure shares or official PCE weights, as the FIRE hypothesis might suggest. While noise may account for $\Lambda_i > 0$ for any individual survey participant, noise cannot explain that the estimated mean for the cross-section is significantly different from zero.

The lowest gap, moreover, is obtained for the non-core aggregation, which is much lower than that for core expectations. This result indicates that non-core expectations—gasoline, energy, and groceries—play an important role in aggregate inflation expectations, in line with the recent literature (e.g., Binder, 2018; D'Acunto et al., 2019; Dietrich, 2023; Trehan, 2011).

As for demographic patterns, the aggregation gap is higher for grocery shoppers, younger respondents, and the less educated. This demographic heterogeneity might point to promising directions for exploring why mean aggregate inflation expectations in major surveys of US households, such as the University of Michigan's Survey of Consumers, have been surprisingly high over the last decade, prior to the COVID pandemic. It raises the possibility that average aggregate inflation expectations for nationally representative samples have been inflated by reporting anomalies among specific demographic segments (such as the young with low education).

Interestingly, our measure of task complexity—inflation uncertainty—is associated with a higher directional aggregation gap. That is, as the task complexity increases, consumers increasingly report aggregate inflation numbers greater than their category-based beliefs. This result might also explain the visual pattern noted in Section 4.2, for Figure 3, where we observe that the discrepancy between aggregate inflation expectations and plausibly rational aggregations is particularly pronounced in April 2021, coinciding with surging realized inflation and inflation news, and following November 2021, which brought a shift to a high-inflation regime.

One might wonder whether another systematic factor explains the aggregation gap in a manner not captured by the above regressors. Respondents, for example, may have a particular variable in mind, the growth rate of which they incorporate into their inflation calculation, but not fully into their consumption basket. This procedure would create an apparent gap between our aggregate and *aggregated* inflation expectations. However, any such variable or omitted category would have to be characterized by extremely high expected inflation rates all the time, both in our low- and high-inflation periods, in order to explain the large, persistent bias—despite representing a small or zero share of the comprehensive PCE basket. While one candidate—house prices—is included in our 11 categories through the housing costs category, another candidate, with no weight in the measured consumption basket, might be asset-price inflation more generally. However, when we consider the return or lagged return on the SP500 as a proxy for asset prices, we do not see any significant relationship with the aggregation gap. This result likely arises from the fact that asset prices are very volatile and often decline (especially during part of our sample period), and hence cannot consistently explain a (positive) gap. Appendix Table A.10 shows this result.

5 Economic Implications

Thus far, we have established that ordinary consumers think differently about aggregate versus category-specific inflation expectations. To gauge the importance of this wedge, we first focus on a central relationship in macroeconomics—the consumption Euler equation. By estimating the Euler equation, we show that aggregated measures of inflation expectations carry superior information for explaining spending plans compared to the conventional aggregate measure of inflation expectations, and that this relative informational advantage increases with the aggregation gap, which—as we show in Section 4.3—relates systematically to heterogeneity in the population, namely socio-demographics and uncertainty. *Aggregated* inflation expectations, moreover, appear to carry additional information regardless of the aggregation chosen as a measure for estimating the Euler equation. At the same time, using *aggregated* expectations implies lower parameter estimates for the intertemporal elasticity of substitution, a key parameter in the main macroeconomic models. In a simple New Keynesian model as in Galí (2015), our preferred estimate of the intertemporal elasticity of substitution implies potentially higher economic volatility.

To show these results, we estimate a consumption Euler equation. We assume consumers follow a standard Euler equation

$$Q_{i,t} = \mathbb{E}_t^i \left[\beta_i \left(\frac{C_{i,t+1}}{C_{i,t}} \right)^{-\frac{1}{\sigma}} \frac{P_t}{P_{t+1}} \right]. \tag{2}$$

This representation of the household Euler equation is widely used in modern macroeconomics (see, for example, Galí, 2015; Woodford, 2003). We adjust the conventional representative-agent version by allowing individual i-specific levels of the discount factor β_i, as well as a nominal interest rate $r_{i,t} = -\log(Q_{i,t})$. $\mathbb{E}_t^i$ represents the expectations operator for respondent i. A log-linearized version of equation (2) is

$$c_{i,t} = \mathbb{E}_t \, c_{i,t+1} - \sigma \left[r_{i,t} - \mathbb{E}_t^i \, \pi_{t+1} - \rho_i \right], \tag{3}$$

where $\pi_t = p_t - p_{t-1}$ denotes the inflation rate. Whereas $\mathbb{E}_t \, c_{i,t+1}$ denotes expected log real consumption, questions Q5 to Q7 of our survey ask respondents about expected expenditure relative to the last month, that is, $\mathbb{E}_t^i \, \Delta s_{i,t+1} = \mathbb{E}_t^i \, (\Delta c_{i,t+1} + \pi_{t+1})$. ρ_i is the log discount factor, $\log \beta_i$. Inserting into equation (3) the expression for the expected change in nominal consumption spending yields a version of the Euler equation that links expected spending to expected inflation

$$\mathbb{E}_t^i \, \Delta s_{i,t+1} - \mathbb{E}_t^i \, \pi_{t+1} = \sigma \left[r_{i,t} - \mathbb{E}_t^i \, \pi_{t+1} - \rho_i \right]. \tag{4}$$

On the left-hand side, we have the expected change in spending, net of the expected rate of inflation. Building on the empirical approach by Crump et al. (2021), we then estimate this equation through the following specification

$$\mathbb{E}_t^i \, \Delta s_{i,t+1} = \gamma_0 + \gamma_1 \, \mathbb{E}_t^i \, \pi_{t+1} + D_i + T_t + \epsilon_{i,t}, \tag{5}$$

where D_i represents demographic fixed effects[16] as well as a control for income expectations, and T_t represents time fixed effects. Including both time and demographic fixed effects relies on the assumption that $r_{i,t} - \rho_i$ may be explained by both variation in time (for example, by changes in the nominal interest rate) and demographic factors, which can impact both the rate of time preference and the nominal interest rate individuals face. The coefficient γ_1 in the estimation equation is equal to $1 - \sigma$ in the model in equation (4).

Estimation of the consumption Euler equation using *aggregated* measures of inflation expectations has clear implications for the estimated intertemporal elasticity of substitution. Estimates based on aggregated expectations all come out lower than the estimate based on aggregate inflation expectations. Table 7 shows the estimation results—using our individual-level, cross-sectional data—for the full array of inflation-expectation measures in the cross-section. The table reports $1 - \hat{\gamma}_1$, which is equal to the intertemporal elasticity of substitution σ. The fourth column gives the R^2 values, the fifth the Akaike information criterion, and the sixth the p-value of a likelihood ratio test, which compares the fit of the respective models to the aggregate inflation-expectation model.

Two results stand out. First, coefficients for inflation expectations are highly significant in all models. Notably, the AIC and the likelihood ratio test suggest improved fit for the *aggregated* measures over aggregate inflation expectations. Moreover, the latter model obtains the lowest R^2. That is, the proportion of variation explained in planned consumption one year ahead is lower for aggregate inflation expectations than for any of the *aggregated* measures; *aggregated* measures of inflation expectations are more informative for future spending plans and can thus better represent effective beliefs. Second, a similar picture emerges when we repeat the estimation for one-year-ahead nondurable and services spending, respectively. The aggregate inflation-expectations model for nondurable spending obtains the highest AIC and the lowest R^2, and *aggregated* models are statistically distinct, according to the likelihood ratio test. Similarly, the aggregate inflation-expectations model for spending on services yields the highest AIC and the lowest R^2. As a robustness exercise to control for possible reporting errors within inflation expectations, Table A.12 in the Appendix reports estimated coefficients for an instrumental variable regression, which takes as an instrument for each measure of inflation expectations the individual mean inflation expectation from the probability distribution question (QDIST, Appendix E).

As with the dispersion of aggregate inflation expectations relative to category-specific expectations, and the absolute aggregation gap between aggregate and *aggregated* inflation expectations, the results can be interpreted through the lens of the model of Bordalo et al. (2022, 2023) on the role of selective recall and memory in belief formation. If recall proves easier in the context of specific consumption categories than in the case of aggregate inflation, for which more interference

[16] Since we rely only on a cross-sectional sample without a panel dimension, we include demographic controls, instead of individual fixed effects.

Table 7: 1 Year Ahead Spending Plans

	$\hat{\sigma} = 1 - \hat{\gamma}_1$	t-stat	R^2	AIC	LR	N
Aggregate	0.960***	7.69	0.057	168157	-	23682
Expenditure	0.821***	15.35	0.083	167499	0.000	23682
Importance	0.786***	16.79	0.087	167390	0.000	23682
PCE	0.788***	15.92	0.085	167439	0.000	23682
Equal	0.777***	16.57	0.088	167381	0.000	23682
Core inflation	0.842***	13.37	0.076	167674	0.000	23682
Non-core inflation	0.874***	14.52	0.076	167679	0.000	23682
Max	0.912***	14.58	0.074	167737	0.000	23682
Second max	0.870***	14.36	0.079	167598	0.000	23682
12-months-ahead nondurable spending						
Aggregate	0.957***	4.96	0.058	33103	-	4696
Expenditure	0.808***	9.34	0.084	32975	0.000	4696
Importance	0.747***	10.67	0.094	32922	0.000	4696
PCE	0.770***	9.71	0.085	32967	0.000	4696
Equal	0.732***	10.27	0.094	32919	0.000	4696
Core inflation	0.845***	8.07	0.073	33027	0.000	4696
Non-core inflation	0.842***	8.64	0.083	32980	0.000	4696
Max	0.907***	6.90	0.071	33039	0.000	4696
Second max	0.851***	8.37	0.086	32964	0.000	4696
12-months-ahead services spending						
Aggregate	0.967***	7.21	0.059	162468	-	23793
Expenditure	0.857***	14.48	0.081	161916	0.000	23793
Importance	0.824***	15.75	0.086	161764	0.000	23793
PCE	0.820***	15.17	0.087	161751	0.000	23793
Equal	0.813***	15.65	0.088	161722	0.000	23793
Core inflation	0.861***	14.07	0.079	161951	0.000	23793
Non-core inflation	0.904***	12.72	0.073	162116	0.000	23793
Max	0.929***	14.06	0.074	162096	0.000	23793
Second max	0.891***	13.82	0.080	161923	0.000	23793

Notes: Estimated Euler equations, relying on various measures of aggregate or aggregated inflation expectations; t statistics in third column, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights to ensure that sample is representative. Data for nondurable spending until 25.02.2021. LR gives the likelihood ratio for the reported aggregate expectations model to minimize the information loss.

occurs, then the measure of aggregate inflation expectations should be noisier, yielding a weaker fit in models that predict spending plans based on aggregate inflation expectations.

Our estimates based on *aggregated* expectations, moreover, imply relatively higher economic volatility than do those based on conventional aggregate expectations. We demonstrate the economic significance of changes in the intertemporal elasticity of substitution in the context of monetary policy, but could also do so in other model contexts, such as forward guidance. We simulate productivity shocks in a simple New Keynesian textbook model, as in Galí (2015),[17] first using an estimate of the intertemporal elasticity of substitution based on an estimation that uses aggregate inflation expectations ($\sigma = 0.960$, see Table 4) and, second, using an elasticity based on an estimation that uses an aggregation of equal weights ($\sigma = 0.777$, see Table 4). We leave all other parameters fixed to highlight the economic importance of the difference in our estimates. We then record two metrics of economic volatility: the variance of inflation and the variance of the output gap. We find large changes: The variance of inflation is 12.2% higher in the simulation that uses the elasticity based on (equal-weight) *aggregated* inflation expectations, relative to simulations using estimates based on conventional aggregate expectations. Similarly, the variance of the output gap is 5.7% higher across theses specifications. In welfare analysis, these changes may be considered costly.

5.1 Superior Information of Aggregated Expectations

The estimation of the Euler equation can also be used to further substantiate the finding that *aggregated* measures of inflation expectations contain superior information for explaining spending plans compared to the conventional aggregate measure of inflation expectations. Specifically, the relative informational advantage of *aggregated* measures of inflation expectations increases with the aggregation gap, which, as shown in Section 4.3, relates systematically to heterogeneity in the population, namely socio-demographics and uncertainty.

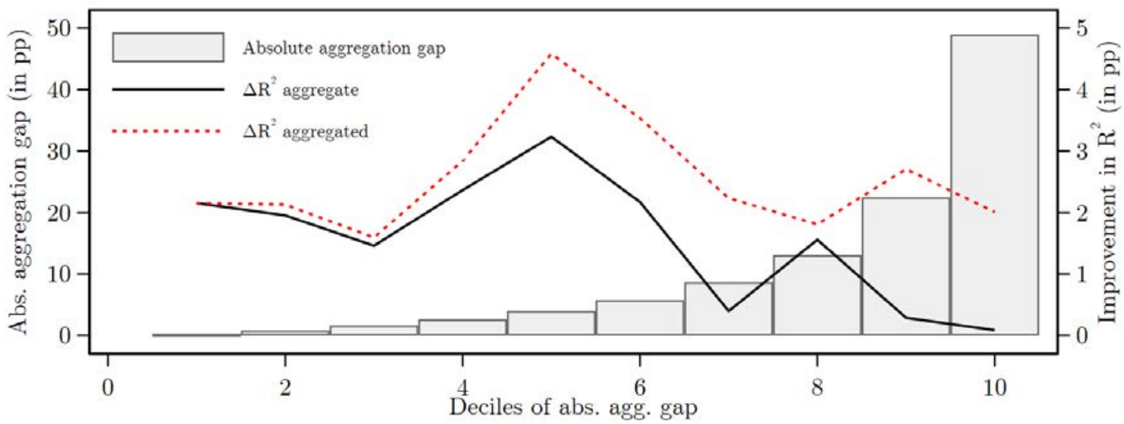
To establish this insight, we split the sample by deciles of the absolute aggregation gap outlined in Section 4.3 and then repeat the estimation of the Euler equation. For each decile, we estimate the Euler equation for planned changes in total spending, as detailed in Equation 5. We are interested in the amount of variation (R^2) of total spending plans explained when we use either the aggregate or *aggregated* measure of inflation expectations as an explanatory variable, benchmarked against a restricted specification with only a constant. In Figure 5, the right axis indicates this difference between a model's R^2 and that of the benchmark. The dashed, red line shows the improvement for *aggregated* inflation expectations, and the black solid line for the conventional, aggregate measure of inflation expectations. As a background, the grey bars in Figure 5 display for each decile (left axis) the mean absolute aggregation gap between the conventional, aggregate inflation expectation and the expenditure-weighted, *aggregated* inflation expectations.[18]

By construction, both lines are always above zero because including an additional independent variable in the estimation improves the share of total variance explained. Moreover, because the measures differ only slightly for respondents with the smallest absolute aggregation gaps, the difference between the two measures in this region is close to zero. However, as the absolute aggregation gap grows, a clear pattern emerges for the conventional measure of aggregate inflation

[17] All parameters are identical to those in chapter 3 of Galí (2015).

[18] We use the expenditure-weighted aggregation as an example. Similar patterns emerge for each of the other aggregations, as Figure A.2 in the appendix shows.

Figure 5: Spending Plan regressions for deciles of the abs. agg. gap



Notes: The figure compares the improvement, for each decile of the aggregation gap, in R^2 achieved by adding a measure of inflation expectations to the estimation in equation (5), relative to an estimation without inflation expectations. Grey bars: Mean aggregation gap (left vertical axis) for each decile on horizontal axis. Black line, right vertical axis: improvement in R^2 by adding the conventional measure of inflation expectations. Red line, right vertical axis: improvement in R^2 by adding the *aggregated* measure of inflation expectations (expenditure-weighted).

expectations: The improvement in R^2 from including the measure into the Euler-equation regression declines substantially, approaching zero for those with the largest aggregation gaps. When using the expenditure-weighted, *aggregated* expectations in the regression, on the other hand, improvement in R^2 does not drop to zero, and the R^2 is also consistently higher than that obtained with aggregate inflation expectations. Against the backdrop of socio-demographic heterogeneity and uncertainty associated with the aggregation gap, this relative informational advantage of aggregated inflation expectations is highly relevant—especially for policymakers—for measuring inflation expectations effectively across diverse parts of the population.

6 Conclusion

Motivated by evidence that consumers struggle to grasp the concept of inflation, we introduce an alternative approach to measuring consumer inflation expectations. Rather than asking consumers outright about aggregate inflation expectations, as conventional measures do, we elicit their expectations for the underlying parts, the full range of personal consumption expenditures, which we then aggregate. Our results indicate that *aggregated* measures, across all of our aggregation procedures, yield more information about consumer spending plans compared to the conventional measure. We obtained this finding in a daily consumer survey of close to 60,000 respondents over two years, extending from the low-inflation environment of the COVID pandemic to the 2021 inflation surge.

In addition, four striking facts stand out. The first is that aggregate inflation expectations are higher than inflation expectations for any single category. For a representative agent, this finding rules out a linear mapping (with non-negative weights) of the category-specific expectations into aggregate inflation expectations. Moreover, disagreement among respondents over aggregate inflation expectations is higher than that over any category. Second, *aggregated* inflation expectations, unsurprisingly with the exception of the max operators, are lower than aggregate expectations—the whole is greater than the sum of the parts. *Aggregated* inflation expectations are also less dispersed. Third, the respondent-specific gap between aggregate and *aggregated* inflation expectations rises with the subjective complexity of the aggregate inflation concept and correlates in a meaningful way with socioeconomic characteristics such as education. Fourth, at higher levels of the respondent-specific gap between *aggregated* and aggregated inflation expectations, the information about consumer spending plans provided by *aggregated* expectations is higher in both absolute and relative terms than that provided by aggregate expectations. In other words, when the two measures differ substantially, the *aggregated* measure become more informative.

Effective inflation expectations—the expectations on which consumers act—therefore, appear *not* best represented by explicit, conventionally reported aggregate inflation expectations, but by aggregations of category-specific inflation expectations. We find that this holds true irrespective of socio-demographic heterogeneity and uncertainty faced by respondents.

Clearly, alternative procedures for decomposing the aggregate inflation concept are worth exploring, as are alternative aggregation mechanisms and measures of consumer planning and behavior. We regard our paper as a first step in this direction, a proof of concept that lays out foundations for future work aimed at improving the measurement of a key macroeconomic aggregate, a vital input to monetary policy that is notoriously difficult to capture.

References

Armantier, Olivier, Giorgio Topa, Wilbert van der Klaauw, and Basit Zafar (2016). *How do people revise their inflation expectations?* Liberty Street Economics, August 22, 2016.

Armantier, Olivier, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar (2017). "An overview of the Survey of Consumer Expectations". *Economic Policy Review* (23-2), 51–72.

Arora, Vipin, Pedro Gomis-Porqueras, and Shuping Shi (2013). "The divergence between core and headline inflation: implications for consumers' inflation expectations". *Journal of Macroeconomics 38* (B), 497–504. DOI: 10.1016/j.jmacro.2013.07.006.

Bachmann, Rüdiger, Tim O. Berg, and Eric R. Sims (2015). "Inflation expectations and readiness to spend: cross-sectional evidence". *American Economic Journal: Economic Policy 7* (1), 1–35. DOI: 10.1257/pol.20130292.

Ben-David, Itzhak, Elyas Fermand, Camelia M. Kuhnen, and Geng Li (2018). "Expectations uncertainty and household economic behaviors". Working Paper 25336. National Bureau of Economic Research. DOI: 10.3386/w25336.

Bertrand, Marianne and Sendhil Mullainathan (2001). "Do people mean what they say? implications for subjective survey data". *The American Economic Review 91* (2), 67–72. DOI: 10.1257/aer.91.2.67.

Binder, Carola and Christos Makridis (2022). "Stuck in the seventies: gas prices and consumer sentiment". *The Review of Economics and Statistics 104* (2), 1–13. DOI: 10.1162/rest_a_00944.

Binder, Carola Conces (2018). "Inflation expectations and the price at the pump". *Journal of Macroeconomics 58*, 1–18. DOI: 10.1016/j.jmacro.2018.08.006.

Bishop, Yvonne, Stephen Fienberg, and Paul Holland (1975). *Discrete multivariate analysis: theory and practice.* DOI: 10.1007/978-0-387-72806-3.

Blanchard, Olivier J. (1986). "The wage price spiral". *The Quarterly Journal of Economics 101* (3), 543–565.

Bordalo, Pedro, Giovanni Burro, Katie Coffman, Nicola Gennaioli, and Andrei Shleifer (2022). "Imagining the future: memory, simulation and beliefs about covid". Mimeo 30353. National Bureau of Economic Research. DOI: 10.3386/w30353.

Bordalo, Pedro, John Conlon, Nicola Gennaioli, Spencer Y. Kwon, and Andrei Shleifer (2023). "Memory and probability". *Quarterly Journal of Economics 138* (1), 265–311.

Bruine de Bruin, Wändi, Wilbert van der Klaauw, and Giorgio Topa (2011). "Expectations of inflation: the biasing effect of thoughts about specific prices". *Journal of Economic Psychology 32* (5), 834–845. DOI: 10.1016/j.joep.2011.07.002.

Bruine de Bruin, Wändi et al. (2010). "Expectations of inflation: the role of demographic variables, expectation formation, and financial literacy". *The Journal of Consumer Affairs 44* (2), 381–402. DOI: 10.1111/j.1745-6606.2010.01174.x.

Cavallo, Alberto, Guillermo Cruces, and Ricardo Perez-Truglia (2017). "Inflation expectations, learning, and supermarket prices: evidence from survey experiments". *American Economic Journal: Macroeconomics 9* (3), 1–35. DOI: 10.1257/mac.20150147.

Clements, Michael P. (2014). "Probability distributions or point predictions? survey forecasts of US output growth and inflation". *International Journal of Forecasting 30* (1), 99–117. DOI: 10.1016/j.ijforecast.2013.07.010.

Coibion, Olivier and Yuriy Gorodnichenko (2012). "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy 120* (1), 116–159. DOI: 10.1086/665662.

——— (2015). "Is the Phillips Curve alive and well after all? Inflation expectations and the missing disinflation". *American Economic Journal: Macroeconomics 7* (1), 197–232. DOI: 10.1257/mac.20130306.

Crump, Richard, Stefano Eusepi, Andrea Tambalotti, and Giorgio Topa (2021). "Subjective intertemporal substitution". *Journal of Monetary Economics 126*. DOI: 10.1016/j.jmoneco.2021.11.008.

Crump, Richard K., Stefano Eusepi, Andrea Tambalotti, and Giorgio Topa (2022). "Subjective intertemporal substitution". *Journal of Monetary Economics 126*, 118–133. DOI: 10.1016/j.jmoneco.2021.11.008.

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber (2019). "Cognitive abilities and inflation expectations". *American Economic Review Papers & Proceedings 109*, 562–566. DOI: 10.1257/pandp.20191050.

——— (2022). "IQ, expectations, and choice". *Review of Economic Studies*. forthcoming. DOI: 10.2139/ssrn.3451486.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber (2021). "Exposure to grocery prices and inflation expectations". *Journal of Political Economy 129* (5), 1615–1639. DOI: 10.1086/713192.

D'Acunto, Francesco and Michael Weber (2022). "Memory and beliefs: evidence from the field".

Deaton, Angus (2019). *The analysis of household surveys : a microeconometric approach to development policy*. Washington, DC: World Bank. DOI: 10.1596/978-1-4648-1331-3.

Dietrich, Alexander (2023). "Consumption categories, household attention and inflation expectations: implications for optimal monetary policy". University of Tübingen Working Papers in Business and Economics 157.

Dietrich, Alexander M., Keith Kuester, Gernot J. Müller, and Raphael Schoenle (2022). "News and uncertainty about covid-19: survey evidence and short-run economic impact". *Journal of Monetary Economics 129*, S35–S51. DOI: https://doi.org/10.1016/j.jmoneco.2022.02.004.

Engelberg, Joseph, Charles F. Manski, and Jared Williams (2009). "Comparing the point predictions and subjective probability distributions of professional forecasters". *Journal of Business & Economic Statistics 27* (1), 30–41. DOI: 10.1198/jbes.2009.0003.

Fischhoff, B. and S. B. Broomell (2020). "Judgment and decision making". *Annual Review of Psychology 71*, 331–355. DOI: 10.1146/annurev-psych-010419-050747.

Gabaix, Xavier and Thomas Graeber (2023). "The complexity of economic decisions". Working paper.

Galí, Jordi (2015). *Monetary policy, inflation, and the business cycle: an introduction to the New Keynesian framework and its applications*. Second. Princeton University Press.

Georganas, Sotiris, Paul J. Healy, and Nan Li (2014). "Frequency bias in consumers perceptions of inflation: an experimental study". *European Economic Review 67*, 144–158. DOI: https://doi.org/10.1016/j.euroecorev.2014.01.014.

Huber, Peter J. (1964). "Robust Estimation of a Location Parameter". *The Annals of Mathematical Statistics 35* (1), 73 –101. DOI: 10.1214/aoms/1177703732.

Hurd, Michael D. and Susann Rohwedder (2008). "Methodological innovations in collecting spending data: the HRS consumption and activities mail survey". Working paper WR-646. Santa Monica, CA: RAND Corporation.

——— (2009). "Methodological innovations in collecting spending data: the HRS consumption and activities mail survey". *Fiscal Studies 30* (3-4), 435–459. DOI: 10.1111/j.1475-5890.2009.00103.x.

Hurd, Michael D. and Susann Rohwedder (2012). “Measuring total household spending in a monthly internet survey: evidence from the American life panel”. Working paper WR-939. Santa Monica, CA: RAND Corporation. DOI: 10.7249/WR939.

Idel, Martin (2016). “A review of matrix scaling and Sinkhorn’s normal form for matrices and positive maps”. Mimeo arXiv:1609.06349. arXiv. DOI: 10.48550/arXiv.1609.06349.

Jonung, Lars (1981). “Perceived and expected rates of inflation in sweden”. *The American Economic Review 71* (5), 961–968.

Knotek, E. S. et al. (2020). “Consumers and COVID-19: a real-time survey”. *Economic Commentary* (2020-08). DOI: 10.26509/frbc-ec-202008.

Kuchler, Theresa and Basit Zafar (2019). “Personal experiences and expectations about aggregate outcomes”. *Journal of Finance 74* (5), 2491–2542. DOI: 10.1111/jofi.12819.

Lorenzoni, Guido and Iván Werning (2023). “Wage price spirals”. Mimeo.

Malmendier, Ulrike and Stefan Nagel (2015). “Learning from inflation experiences”. *The Quarterly Journal of Economics 131* (1), 53–87. DOI: 10.1093/qje/qjv037.

Menon, G. (1997). “Are the parts better than the whole? the effects of decompositional questions on judgments of frequent behaviors”. *Journal of Marketing Research 34* (3), 335–346.

Niu, X. and N. Harvey (2023). “Are lay expectations of inflation based on recall of specific prices? if so, how and under what conditions?” *Journal of Economic Psychology.* DOI: 10.1016/j.joep.2023.102662.

Payne, J., J. Bettman, and E. Johnson (1993). *The adaptive decision maker.* Cambridge University Press. DOI: 10.1017/CBO9781139173933.

Powell, Jerome (2021). *Transcript of chair powell’s press conference, september 21, 2021.* speech.

Qualtrics (2019). *ESOMAR 28: 28 questions to help buyers of online samples.*

Ryngaert, Jane M. (2022). “Inflation disasters and consumption”. *Journal of Monetary Economics 129*, S67–S81. DOI: https://doi.org/10.1016/j.jmoneco.2022.03.002.

Shanmugam, Ramalingam (2008). “Correlation between the sample mean and sample variance”. *Journal of Modern Applied Statistical Methods 7* (2).

Stanisławska, Ewa, Maritta Paloviita, and Tomasz Łyziak (2021). "Consumer inflation views: micro-level inconsistencies and macro-level measures". *Economics Letters 206*, 110004. DOI: `10.1016/j.econlet.2021.110004`.

Trehan, Bharat (2011). "Household inflation expectations and the price of oil: it's déjà vu all over again". *FRBSF Economic Letter.*

Tversky, Amos and Daniel Kahneman (1974). "Judgment under uncertainty: heuristics and biases: biases in judgments reveal some heuristics of thinking under uncertainty". *Science 185* (4157), 1124–1131. DOI: `10.1126/science.185.4157.1124`.

Weber, Michael, Francesco D'Acunto, Yuriy Gorodnichenko, and Olivier Coibion (2022). "The subjective inflation expectations of households and firms: measurement, determinants, and implications". *Journal of Economic Perspectives 36* (3), 157–84. DOI: `10.1257/jep.36.3.157`.

Winter, Joachim (2004). "Response bias in survey-based measures of household consumption". *Economics Bulletin 3* (9), 1–12.

Woodford, Michael (2003). *Interest and prices: foundations of a theory of monetary policy.* Princeton University Press.

Zhang, Lingyun (2007). "Sample mean and sample variance: their covariance and their (in)dependence". *The American Statistician 61* (2), 159–160.

A Additional Tables

A.1 Socio-Demographic Summary Statistics

Table A.1: Summary Statistics - Mean Demographics

	Gender		Grocery		Education			Income	
	Female	Male	Yes	No	High	Low	High	Middle	Low
Aggregate expectation	7.31	6.01	6.70	5.87	6.11	7.46	7.65	6.36	6.97
Category expectations									
Motor vehicles	5.68	5.41	5.54	5.67	5.71	5.35	5.82	5.51	5.37
Recreational goods	4.45	3.72	4.01	4.04	4.23	3.82	4.27	4.18	3.85
Other durable goods	4.32	3.96	4.15	3.80	4.32	3.94	4.62	4.21	3.90
Food and beverages	5.79	4.88	5.28	5.60	5.39	5.27	5.60	5.52	5.25
Gasoline	5.78	4.96	5.28	5.74	5.40	5.35	5.39	5.72	5.27
Other nondurable	4.41	3.95	4.20	3.97	4.33	4.05	4.58	4.28	3.94
Housing and util.	5.28	4.66	4.99	4.94	5.22	4.77	5.30	5.34	4.69
Health care	4.15	3.90	4.03	3.95	4.21	3.81	4.53	4.13	3.70
Transportation	5.26	4.46	4.87	4.73	4.89	4.82	4.78	5.09	4.87
Food services	5.02	4.57	4.81	4.92	5.05	4.56	5.23	4.87	4.52
Other services	4.58	4.07	4.37	4.22	4.39	4.27	4.56	4.51	4.23
Aggregated expectations									
Plausibly rational aggregation									
Expenditure weights	5.40	4.63	4.99	5.14	5.08	4.95	5.22	5.22	4.90
Importance weights	5.02	4.28	4.61	4.77	4.78	4.51	4.88	4.81	4.49
PCE weights	4.90	4.15	4.48	4.53	4.63	4.41	4.75	4.65	4.36
Behavioral aggregation									
Equal weights	4.89	4.21	4.52	4.60	4.69	4.42	4.75	4.65	4.41
Core inflation	5.10	4.47	4.77	4.71	4.88	4.68	4.96	4.92	4.61
Non-core inflation	6.30	5.25	5.72	6.27	5.68	5.86	5.83	6.00	5.85
Max	12.54	10.58	11.29	12.75	11.29	11.76	11.52	11.48	11.94
Second max	7.64	6.51	6.97	7.57	7.06	7.01	7.06	7.21	7.16

Notes: This table presents summary statistics on the Huber-robust and survey-weighted mean on expectations across demographics. “Grocery” refers to weather a respondent indicates to be the main grocery shopper of its household.

Table A.2: Summary Statistics - Standard Deviation Demographics

	Gender		Grocery		Education		Income		
	Female	Male	Yes	No	High	Low	High	Middle	Low
Aggregate expectation	10.18	5.86	7.88	7.03	5.98	10.22	7.61	6.22	9.67
Category expectations									
Motor vehicles	6.51	5.26	5.92	5.98	5.44	6.22	5.34	5.38	6.29
Recreational goods	6.83	5.62	6.34	5.85	5.67	6.79	5.57	5.61	6.90
Other durable goods	6.73	5.43	6.21	5.51	5.49	6.56	5.57	5.45	6.82
Food and beverages	7.06	5.86	6.49	6.08	5.83	6.89	5.93	5.76	6.99
Gasoline	7.97	7.24	7.49	8.30	7.47	7.65	7.33	7.56	7.77
Other nondurable	6.55	5.33	6.03	5.67	5.34	6.50	5.31	5.43	6.48
Housing and util.	6.93	5.76	6.47	6.13	5.77	6.94	5.79	5.90	6.92
Health care	6.99	5.90	6.51	6.21	6.09	6.80	6.15	5.99	6.83
Transportation	6.80	5.50	6.21	5.97	5.63	6.69	5.55	5.60	6.83
Food services	6.87	5.85	6.49	6.03	6.03	6.70	6.04	5.80	6.88
Other services	6.20	4.98	5.70	5.31	5.01	6.12	5.11	5.04	6.13
Aggregated expectations									
Plausibly rational aggregation									
Expenditure weights	5.87	4.64	5.29	4.92	4.74	5.72	4.58	4.81	5.78
Importance weights	5.16	4.20	4.72	4.40	4.40	4.83	4.16	4.42	4.98
PCE weights	5.11	4.06	4.63	4.23	4.23	4.84	4.10	4.23	4.91
Behavioral aggregation									
Equal weights	5.00	4.06	4.59	4.26	4.28	4.68	4.00	4.29	4.84
Core inflation	5.76	4.56	5.22	4.76	4.66	5.69	4.40	4.70	5.67
Non-core inflation	6.78	5.65	6.26	6.02	5.78	6.64	5.87	5.77	6.77
Max	9.57	7.62	8.41	8.95	7.99	8.98	7.94	8.11	9.30
Second max	7.13	5.82	6.43	6.51	5.87	6.90	5.81	6.05	7.11

Notes: This table presents summary statistics on the Huber-robust and survey-weighted standard deviation on expectations across demographics. "Grocery" refers to weather a respondent indicates to be the main grocery shopper of its household.

Table A.3: Summary Statistics - Age Groups

	Mean				Disagreement (SD)			
	18-34	35-44	45-54	above 55	18-34	35-44	45-54	above 55
Aggregate expectation	7.95	9.00	8.42	5.74	11.64	11.63	9.62	4.36
Category expectations								
Motor vehicles	4.62	5.89	5.98	6.35	6.38	6.26	6.04	4.97
Recreational goods	2.47	4.11	4.81	5.28	7.15	6.89	6.25	4.50
Other durable goods	2.82	4.25	4.77	5.23	6.76	6.87	6.13	4.64
Food and beverages	3.80	5.41	6.32	7.06	6.99	7.18	6.78	5.19
Gasoline	3.81	5.16	6.50	7.60	7.42	7.32	7.63	7.98
Other nondurable	2.85	4.26	5.11	5.25	7.00	6.57	6.00	4.34
Housing and util.	3.66	4.77	5.96	6.30	7.08	7.09	6.48	5.09
Health care	2.61	4.17	4.59	5.26	7.02	6.79	6.15	5.30
Transportation	3.51	4.80	5.59	6.27	6.86	6.74	6.41	4.95
Food services	3.06	4.65	5.55	6.62	6.94	6.77	6.32	5.32
Other services	3.37	4.24	5.06	5.20	6.35	6.26	5.56	4.08
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	3.66	4.95	6.02	6.47	5.11	5.28	5.31	4.65
Importance weights	3.05	4.51	5.61	6.46	3.81	4.60	4.82	4.53
PCE weights	3.10	4.36	5.37	6.17	4.02	4.75	4.66	4.14
Behavioral aggregation								
Equal weights	2.99	4.42	5.42	6.32	3.68	4.54	4.69	4.32
Core inflation	3.63	4.68	5.64	5.96	5.32	5.26	5.13	4.42
Non-core inflation	4.16	5.66	6.73	7.36	6.36	6.43	6.41	5.66
Max	10.63	11.61	12.23	13.22	8.09	8.51	9.19	8.89
Second max	6.13	7.27	7.47	8.50	6.65	6.64	7.01	5.88

Notes: This table presents summary statistics on the Huber-robust and survey-weighted mean and standard deviation on expectations across age groups.

A.2 Robustness

We applied a number of robustness exercises in order to ensure that results, especially as reported in section 3, are not subject to the wording of questions. Table A.4 reports results on an additional survey that asked respondents about more specific concepts of aggregate inflation, such as PCE or CPI price-index inflation. Table A.5 reports results on a subset of respondents from the main survey that were asked about category-specific inflation expectations before the question on aggregate expectations. In addition, in Table A.6 we report results on a replication study with funds managers.

Table A.4: Robustness - Wording of Aggregate Inflation Question

	Mean	Std. Dev. (Disagreement)
Aggregate expectation		
"Inflation"	9.54	8.20
"PCE Inflation"	9.86	7.35
"CPI Inflation"	10.17	7.08
Category expectations		
Motor vehicles	6.60	5.13
Recreational goods	5.05	7.01
Other durable goods	5.31	6.50
Food and beverages	7.00	7.19
Gasoline	5.62	7.87
Other nondurable goods	5.13	6.67
Housing and utilities	6.01	6.12
Health care	5.52	6.59
Transportation services	6.28	6.92
Food services	6.62	5.93
Other services	5.55	6.49
Aggregated expectation		
Plausibly rational aggregation		
Expenditure	6.28	5.78
Importance	5.99	5.49
PCE	5.97	5.39
Behavioral aggregation		
Equal	6.01	5.39
Core inflation	6.02	5.80
Non-core inflation	7.07	7.82
Max	12.68	9.00
Second max	8.63	6.64

Notes: This table presents the cross section mean and disagreement for a replication survey conducted between July 5 and July 28, 2022. The survey uses three different wordings ("Inflation", "PCE Inflation", "CPI Inflation") for aggregate inflation, randomly assigned to respondents. Other questions are identical to the main survey. Answers to "CPI Inflation" are significantly higher, after controlling for socio-demographic factors (t-stat=2.02).

Table A.5: Robustness - Ordering of Inflation Questions

	Aggregate First	Category First	
	Mean		p-val
Aggregate expectation	6.52	6.86	0.086
Category expectations			
Motor vehicles	6.38	6.35	0.707
Recreational goods	4.80	4.35	0.008
Other durable goods	4.73	4.83	0.357
Food and beverages	6.03	5.60	0.025
Gasoline	5.70	5.63	0.299
Other nondurable goods	4.84	4.71	0.855
Housing and utilities	4.92	5.17	0.297
Health care	4.55	4.36	0.320
Transportation services	5.14	5.21	0.982
Food services	5.30	4.97	0.075
Other services	4.53	4.59	0.896
Aggregated expectation			
Plausibly rational aggregation			
Expenditure	5.38	5.19	0.175
Importance	5.07	5.04	0.794
PCE	4.95	4.92	0.935
Behavioral aggregation			
Equal	5.08	4.97	0.539
Core inflation	5.11	4.88	0.143
Non-core inflation	5.91	5.91	0.047
Max	11.1	10.52	0.588
Second max	7.5	7.16	0.620

Notes: This table presents the cross-sectional mean for survey respondents asked between January 19 and March 03, 2022, dependent on weather they received the question on aggregate inflation expectations first or after the question on category inflation expectations. The third column shows the p-value for a difference in cross-sectional Huber-robust and survey-weighted means, controlling for time and demographic fixed effects.

Table A.6: Replication Study - Financial Experts Panel

	Mean	Std. Dev. (Disagreement)	N
Aggregate expectations	5.81	1.38	35
Category expectations			
Motor vehicles	1.77	5.11	35
Recreational goods	3.13	3.11	35
Other durable goods	3.54	3.05	35
Food and beverages	6.04	4.22	35
Gasoline	2.99	6.41	35
Other nondurable goods	3.91	1.72	35
Housing and utilities	5.38	3.69	35
Health care	2.88	5.33	35
Transportation services	4.18	5.20	35
Food services	5.45	5.13	35
Other services	3.68	3.01	35
Aggregated expectation			
Plausibly rational aggregation			
Expenditure	4.93	3.80	34
PCE	4.08	3.68	35
Behavioral aggregation			
Equal	3.63	3.49	35
Core inflation	4.85	3.89	34
Non-core inflation	5.39	4.51	34
Max	7.89	3.26	35
Second max	5.45	2.98	35

Notes: This table presents the cross section mean and disagreement (Huber-robust estimates) for a replication survey conducted between November 4 and November 9 2022, with a group of financial market experts (fund managers).

A.3 Aggregation Gap

A.3.1 The Directional Aggregation Gap

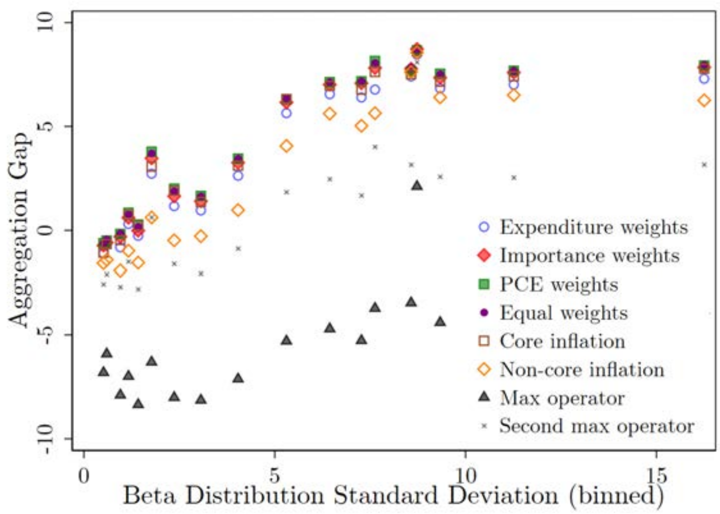
In this section we report results for the aggregation gap between aggregate and aggregated inflation expectations per respondent, without taking the absolute value of the gap. Table A.7 displays the mean aggregation gap, as compared to the mean absolute aggregation gap (see also Table **??** in the main text). While the absolute aggregation gap provides a measure of the discrepancy between aggregate and aggregated inflation expectations irrespective of sign, the aggregation gap indicates the direction of the gap.

Table A.7: Summary Statistics

	Mean Aggregation Gap (Λ_i)	Mean Absolute Aggregation Gap abs(Λ_i)
Plausibly rational aggregation		
Expenditure weights	1.33***	5.64***
Importance weights	1.55***	5.49***
PCE weights	1.62***	5.33***
Behavioral aggregation		
Equal weights	1.64***	5.34***
Core inflation	1.62***	5.67***
Non-core inflation	0.68***	6.06***
Max	-3.97***	9.09***
Second max	-0.42***	6.51***

Notes: This table presents Huber-robust and survey-weighted estimates for the mean aggregation gap and mean root square aggregation gap; Stars: significance level of a t-test that numbers are different from zero. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$;

Figure A.1: The Aggregation Gap and Uncertainty



Notes: The figure shows the correlation between the aggregation gap Λ_i and the subjective uncertainty over aggregate inflation expectations obtained via a beta distribution over a probabilistic question.

A.3.2 Aggregation Gap - Socio-Demographic Effects

Tables A.8 and A.9 link the (absolute) aggregation gap to socio-demographic factors as well as respondents' grocery shopping habits.

Table A.8: Demographics and the Aggregation Gap

	(1) Expenditure	(2) Importance	(3) PCE	(4) Equal	(5) Core	(6) Non-core	(7) 1^{st}max	(8) 2^{nd} max
Female	-0.0288	-0.0639	-0.0523	-0.0378	0.146	-0.239**	-0.808***	-0.349***
	(-0.37)	(-0.86)	(-0.72)	(-0.52)	(1.87)	(-2.68)	(-8.04)	(-4.12)
Grocery Shopper	0.895***	0.909***	0.898***	0.867***	0.825***	0.908***	1.721***	1.293***
	(7.57)	(8.07)	(8.08)	(7.85)	(6.95)	(6.64)	(11.00)	(9.84)
35 to 44 years	-0.219*	-0.376***	-0.289**	-0.377***	-0.148	-0.661***	-0.111	-0.0852
	(-2.10)	(-3.75)	(-2.92)	(-3.83)	(-1.39)	(-5.37)	(-0.86)	(-0.76)
45 to 54 years	-1.174***	-1.257***	-1.182***	-1.212***	-1.013***	-1.695***	-0.774***	-0.744***
	(-9.57)	(-10.59)	(-10.02)	(-10.40)	(-8.13)	(-11.99)	(-4.90)	(-5.50)
above 55 years	-2.798***	-3.047***	-2.749***	-2.952***	-2.453***	-3.802***	-3.847***	-3.032***
	(-33.48)	(-37.95)	(-34.90)	(-37.53)	(-29.04)	(-38.00)	(-34.77)	(-32.88)
High Educated	-0.673***	-0.705***	-0.722***	-0.711***	-0.699***	-0.581***	-0.873***	-0.830***
	(-7.80)	(-8.49)	(-8.83)	(-8.73)	(-7.99)	(-5.84)	(-7.61)	(-8.68)
Middle Income	0.0624	0.142	0.188*	0.175*	0.112	-0.201*	0.110	0.188
	(0.70)	(1.64)	(2.20)	(2.07)	(1.25)	(-1.99)	(0.93)	(1.88)
High Income	0.110	0.0624	0.0902	0.103	0.0815	0.0826	0.433**	0.205
	(0.96)	(0.57)	(0.83)	(0.96)	(0.71)	(0.64)	(2.88)	(1.64)
Constant	2.098***	2.430***	2.358***	2.474***	2.201***	2.307***	-3.242***	0.198
	(15.49)	(18.74)	(18.43)	(19.42)	(16.08)	(14.36)	(-18.47)	(1.33)
N	54453	54183	52857	54205	53152	49216	55995	55009
r2	0.0346	0.0430	0.0389	0.0421	0.0289	0.0471	0.0377	0.0353

Notes: This table presents Huber-robust and survey-weighted regressions of the aggregation gap on several demographic characteristics. The headers for each column represent the aggregation mechanism. For details on aggregated expectations, see Table 4. t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.9: Demographics and the Absolute Aggregation Gap

	(1) Expenditure	(2) Importance	(3) PCE	(4) Equal	(5) Core	(6) Non-core	(7) 1^{st}max	(8) 2^{nd} max
Female	0.764***	0.658***	0.734***	0.687***	0.819***	0.602***	0.923***	0.593***
	(12.40)	(11.41)	(12.74)	(12.09)	(12.92)	(8.52)	(9.64)	(8.47)
Grocery Shopper	-0.0469	0.00631	-0.0275	-0.0189	-0.0474	-0.227*	-0.551***	-0.230*
	(-0.51)	(0.07)	(-0.32)	(-0.22)	(-0.50)	(-2.16)	(-3.74)	(-2.20)
35 to 44 years	-0.276***	-0.272***	-0.206**	-0.272***	-0.211*	-0.408***	-0.321**	-0.384***
	(-3.31)	(-3.45)	(-2.62)	(-3.50)	(-2.43)	(-4.15)	(-2.59)	(-4.12)
45 to 54 years	-0.906***	-0.908***	-0.850***	-0.912***	-0.947***	-1.210***	-0.706***	-0.692***
	(-9.34)	(-9.98)	(-9.26)	(-10.14)	(-9.44)	(-10.98)	(-4.86)	(-6.34)
above 55 years	-2.001***	-1.993***	-2.068***	-2.022***	-2.140***	-2.007***	-0.721***	-1.689***
	(-30.31)	(-32.19)	(-33.64)	(-33.18)	(-31.62)	(-25.44)	(-6.90)	(-22.47)
High Educated	-0.591***	-0.624***	-0.724***	-0.663***	-0.651***	-0.477***	-0.164	-0.605***
	(-8.54)	(-9.65)	(-11.19)	(-10.43)	(-9.15)	(-6.01)	(-1.49)	(-7.65)
Middle Income	-0.284***	-0.205**	-0.192**	-0.217**	-0.284***	-0.310***	-0.376***	-0.189*
	(-3.97)	(-3.05)	(-2.85)	(-3.28)	(-3.87)	(-3.86)	(-3.41)	(-2.32)
High Income	-0.0311	-0.0785	-0.0265	-0.0668	-0.0254	-0.0936	0.218	0.0174
	(-0.34)	(-0.93)	(-0.31)	(-0.80)	(-0.27)	(-0.92)	(1.52)	(0.17)
Constant	6.492***	6.314***	6.201***	6.258***	6.602***	7.363***	9.387***	7.244***
	(60.79)	(63.81)	(62.91)	(63.93)	(60.17)	(59.05)	(56.49)	(60.23)
N	54348	54034	52846	54172	53175	49148	57473	55146
r2	0.0368	0.0401	0.0465	0.0430	0.0415	0.0277	0.00469	0.0197

Notes: This table presents Huber-robust and survey-weighted regressions of the absolute aggregation gap on several demographic characteristics. The headers for each column represent the aggregation mechanism. For details on aggregated expectations, see Table 4. t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.10: Aggregation Gap and Asset Price Changes

Absolute aggregation gap (expenditure weighted)	(1) $\mid \Lambda_i^{exp} \mid$	(2) $\mid \Lambda_i^{exp} \mid$	(3) $\mid \Lambda_i^{exp} \mid$	(4) $\mid \Lambda_i^{exp} \mid$
$\Delta_{t,t-1}$ ln(SP500)	-1.176			
	(-0.54)			
$\Delta_{t-1,t-2}$ ln(SP500)		0.560		
		(0.56)		
$\Delta_{t,t-365}$ ln(SP500)			0.0430	
			(0.53)	
$\Delta_{t-1,t-366}$ ln(SP500)				0.0421
				(0.55)
Constant	6.995***	7.225***	7.398***	7.422***
	(9.88)	(15.60)	(14.23)	(14.81)
Time + Demographic FE	yes	yes	yes	yes
N	35911	35952	36641	36864
r2	0.0752	0.0729	0.0751	0.0748

Directed aggregation gap (expenditure weighted)	(5) Λ_i^{exp}	(6) Λ_i^{exp}	(7) Λ_i^{exp}	(8) Λ_i^{exp}
$\Delta_{t,t-1}$ ln(SP500)	0.811			
	(0.29)			
$\Delta_{t-1,t-2}$ ln(SP500)		-0.407		
		(-0.31)		
$\Delta_{t,t-365}$ ln(SP500)			-0.0301	
			(-0.28)	
$\Delta_{t-1,t-366}$ ln(SP500)				-0.0298
				(-0.30)
Constant	2.285*	2.076**	2.036**	1.958**
	(2.19)	(3.04)	(3.07)	(3.02)
Time + Demographic FE	yes	yes	yes	yes
N	35899	35946	36630	36862
r2	0.0673	0.0672	0.0678	0.0672

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: This table presents Huber-robust and survey-weighted regressions of the absolute aggregation gap (expenditure weighted) on daily asset price changes (SP500). Time (daily) and demographic fixed effects used. For details on aggregated expectations, see Table 4. t statistics in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

A.4 Aggregate vs. Aggregated Inflation Expectations

Table A.11: Aggregate vs. Aggregated Inflation Expectations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure	0.557***								0.116
	(34.72)								(1.61)
Importance		0.650***							-0.220*
		(37.02)							(-2.15)
PCE			0.639***						-0.378***
			(35.12)						(-4.48)
Equal				0.685***					1.042***
				(38.07)					(8.44)
Core Inflation					0.514***				-0.0122
					(31.32)				(-0.22)
Non-core Inflation						0.389***			-0.0429
						(30.63)			(-1.40)
Max							0.314***		0.126***
							(33.93)		(7.01)
Second max								0.411***	0.00485
								(31.63)	(0.17)
Constant	7.589***	7.174***	7.370***	7.067***	8.123***	8.464***	6.446***	7.440***	6.187***
	(48.97)	(45.39)	(46.10)	(44.83)	(53.66)	(55.85)	(38.32)	(44.68)	(37.39)
N	50701	50701	50701	50701	50701	50701	50701	50701	50701
R2	0.0721	0.0807	0.0756	0.0840	0.0621	0.0560	0.0668	0.0652	0.0906
AIC	441745	441270	441551	441088	442285	442618	442033	442117	440735

Notes: The table presents estimates on a micro level for a linear regression of reported aggregate on one (column 1 to 8) or multiple (column 9) aggregated, category-based measures of inflation. t statistics in parentheses, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights to ensure that sample is representative.

A.5 Spending Plans - Instrumental Variable Regression

Table A.12 reports results on an estimation of the Euler equation, using measures of inflation expectations in first column instrumented with the mean inflation expectation from the distribution question (QDIST, see Appendix E).

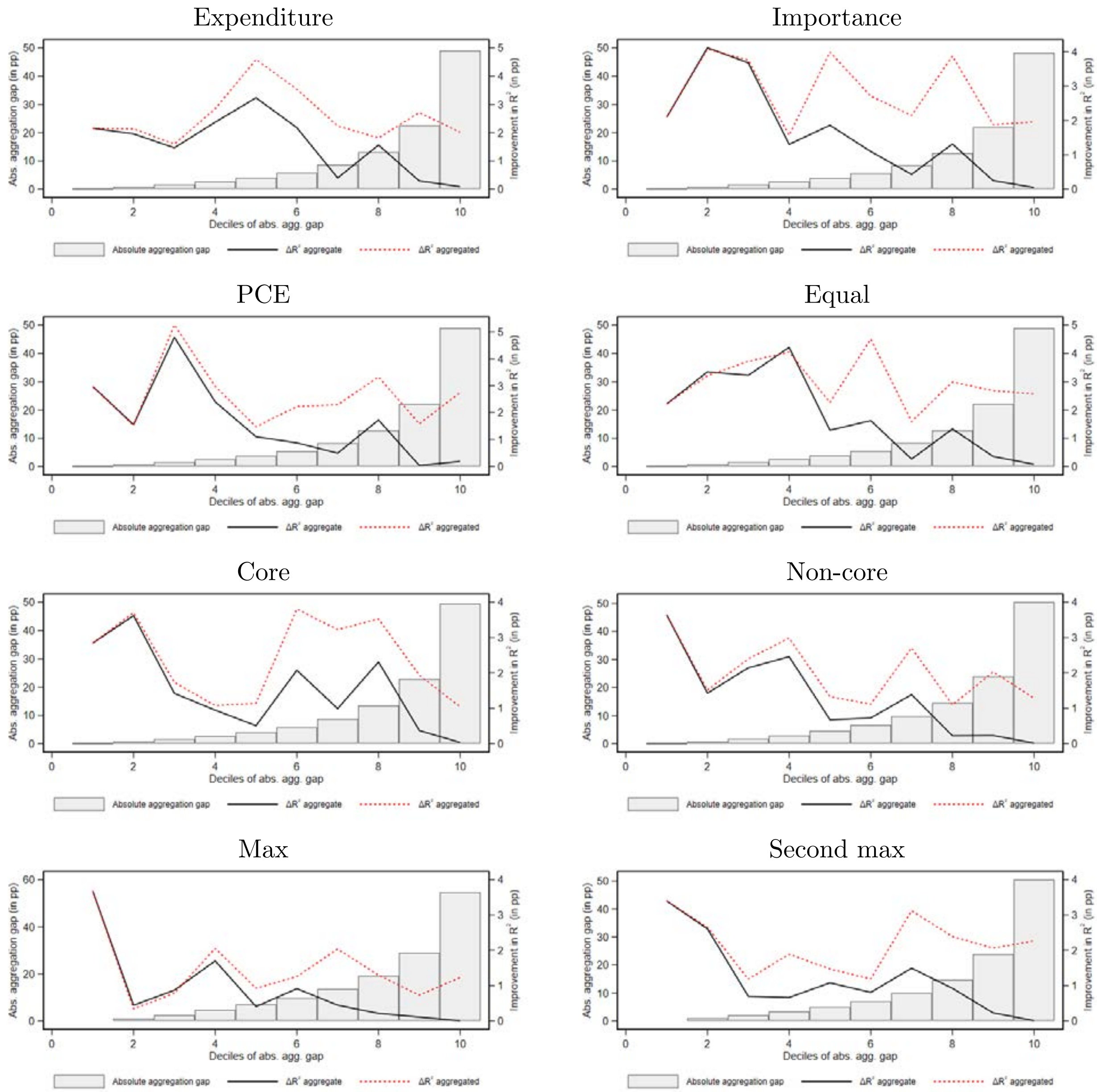
Table A.12: Instrumental Variable regression: 1 Year Ahead Spending Plans

	$\hat{\sigma} = 1 - \hat{\gamma}_{OLS}$ (OLS)	$\hat{\sigma} = 1 - \hat{\gamma}_{IV}$ (IV)	t-stat	F-stat (first stage)	N
12-month-ahead aggregate spending					
Aggregate	0.960***	0.863***	7.34	478	23053
Expenditure	0.821***	0.700***	7.55	364	23053
Importance	0.786***	0.696***	7.63	445	23053
PCE	0.788***	0.673***	7.60	410	23053
Equal	0.777***	0.684***	7.63	463	23053
Core inflation	0.842***	0.667***	7.38	279	23053
Non-core inflation	0.874***	0.758***	7.45	364	23053
Max	0.912***	0.748***	7.29	198	23053
Second max	0.870***	0.717***	7.44	261	23053
12-month-ahead nondurable spending					
Aggregate	0.957***	0.889***	4.12	144	4567
Expenditure	0.808***	0.551***	3.99	42	4567
Importance	0.747***	0.492***	3.96	38	4567
PCE	0.770***	0.490***	3.99	43	4567
Equal	0.732***	0.479***	4.00	45	4567
Core inflation	0.845***	0.590***	3.93	47	4567
Non-core inflation	0.842***	0.468***	3.58	21	4567
Max	0.907***	0.523***	3.14	14	4567
Second max	0.851***	0.552***	3.70	26	4567
12-month-ahead services spending					
Aggregate	0.967***	0.927***	4.90	503	23168
Expenditure	0.857***	0.838***	5.01	372	23168
Importance	0.824***	0.837***	5.01	445	23168
PCE	0.820***	0.824***	5.02	412	23168
Equal	0.813***	0.831***	5.03	464	23168
Core inflation	0.861***	0.818***	4.96	286	23168
Non-core inflation	0.904***	0.871***	4.97	374	23168
Max	0.929***	0.869***	4.92	215	23168
Second max	0.891***	0.850***	4.98	275	23168

Notes: Estimated Euler equations, based on cross-sectional data; measures of inflation expectations in first column instrumented with the mean inflation expectation from the distribution question; see Table 7 for details on OLS results. t statistics in third column, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$;

A.6 Spending Plan Regressions - Deciles of Absolute Aggregation Gaps

Figure A.2: Spending Plan Regressions - Deciles of Absolute Aggregation Gaps



Notes: The figure reproduces Figure 5 from the main text for different aggregated inflation expectations. Each panel compares the improvement, for each decile of the aggregation gap, in R^2 achieved by adding a measure of inflation expectations to the estimation in equation (5), relative to an estimation without inflation expectations. Grey bars: Mean aggregation gap (left vertical axis) for each decile on horizontal axis. Black line, right vertical axis: improvement in R^2 by adding the conventional measure of inflation expectations. Red line, right vertical axis: improvement in R^2 by adding the *aggregated* measure of inflation expectations.

B Low and High Inflation Environment

We split the sample in November 2021 and define the period from June 2020 to October 2021 as the "low inflation environment." The period after November 2021 (until August 2022) is defined as a "high inflation environment." We reproduce key statistics from the paper for both periods, to check for consistency.

B.1 Summary Statistics

Table B.1: Summary Statistics - Low and High Inflation Environment

	Mean		Std. Dev. (Disagreement)		Time Series Volatility	
Inflation environment	Low	High	Low	High	Low	High
Aggregate expectation	5.62	7.62	7.60	7.41	2.70	1.61
Category expectations						
Motor vehicles	4.88	6.46	5.86	6.11	1.90	0.97
Recreational goods	3.53	4.74	6.15	6.63	1.75	0.99
Other durable goods	3.60	4.93	5.97	6.42	1.83	1.00
Food and beverages	4.85	5.93	6.28	6.80	1.86	1.18
Gasoline	4.95	5.79	7.15	8.24	2.19	1.64
Other nondurable goods	3.71	4.85	5.83	6.32	1.48	0.94
Housing and utilities	4.73	5.25	6.42	6.53	1.76	0.88
Health care	3.43	4.81	6.53	6.51	1.62	1.08
Transportation services	4.36	5.53	5.89	6.68	1.60	1.06
Food services	4.36	5.45	6.42	6.53	1.68	1.01
Other services	3.95	4.91	5.48	5.90	1.39	0.85
Aggregated expectation						
Plausibly rational aggregation						
Expenditure	4.58	5.53	5.07	5.61	1.38	0.93
Importance	4.13	5.32	4.37	5.31	1.31	0.98
PCE	4.03	5.15	4.29	5.14	1.23	0.91
Behavioral aggregation						
Equal	4.00	5.28	4.20	5.19	1.31	0.96
Core inflation	4.39	5.23	5.10	5.38	1.36	0.83
Non-core inflation	5.30	6.38	5.80	6.98	1.72	1.35
Max	11.01	11.73	8.17	9.04	3.14	2.13
Second max	6.58	7.56	6.20	6.83	1.95	1.34

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.1.1 Socio-Demographics - Gender

Table B.2: Summary Statistics - Gender - High and Low Inflation Environment

	Male				Female			
	Mean		Disagreement		Mean		Disagreement	
Inflation environment	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	5.41	6.97	5.84	5.90	6.23	9.00	10.28	10.01
Category expectations								
Motor vehicles	4.96	6.11	5.20	5.35	4.88	6.94	6.27	6.89
Recreational goods	3.34	4.32	5.41	5.95	3.89	5.32	6.70	7.03
Other durable goods	3.51	4.67	5.27	5.69	3.70	5.30	6.58	6.96
Food and beverages	4.56	5.38	5.67	6.17	5.19	6.71	6.72	7.60
Gasoline	4.87	5.10	6.81	7.92	5.15	6.76	7.37	8.92
Other nondurable	3.59	4.52	5.20	5.52	3.85	5.28	6.36	6.84
Housing and util.	4.43	5.02	5.73	5.82	5.08	5.60	6.90	6.99
Health care	3.51	4.52	5.92	5.87	3.47	5.20	7.03	6.91
Transportation	4.08	5.06	5.18	5.98	4.59	6.31	6.41	7.40
Food services	4.23	5.10	5.78	5.96	4.48	5.88	6.82	6.97
Other services	3.76	4.57	4.81	5.25	4.12	5.31	6.11	6.35
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	4.35	5.09	4.42	5.00	4.90	6.19	5.61	6.30
Importance weights	3.95	4.80	3.90	4.68	4.40	5.99	4.72	5.87
PCE weights	3.82	4.69	3.77	4.54	4.35	5.77	4.73	5.74
Behavioral aggregation								
Equal weights	3.85	4.78	3.72	4.59	4.23	5.94	4.53	5.75
Core inflation	4.19	4.91	4.41	4.79	4.70	5.73	5.59	6.01
Non-core inflation	4.97	5.69	5.18	6.40	5.66	7.31	6.23	7.65
Max	10.55	10.62	7.47	7.84	11.81	13.69	8.76	10.84
Second max	6.27	6.90	5.68	6.06	6.99	8.65	6.63	7.93

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.1.2 Socio-Demographics - Grocery Shopper

Table B.3: Summary Statistics - Grocery Shopper - High and Low Inflation Environment

	Grocery Shopper				Not Grocery Shopper			
	Mean		Disagreement		Mean		Disagreement	
Inflation environment	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	5.29	6.66	7.14	6.89	5.95	7.88	8.04	7.64
Category expectations								
Motor vehicles	4.88	6.71	5.54	6.57	4.96	6.45	5.86	6.01
Recreational goods	3.46	4.83	5.48	6.34	3.54	4.75	6.18	6.60
Other durable goods	2.88	5.01	5.21	5.91	3.63	4.95	6.08	6.41
Food and beverages	4.92	6.48	5.56	6.75	4.88	5.92	6.32	6.77
Gasoline	5.40	6.18	7.63	9.19	4.96	5.78	7.09	8.13
Other nondurable	3.35	4.78	5.48	5.92	3.77	4.88	5.85	6.31
Housing and util.	4.53	5.48	6.00	6.32	4.84	5.22	6.46	6.48
Health care	3.27	4.83	6.28	6.14	3.52	4.82	6.53	6.47
Transportation	3.82	5.95	5.46	6.67	4.45	5.53	5.94	6.64
Food services	4.29	5.76	5.71	6.46	4.41	5.44	6.47	6.51
Other services	3.75	4.81	4.98	5.75	4.04	4.91	5.59	5.87
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	4.60	5.89	4.35	5.72	4.66	5.52	5.11	5.58
Importance weights	4.24	5.51	3.83	5.17	4.17	5.31	4.37	5.27
PCE weights	3.86	5.43	3.71	4.94	4.09	5.13	4.33	5.11
Behavioral aggregation								
Equal weights	3.92	5.47	3.67	5.03	4.03	5.27	4.22	5.17
Core inflation	4.11	5.50	4.29	5.39	4.48	5.23	5.15	5.34
Non-core inflation	5.48	7.28	5.02	7.31	5.31	6.37	5.82	6.96
Max	12.40	13.22	8.63	9.38	11.04	11.68	8.07	8.95
Second max	7.07	8.26	5.99	7.21	6.59	7.56	6.17	6.84

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.1.3 Socio-Demographics - Education

Table B.4: Summary Statistics - Education - High and Low Inflation Environment

	Low Education				High Education			
	Mean		Disagreement		Mean		Disagreement	
Inflation environment	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	6.59	8.83	10.25	10.19	5.46	7.15	6.12	5.76
Category expectations								
Motor vehicles	4.74	6.31	6.07	6.47	5.07	6.72	5.35	5.59
Recreational goods	3.41	4.49	6.62	7.04	3.68	5.09	5.54	5.89
Other durable goods	3.49	4.62	6.37	6.85	3.68	5.33	5.44	5.57
Food and beverages	4.86	5.90	6.60	7.34	4.89	6.16	5.61	6.18
Gasoline	5.03	5.85	7.18	8.39	5.00	6.02	6.93	8.32
Other nondurable	3.67	4.65	6.32	6.78	3.84	5.08	5.24	5.49
Housing and util.	4.62	5.01	6.95	6.94	4.98	5.58	5.80	5.74
Health care	3.35	4.54	6.83	6.77	3.61	5.15	6.21	5.90
Transportation	4.38	5.51	6.36	7.21	4.33	5.75	5.40	5.99
Food services	4.16	5.18	6.61	6.85	4.54	5.83	6.06	5.96
Other services	3.96	4.75	5.95	6.40	3.95	5.08	4.97	5.08
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	4.62	5.49	5.57	5.97	4.68	5.72	4.48	5.14
Importance weights	4.08	5.18	4.46	5.41	4.27	5.60	4.03	5.00
PCE weights	4.05	5.00	4.54	5.32	4.13	5.44	3.89	4.77
Behavioral aggregation								
Equal weights	3.98	5.13	4.30	5.28	4.14	5.55	3.90	4.88
Core inflation	4.40	5.12	5.64	5.76	4.50	5.47	4.52	4.88
Non-core inflation	5.40	6.57	6.18	7.35	5.24	6.36	5.28	6.56
Max	11.50	12.18	8.53	9.69	10.96	11.80	7.64	8.55
Second max	6.63	7.62	6.54	7.46	6.58	7.82	5.61	6.27

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.1.4 Socio-Demographics - Income

Table B.5: Summary Statistics - Income - High and Low Inflation Environment

	Low Income				Middle Income				High Income			
	Mean		Disag.		Mean		Disag.		Mean		Disag.	
Inflation environment	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	6.24	8.11	9.94	9.27	5.67	7.41	6.40	5.96	7.09	8.50	7.61	7.62
Category expectations												
Motor vehicles	4.72	6.39	6.11	6.57	4.95	6.32	5.16	5.70	5.20	6.75	5.35	5.33
Recreational goods	3.42	4.53	6.83	7.03	3.58	5.07	5.38	5.94	3.80	4.95	5.42	5.79
Other durable goods	3.37	4.73	6.77	6.89	3.69	4.99	5.27	5.73	4.12	5.37	5.68	5.41
Food and beverages	4.83	5.91	6.69	7.47	5.13	6.13	5.36	6.37	5.32	6.03	5.90	6.00
Gasoline	4.86	5.92	7.27	8.55	5.40	6.20	6.93	8.50	5.10	5.82	6.90	7.96
Other nondurable	3.46	4.69	6.29	6.78	3.85	4.93	5.32	5.58	4.26	5.06	5.23	5.42
Housing and util.	4.48	5.03	6.92	6.91	5.29	5.42	5.88	5.93	5.06	5.65	5.85	5.71
Health care	3.17	4.54	6.83	6.83	3.60	4.93	6.03	5.93	4.15	5.11	6.33	5.89
Transportation	4.37	5.65	6.57	7.23	4.71	5.65	5.23	6.15	4.30	5.46	5.30	5.92
Food services	4.09	5.22	6.83	6.96	4.38	5.61	5.65	6.03	4.83	5.82	6.21	5.80
Other services	3.90	4.75	5.98	6.35	4.22	4.94	4.89	5.24	4.14	5.18	5.15	5.05
Aggregated expectations												
Plausibly rational aggregation												
Expenditure weights	4.48	5.55	5.57	6.10	4.95	5.63	4.50	5.28	4.96	5.63	4.38	4.88
Importance weights	4.00	5.28	4.59	5.60	4.40	5.45	3.98	5.10	4.54	5.40	3.83	4.67
PCE weights	3.89	5.13	4.54	5.50	4.26	5.26	3.82	4.89	4.40	5.29	3.84	4.51
Behavioral aggregation												
Equal weights	3.87	5.25	4.42	5.49	4.16	5.39	3.83	4.98	4.35	5.35	3.66	4.52
Core inflation	4.25	5.18	5.54	5.88	4.69	5.28	4.53	4.95	4.64	5.44	4.22	4.68
Non-core inflation	5.34	6.65	6.27	7.53	5.62	6.58	5.16	6.70	5.61	6.15	5.62	6.24
Max	11.54	12.58	8.73	10.21	11.15	11.97	7.54	8.98	11.66	11.32	8.09	7.73
Second max	6.70	7.87	6.73	7.70	6.87	7.73	5.70	6.59	6.78	7.47	5.72	5.95

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.1.5 Socio-Demographics - Age

Table B.6: Summary Statistics - Age (mean) - High and Low Inflation Environment

	Below 35		35 to 44		45 to 54		Above 55	
Inflation environment	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	7.12	9.23	8.69	9.49	7.05	10.29	4.83	7.12
Category expectations								
Motor vehicles	4.11	5.39	5.69	6.20	5.03	7.23	5.26	7.99
Recreational goods	2.14	2.97	3.88	4.45	3.92	6.02	4.45	6.54
Other durable goods	2.49	3.33	3.95	4.71	3.92	5.89	4.35	6.56
Food and beverages	3.63	4.04	5.30	5.59	5.41	7.53	6.14	8.45
Gasoline	3.65	4.06	4.87	5.59	5.81	7.41	6.76	8.90
Other nondurable	2.55	3.31	3.99	4.68	4.40	6.07	4.59	6.26
Housing and util.	3.62	3.74	4.66	4.94	5.56	6.50	5.92	6.88
Health care	2.14	3.34	3.81	4.73	3.92	5.50	4.52	6.40
Transportation	3.30	3.80	4.49	5.26	4.50	7.01	5.36	7.65
Food services	2.66	3.69	4.45	4.95	4.89	6.45	5.94	7.67
Other services	3.17	3.66	3.72	5.02	4.51	5.79	4.67	6.01
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	3.51	3.90	4.78	5.21	5.39	6.90	5.76	7.60
Importance weights	2.81	3.43	4.28	4.87	4.85	6.66	5.65	7.72
PCE weights	2.92	3.38	4.11	4.76	4.57	6.46	5.41	7.38
Behavioral aggregation								
Equal weights	2.71	3.41	4.13	4.86	4.56	6.56	5.48	7.62
Core inflation	3.48	3.86	4.51	4.93	5.04	6.42	5.35	6.90
Non-core inflation	3.97	4.44	5.46	5.95	5.78	7.94	6.48	8.74
Max	11.02	10.01	12.02	11.01	11.34	13.42	12.04	15.05

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

Table B.7: Summary Statistics - Age (disagreement) - High and Low Inflation Environment

	Below 35		35 to 44		45 to 54		Above 55	
Inflation environment	Low	High	Low	High	Low	High	Low	High
Aggregate expectation	11.60	11.72	11.86	11.30	9.45	9.85	4.45	4.24
Category expectations								
Motor vehicles	6.76	5.79	6.47	5.95	5.79	6.37	4.36	5.89
Recreational goods	7.24	7.03	6.94	6.83	5.97	6.64	4.21	4.93
Other durable goods	6.80	6.70	7.10	6.54	6.04	6.26	4.37	5.06
Food and beverages	7.08	6.85	7.36	6.89	6.04	7.76	4.56	6.15
Gasoline	7.52	7.26	7.27	7.39	6.84	8.69	6.57	10.12
Other nondurable	7.09	6.86	6.67	6.41	5.70	6.42	4.03	4.80
Housing and util.	7.31	6.71	7.36	6.70	6.33	6.69	4.89	5.39
Health care	7.26	6.64	7.04	6.40	5.97	6.42	5.19	5.47
Transportation	6.87	6.84	6.79	6.67	5.90	7.07	4.37	5.84
Food services	7.19	6.54	7.03	6.39	6.07	6.66	5.06	5.70
Other services	6.40	6.28	6.38	6.08	5.31	5.90	3.83	4.47
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	5.24	4.90	5.25	5.32	4.75	6.09	4.15	5.43
Importance weights	3.70	3.98	4.46	4.83	4.14	5.78	3.92	5.47
PCE weights	3.95	4.13	4.68	4.85	4.02	5.54	3.63	4.96
Behavioral aggregation								
Equal weights	3.54	3.90	4.38	4.77	4.02	5.59	3.73	5.25
Core inflation	5.55	4.96	5.33	5.14	4.72	5.68	4.10	4.92
Non-core inflation	6.39	6.30	6.31	6.60	5.42	7.68	4.74	7.08
Max	8.32	7.75	8.76	8.16	8.56	10.04	7.83	10.52
Second max	6.83	6.36	6.73	6.51	6.55	7.63	5.16	6.99

Notes: This table presents summary statistics on the demographic distribution of expectations. Statistics based on averages of Huber-robust and survey-weighted daily means on expectations across demographics.

B.2 Aggregate vs. Aggregated Inflation Expectations

B.2.1 Low Inflation Environment

Table B.8: Aggregate vs. Aggregated Inflation Expectations - Before November 2021

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure	0.444***								0.0227
	(14.75)								(0.18)
Importance		0.567***							-0.0461
		(16.22)							(-0.28)
PCE			0.533***						-0.442***
			(14.89)						(-3.43)
Equal				0.605***					0.930***
				(16.56)					(5.09)
Core Inflation					0.388***				-0.00417
					(13.67)				(-0.04)
Non-core Inflation						0.321***			-0.0209
						(12.42)			(-0.38)
Max							0.281***		0.179***
							(16.52)		(5.68)
Second max								0.327***	-0.0643
								(13.24)	(-1.34)
Constant	6.287***	5.843***	6.088***	5.794***	6.734***	6.891***	4.885***	6.180***	4.696***
	(24.88)	(22.67)	(23.42)	(22.68)	(28.05)	(27.26)	(17.63)	(22.77)	(16.88)
N	20685	20685	20685	20685	20685	20685	20685	20685	20685
R2	0.0383	0.0476	0.0416	0.0499	0.0326	0.0291	0.0426	0.0345	0.0586
AIC	180209.6	180008.8	180139.3	179957.3	180331.9	180407.3	180116.8	180289.9	179783.1

Notes: The table presents estimates on a micro level for a linear regression of reported aggregate on one (column 1 to 8) or multiple (column 9) aggregated, category-based measures of inflation. t statistics in parentheses, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights and Huber-robust weights to ensure that sample is representative and independent of outliers, respectively.

B.2.2 High Inflation Environment

Table B.9: Aggregate vs. Aggregated Inflation Expectations - After November 2021

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure	0.590***								0.168
	(31.64)								(1.94)
Importance		0.657***							-0.283*
		(32.64)							(-2.37)
PCE			0.658***						-0.241*
			(31.49)						(-2.17)
Equal				0.688***					0.883***
				(33.58)					(5.23)
Core					0.565***				0.0171
					(28.49)				(0.25)
Non core						0.399***			-0.0417
						(27.73)			(-1.13)
First max							0.316***		0.0956***
							(28.34)		(4.43)
Second max								0.434***	0.0471
								(28.61)	(1.29)
Constant	8.737***	8.383***	8.542***	8.241***	9.249***	9.800***	7.858***	8.595***	7.452***
	(44.84)	(41.89)	(42.17)	(41.05)	(47.78)	(51.94)	(36.85)	(40.88)	(35.64)
N	29936	29936	29936	29936	29936	29936	29936	29936	29936
r2	0.0901	0.0948	0.0915	0.0977	0.0797	0.0684	0.0772	0.0807	0.104
AIC	260403.5	260247.5	260355.8	260151.8	260742.0	261110.2	260823.7	260709.9	259956.8

Notes: The table presents estimates on a micro level for a linear regression of reported aggregate on one (column 1 to 8) or multiple (column 9) aggregated, category-based measures of inflation. t statistics in parentheses, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights and Huber-robust weights to ensure that sample is representative and independent of outliers, respectively.

B.3 The Aggregation Gap

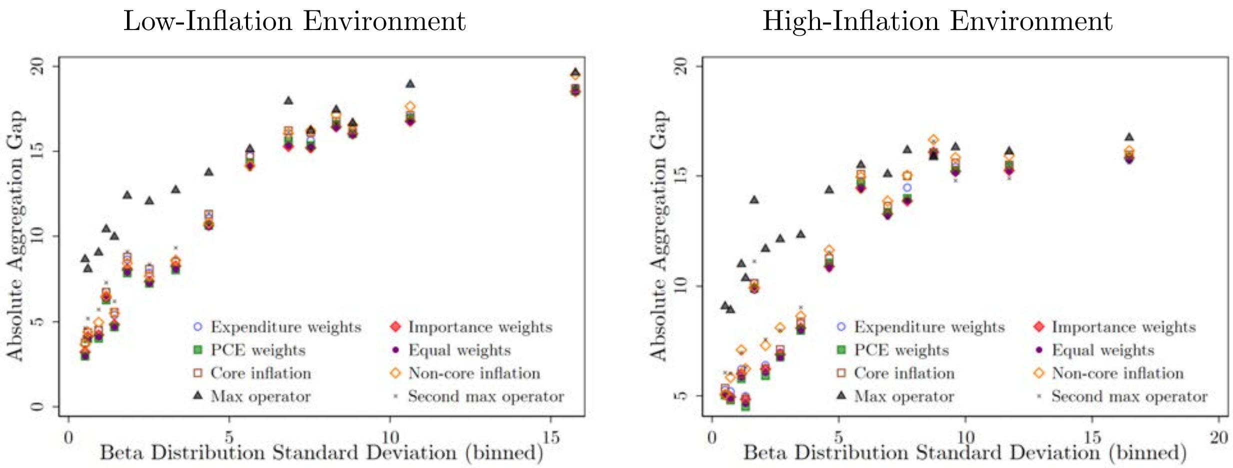
Table B.10: Summary Statistics

Inflation environment	Mean Aggregation Gap Λ_i Low	High	Mean Absolute Aggregation Gap $\mid \Lambda_i \mid$ Low	High
Plausibly rational aggregation				
Expenditure	0.65***	1.77***	5.62***	5.65***
Importance	1.01***	1.91***	5.35***	5.55***
PCE	1.17***	1.95***	5.16***	5.49***
Behavioral aggregation				
Equal	1.07***	1.99***	5.18***	5.42***
Core inflation	-4.26***	-3.61***	9.20***	9.03***
Non-core inflation	-0.78***	-0.17***	6.52***	6.51***
Max	0.63***	1.36***	6.02***	6.16***
Second max	0.83***	2.07***	5.72***	5.64***

Notes: This table presents Huber-robust and survey-weighted estimates for the mean aggregation gap and mean absolute aggregation gap; Stars: significance level of a t-test that numbers are different from zero. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$;

B.3.1 The Absolute Aggregation Gap

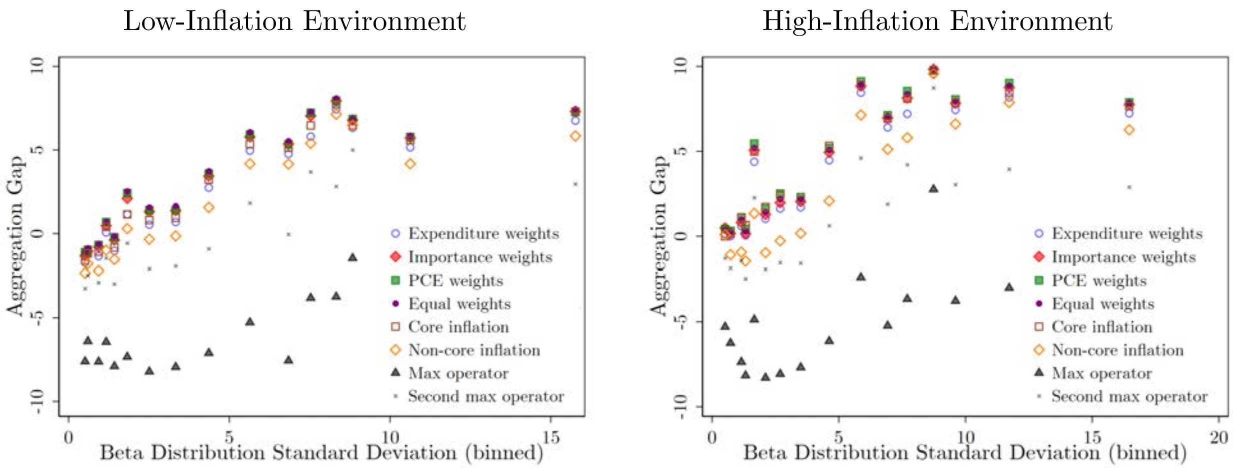
Figure B.1: The Absolute Aggregation Gap and Aggregate Uncertainty



Notes: The panel display the correlation between the absolute aggregation gap $\mid \Lambda_i \mid$ and the subjective uncertainty over aggregate inflation expectations obtained via a beta distribution over a probabilistic question;; variable on horizontal axis binned; the left panel displays the relation in the low inflation environment, the right panel in the high inflation environment.

B.3.2 The Directed Aggregation Gap

Figure B.2: The Aggregation Gap and Aggregate Uncertainty



Notes: The panel display the correlation between the aggregation gap Λ_i and the subjective uncertainty over aggregate inflation expectations obtained via a beta distribution over a probabilistic question;; variable on horizontal axis binned; the left panel displays the relation in the low inflation environment, the right panel in the high inflation environment.

B.4 Spending Plans

Table B.11: 1 Year Ahead Spending Plans

	$\hat{\sigma} = 1 - \hat{\gamma}_1$	t-stat	R^2	AIC	p-val (LR)	N
June 2020 - October 2021						
Aggregate	0.956***	5.90	0.071	76485	-	10767
Expenditure	0.826***	9.67	0.090	76271	0	10767
Importance	0.778***	10.84	0.095	76204	0	10767
PCE	0.781***	10.22	0.094	76221	0	10767
Equal	0.765***	10.83	0.096	76196	0	10767
Core inflation	0.847***	8.68	0.086	76317	0	10767
Non-core inflation	0.884***	8.28	0.080	76387	0	10767
Max	0.916***	8.52	0.080	76381	0	10767
Second max	0.865***	9.72	0.090	76261	0	10767
November 2021 - August 2022						
Aggregate	0.964***	4.91	0.047	91488	-	12889
Expenditure	0.818***	11.84	0.079	91052	0	12889
Importance	0.792***	12.82	0.082	91012	0	12889
PCE	0.793***	12.20	0.080	91042	0	12889
Equal	0.785***	12.57	0.082	91010	0	12889
Core inflation	0.839***	10.12	0.070	91178	0	12889
Non-core inflation	0.869***	11.86	0.075	91116	0	12889
Max	0.910***	11.85	0.070	91178	0	12889
Second max	0.874***	10.61	0.071	91159	0	12889

Notes: Estimated Euler equations, relying on various measures of aggregate or aggregated inflation expectations; t statistics in third column, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights to ensure that sample is representative.

Table B.12: 1 Year Ahead Services Spending Plans

	$\hat{\sigma} = 1 - \hat{\gamma}_1$	t-stat	R^2	AIC	p-val (LR)	N
June 2020 - October 2021						
Aggregate	0.970***	5.03	0.071	73966	-	10809
Expenditure	0.866***	9.24	0.086	73783	0	10809
Importance	0.825***	10.66	0.092	73711	0	10809
PCE	0.823***	9.97	0.093	73710	0	10809
Equal	0.812***	10.55	0.094	73693	0	10809
Core inflation	0.878***	8.64	0.084	73809	0	10809
Non-core inflation	0.904***	7.87	0.081	73846	0	10809
Max	0.925***	9.53	0.084	73808	0	10809
Second max	0.885***	9.85	0.092	73711	0	10809
November 2021 - August 2022						
Aggregate	0.965***	5.25	0.052	88318	-	12958
Expenditure	0.851***	11.18	0.079	87952	0	12958
Importance	0.823***	11.84	0.084	87873	0	12958
PCE	0.818***	11.54	0.085	87862	0	12958
Equal	0.814***	11.79	0.086	87851	0	12958
Core inflation	0.848***	11.09	0.078	87955	0	12958
Non-core inflation	0.904***	10.03	0.069	88091	0	12958
Max	0.931***	10.48	0.067	88109	0	12958
Second max	0.896***	9.92	0.073	88033	0	12958

Notes: Estimated Euler equations, relying on various measures of aggregate or aggregated inflation expectations; t statistics in third column, based on robust standard errors; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; regression adjusted with survey weights and Huber-robust weights to ensure that sample is representative and independent of outliers, respectively.

C Additional Proofs

C.1 Second-Order Approximation to the Price Index

A second-order log-linear approximation for conventional price indices, such as

$$
\begin{aligned}
1 &= \int_0^1 \left(\frac{P(i)}{P}\right)^{1-\epsilon} di && (6) \\
&= \int_0^1 e^{(1-\epsilon)(p(i)-p)} di && (7) \\
&\approx 1 + (1-\epsilon)\int_0^1 (p(i)-p)di + \frac{(1-\epsilon)^2}{2}\int_0^1 (p(i)-p)^2 di && (8)
\end{aligned}
$$

where $p(i)$ and p denote logs of the respective prices.

As a result,

$$
\begin{aligned}
p_t &\approx \bar{p}_{i,t} + \frac{1-\epsilon}{2}\int_0^1 (p(i)-p)^2 di && (9) \\
&= \bar{p}_{i,t} + \frac{1-\epsilon}{2}\text{var}(p(i)) && (10)
\end{aligned}
$$

where $\bar{p}_{i,t}$ denotes the average of log prices.

C.2 Model Fit

In order to compare the model fit of different expectations measures, we rely on the Akaike Information Criterion (AIC). This is equal to:

$$AIC = 2k - 2\ln(\hat{L})$$

Where k is the number of estimated parameters in the model and $\hat{L}$ represents the maximized value of the likelihood function.

Similarly, to study the statistical significance of differences in the model fit between various measures of expectations, we compute the likelihood ratio of models. Specifically assume that the AIC is lower for model 2 than for model 1, $AIC_1 > AIC_2$. Then, the likelihood ratio is defined as:

$$
\begin{aligned}
LR &= \frac{\hat{L}_1}{\hat{L}_2} \\
&= \exp\left(\frac{AIC_2 - AIC_1}{2}\right) \text{ if } k_1 = k_2
\end{aligned}
$$

where the second line links the likelihood ratio to the AIC, given that both models estimate the same number of parameters. The likelihood ratio then shows how probable model 1 is to minimize the information loss, relative to model 2.

D Model of Cognitive Recall - Aggregation Levels

This section outlines a simple model that embodies the intuition of Bordalo et al. (2022, 2023) in the context of aggregation levels. The model shows that framing survey questions in terms of category-specific rather than aggregate inflation exploits a variation in the ease of recall—i.e., survey participants' ability to relate to relevant information from their memory—sufficient to account for some findings in this paper.

The main intuition of the model is as follows: consumers form inflation expectations—both at the aggregate and category-specific level—by simulating the future, based on recalled experiences from their memory. As recall is imperfect, simulations from experiences irrelevant to the aggregation level forecasted interfere with those that are relevant for inflation expectations. The severity of interference increases in the frequency of irrelevant recalls, the incidence of which is likely higher for more complex, abstract concepts of inflation.

D.1 Memory Database

Assume a decision maker (DM) with a memory database E, a set of $N \geq 1$ experiences e. Individual experiences e are represented by F$\geq$ 1 features; each feature f takes a value in $f \in \{0, 1\}$.

A cue or concept that the DM attempts to recall is described by C. A concept partitions the memory database E along a number of features $L \geq 1$. The subset $H_C \subset E$ is the set of experiences relevant for concept C. We denote the disjoint alternative to concept C by $\bar{C}$; the subset $\bar{H}_C = E \setminus H_C$ contains experiences not relevant for C, such that $E = H_C \cup \bar{H}_C$. A concept may, for example, be *aggregate inflation*, or *food price inflation*. In the context of personal expenditure recall (Hurd and Rohwedder, 2009; Winter, 2004), concepts might be *total expenditure* or *expenditure on food items*.

A function $S(u, v) : E \times E \mapsto [0, \bar{S}]$ measures the similarity between two experiences u and v. The similarity increases in the number of shared features F. The similarity between two subsets of the database $A \subset E$, $B \subset E$ is the average pairwise similarity between their elements.

$$S(A, B) = \sum_{u \in A} \sum_{v \in B} S(u, v) \frac{1}{\mid A \mid} \frac{1}{\mid B \mid} \tag{11}$$

By $S(e, C)$, we denote the similarity between an experience e and the concept C.

When cued with concept C, the probability that the DM recalls experience e is r(e,C), a function of the similarity between experience e and the cued concept C.

$$r(e, C) = \frac{S(e, C)}{\sum_{u \in E} S(u, C)} \tag{12}$$

If experience e is more similar to concept C, it is recalled with a higher probability.

When aiming to recall experiences relevant to concept C, the retrieval fluency $r(C)$ measures the probability that any experience e recalled is part of H_C, the relevant set of experiences. The retrieval fluency is equal to the sum of probabilities of recalling each element of H_C:

$$
\begin{aligned}
r(H_C) &= \sum_{e\in H_C} r(e,C) \\
&= \frac{\sum_{e\in H_C} S(e,C)}{\sum_{u\in H_C} S(u,C) + \sum_{u\in \bar{H}_C} S(u,C)} \\
&= \frac{\alpha(H_C)}{\alpha(H_C) + \frac{S(C,\bar{C})}{S(C,C)}\alpha(\bar{H}_C)} \\
&= \frac{\alpha(H_C)}{\alpha(H_C) + \mathbb{I}(C)^{-1}[1-\alpha(H_C)]}
\end{aligned} \tag{13}
$$

Where $\alpha(H_C) = \frac{|H_C|}{|E|}$ denotes the frequency of elements in E relevant for C. The variable $\mathbb{I}(C) = \frac{S(C,C)}{S(C,\bar{C})} \in [0,1]$ denotes the *relative experience homogeneity*, a measure of the average homogeneity of relevant experiences in H_C, relative to their average similarity to irrelevant experiences. $S(C,C)$ is the self-similarity between experiences in C (homogeneity) and $S(C,\bar{C})$ the similarity between experiences relevant for C and irrelevant experiences. Following Bordalo et al. (2023), we assume that $S(C,C) \geq S(C,\bar{C})$.

Proposition 1 (Retrieval Fluency)**.** *The retrieval fluency of relevant experiences when cued with C is a increasing function of the relative experience homogeneity of C, $\mathbb{I}(C)$, as well the prevalence of relevant experiences in E, $\alpha(H_C) = \frac{|H_C|}{|E|}$.*

$$
r(H_C) = \frac{\alpha(H_C)}{\alpha(H_C) + \mathbb{I}(C)^{-1}[1-\alpha(H_C)]} \tag{14}
$$

It holds that

$$
\frac{\partial r(H_C)}{\partial \mathbb{I}(C)} = \frac{\alpha(H_C)[1-\alpha(H_C)]}{[\alpha(H_C) + \mathbb{I}(C)^{-1}[1-\alpha(H_C)]]^2} \geq 0 \tag{15}
$$

$$
\frac{\partial r(H_C)}{\partial \alpha(C)} = \frac{\mathbb{I}(C)^{-1}}{[\alpha(H_C) + \mathbb{I}^{-1}(C)[1-\alpha(H_C)]]^2} \geq 0 \tag{16}
$$

Proof. The result follows directly from the preceding definitions. □

For the aggregation levels of inflation, we assume that the retrieval fluency is higher for category-specific recall, as the self-similarity of relevant experiences is larger, and thus, the $\mathbb{I}(C)$ is smaller compared to cuing respondents with aggregate inflation. For example, prices for food are arguably more similar to each other than are prices in general. A similar argument can be made in terms of expenditure recall (Hurd and Rohwedder, 2009; Winter, 2004); cuing survey participants with categories rather than asking for total spending, increases the retrieval fluency of relevant experiences, thereby reducing the interference from irrelevant ones.

Note that when recalling $D < \infty$ experiences, the fraction of relevant recalls is a random variable, converging towards the recall fluency for $D \to \infty$. For simplicity, in what follows, we assume the recall fluency to be similar across all DM's for a concept C.

D.2 Simulation

Assume that the DM uses recalled experiences to simulate the future. Given a recall e, the simulated future is $s(e)$. Assume that the simulation is a function of the features of experience e, such that $s(e) : F \mapsto \mathbb{R}$. For example, a survey participant cued with gasoline and energy inflation might recall a recent trip to the gas station (the experience e) and simulate an energy price inflation of $s(e) = 6.8\%$, based on the features of his specific recall (single trip to the gas station).

For a given set H_C we define η_i as the mean simulation of DM i and σ_η^2 as the the variance of simulations. Further, we assume that σ_η^2 decreases in the homogeneity of experiences $S(C, C)$. Intuitively, this means that the more similar experiences in H_C are, the more similar simulations based on the respective experiences will be. By analogy, we define ι_i to be the mean simulation from the set of irrelevant experiences, $\bar{H}_C$. The variable σ_ι^2 denotes the variance of simulations from irrelevant experiences. From the assumption of $S(C, C) > S(\bar{C}, \bar{C})$ it follows that $\sigma_\iota^2 > \sigma_\eta^2$. Note that the mean of simulations is specific to an individual; we assume the variance of simulations to be equal across respondents.

Assumption 1. *Simulation The DM uses draws e from the memory database E to simulate, relying on simulation function $s(e) : F \mapsto \mathbb{R}$. We assume that*

$$\mathbb{E}_i[s(e) \mid e \in H_C] = \eta_i; \quad Var[s(e) \mid e \in H_C] = \sigma_\eta^2$$
$$\mathbb{E}_i[s(e) \mid e \in \bar{H}_C] = \iota_i; \quad Var[s(e) \mid e \in \bar{H}_C] = \sigma_\iota^2$$

For the mean of relevant (irrelevant) simulations across DM's it holds that $\mathbb{E}\eta_i = \eta$ ($\mathbb{E}\iota_i = \iota$).

Assumption 1 adds some structure to the simulation based on experiences. In our example of gasoline and energy price recall, trips to the gas station, electricity bills and forecasts for energy prices would be examples of relevant experiences, with a mean simulation of η_i. Instead, a specific oil painting that the DM has seen, the probability of winning the lottery he read about in the newspaper or his recall of last weeks' sports results might be experiences irrelevant to forecasting energy and gasoline inflation that might still come to mind. For these experiences we assume that the DM has a mean simulation of ι_i, based on the features of the irrelevant experiences. We also assume that the simulations based on relevant experiences are more similar quantitatively than those based on irrelevant recalls, that is have a smaller variance.

It follows from Assumption 1 that the variance of relevant simulations across DM's is $\sigma_\eta^2 + \upsilon_\eta^2$ ($\sigma_\iota^2 + \upsilon_\iota^2$). Here, $Var(\eta_i) = \upsilon_\eta^2$ ($Var(\iota_i) = \upsilon_\iota^2$) denotes the variance of mean relevant (irrelevant) simulations across DM's. We assume that the variance of irrelevant experiences across DM's is larger, $\sigma_\iota^2 + \upsilon_\iota^2 \geq \sigma_\eta^2 + \upsilon_\eta^2$.

D.3 Moments

The simulated, expected value (forecast) for concept C of DM i is the average over performed simulations $s(e_d)$:

$$\pi_C^i = \frac{1}{D}\sum_{d=1}^{D} s(e_d) \tag{17}$$

The parameter $D > 0$ denotes the number of experience draws (and thus simulations performed) that the DM retrieves from the memory dataset E. Note that π_C^i is a random variable in the model. The expected value for DM i is

$$\begin{aligned}
\mathbb{E}_i\pi_C &= \frac{1}{D}\sum_{d=1}^{D}\left[\sum_{u\in E} r(u,C)s(u)\right] \\
&= \sum_{u\in H_C} r(u,C)s(u) + \sum_{u\in \bar{H}_C} r(u,C)d(u) \\
&= r(H_C)\sum_{u\in H_C}\frac{r(u,C)}{r(H_C)}s(u) + [1-r(H_C)]\sum_{u\in \bar{H}_C}\frac{r(u,C)}{1-r(H_C)}s(u) \\
&= r(H_C)\mathbb{E}\left[s(e) \mid e\in H_C\right] + [1-r(H_C)]\mathbb{E}\left[s(e) \mid e\notin H_C\right]
\end{aligned}$$

Where $\mathbb{E}_i$ denotes the expectations operator conditional on DM i. We replace by the mean simulation from relevant experiences $\mathbb{E}\left[s(e) \mid e\in H_C\right] = \sum_{u\in H_C}\frac{r(u,C)}{r(H_C)}s(u) = \eta_i$ and irrelevant experiences $\mathbb{E}\left[s(e) \mid e\notin H_C\right] = \sum_{u\in \bar{H}_C}\frac{r(u,C)}{1-r(H_C)}s(u) = \iota_i$. The term $\frac{r(u,C)}{r(H_C)}$ is the probability of recalling experience $u\in H_C$, conditional on recalling a relevant experience. It follows that

$$\mathbb{E}_i\pi_C = r(H_C)\eta_i + [1-r(H_C)]\iota_i \tag{18}$$

Next, the expected value across respondents is given by $\mathbb{E}\left[\mathbb{E}_i\pi_C\right] = \mathbb{E}\pi_C$. This corresponds to the limit of the population mean for a large number of DM's. We may write the expected value as

$$\mathbb{E}\pi_C = r(H_C)\eta + [1-r(H_C)]\iota \tag{19}$$

Proposition 2 (Expected Beliefs and Interference)**.** *The expected value of beliefs among respondents for concept C,* $\mathbb{E}\pi_C$ *is defined by* (20)

$$\mathbb{E}\pi_C = \eta + [1-r(H_C)][\iota-\eta] \tag{20}$$

where η *is the mean relevant experience simulation,* $\iota-\eta$ *measures the mean interference from irrelevant experiences on the intensive margin (bias) and* $[1-r(H_C)]\in[0,1]$ *the degree of interference on the extensive margin.*

Proof. From the above, we have that

$$\mathbb{E}\pi_C = r(H_C)\eta + [1 - r(H_C)]\iota$$

From which it follows that

$$\begin{aligned}
\mathbb{E}\pi_C &= r(H_C)\eta + \iota - \iota r(H_C) \\
\mathbb{E}\pi_C &= \eta + \iota - \eta + r(H_C)(\eta - \iota) \\
\mathbb{E}\pi_C &= \eta + (\iota - \eta)[1 - r(H_C)]
\end{aligned}$$

□

Proposition 3 discusses the disagreement among DM's beliefs. The disagreement decreases in the number of draws per individual D from the memory database E. Under standard assumptions, disagreement decreases with the retrieval fluency and thus the relative experience homogeneity. In line with the model, in the data, we document lower disagreement for category-specific inflation expectations.

Proposition 3 (Belief Disagreement)**.** *For a common set of experiences E, the inter-personal disagreement for π_C^i, $Var(\pi_C^i)$ is defined by* (21)*.*

$$Var(\pi_C^i) = \frac{1}{D}\left[r(H_C)(\sigma_\eta^2 + \upsilon_\eta^2) + [1 - r(H_C)](\sigma_\iota^2 + \upsilon_\iota^2) + [1 - r(H_C)]r(H_C)(\iota - \eta)^2\right] \quad (21)$$

where the term $\sigma_\eta^2 + \upsilon_\eta^2$ ($\sigma_\iota^2 + \upsilon_\iota^2$) denotes the variance of relevant (irrelevant) experience simulations across respondents. Thus, the change in disagreement with $r(H_C)$ is

$$\frac{\partial Var(\pi_C)}{\partial r(H_C)} = \frac{1}{D}\left[\sigma_\eta^2 + \upsilon_\eta^2 - \sigma_\iota^2 - \upsilon_\iota^2 + (1 - 2r(H_C))(\iota - \eta)^2\right] > 0 \quad \forall r(H_C) \geq \bar{r} \quad (22)$$

Disagreement decreases in the retrieval fluency as long as the retrieval fluency is sufficiently large. ($\bar{r} = 0.5\left[1 - \frac{\sigma_\iota^2 + \upsilon_\iota^2 - \sigma_\eta^2 - \upsilon_\eta^2}{(\iota - \eta)^2}\right] \leq 0.5$). For $0 < r(H_C) < 1$ disagreement increases in the absolute value of the intensive margin of interference $(\iota - \eta)^2$.

Proof.

$$
\begin{aligned}
Var(\pi_C^i) &= \mathbb{E}\left(\frac{1}{D}\sum_{d=1}^{D} s(e_d) - \mathbb{E}\pi_C\right)^2 \\
&= \frac{1}{D^2}\mathbb{E}\left(\sum_{d=1}^{D} (s(e_d) - \mathbb{E}\pi_C)\right)^2 \\
&= \frac{1}{D}\left[\mathbb{E}\left(s(e) - \mathbb{E}\pi_C\right)^2\right] \\
&= \frac{1}{D}\left[\mathbb{E}(s(e)^2) - (\mathbb{E}\pi_C)^2\right] \\
&= \frac{1}{D}\left[r(H_C)\mathbb{E}\left[s(e)^2 \mid e \in H_C\right] + (1 - r(H_C))\mathbb{E}\left[s(e)^2 \mid e \in \bar{H}_C\right] - \mathbb{E}(\pi_C)^2\right] \\
&= \frac{1}{D}\left[r(H_C)(\sigma_\eta^2 + \eta^2 + \upsilon_\eta^2) + (1 - r(H_C))(\sigma_\iota^2 + \iota^2 + \upsilon_\iota^2) - \mathbb{E}(\pi_C)^2\right] \\
&= \frac{1}{D}\left[r(H_C)(\sigma_\eta^2 + \upsilon_\eta^2) + [1 - r(H_C)](\sigma_\iota^2 + \upsilon_\iota^2) + (1 - r(H_C))r(H_C)(\iota - \eta)^2\right] \quad (23)
\end{aligned}
$$

In equation (23), we use that draws from the relevant and irrelevant set of memories are independent. The derivative part of the proposition follows directly from this last line. □

Proposition 4 discusses the subjective uncertainty of DM i – the variance among her D simulations – denoted by $S_i(C)$. The subjective uncertainty is a random variable, dependent on the experiences retrieved from memory. Proposition 4 shows that the expected value of subjective uncertainty for a decision maker is a function of the retrieval fluency $r(H_C)$.

Proposition 4 (Subjective Uncertainty)**.** *A DM retrieves $D > 0$ experiences from her memory dataset E, which she uses for simulation.The subjective uncertainty $S_i(C)$ over individual outcomes (simulations) is*

$$
S_i(C) = \frac{1}{D}\sum_{d=1}^{D}\left(s(e_d) - \frac{1}{D}\sum_{d=1}^{D} s(e_d)\right)^2 \tag{24}
$$

The expected subjective uncertainty of DM i is defined by equation (25)*.*

$$
\mathbb{E}_i S(C) = r(H_C)\sigma_\eta^2 + [1 - r(H_C)]\sigma_\iota^2 + (1 - r(H_C))r(H_C)(\iota_i - \eta_i)^2 \tag{25}
$$

The change in the expected subjective uncertainty with $r(H_C)$ is

$$
\frac{\partial \mathbb{E}_i S(C)}{\partial r(H_C)} = \left[\sigma_\eta^2 - \sigma_\iota^2 + (1 - 2r(H_C))(\iota_i - \eta_i)^2\right] > 0 \quad \forall r(H_C) \geq \bar{r} \tag{26}
$$

The expected subjective uncertainty decreases in the retrieval fluency as long as the retrieval fluency is sufficiently large ($\bar{r} = 0.5\left[1 - \frac{\sigma_\iota^2 - \sigma_\eta^2}{(\iota_i - \eta_i)^2}\right] \leq 0.5$).

Proof. The result directly follows from the definitions and preceding propositions. □

Proposition 2-4 state that with lower retrieval fluency $r(H_C)$, the subjective uncertainty increases, disagreement among respondents increases, and the extensive margin of interference increases.

D.4 Aggregation Levels

We now turn to model predictions related to different levels of inflation expectations elicited in the survey.

Assumption 2. *Assume that concept $C = agg.$ refers to aggregate inflation expectations, while $C = k$ refers to category-specific inflation expectations for consumption category $k \in (1, K)$. Further assume that $\eta_i = \sum_k^{11} \omega_k \eta_{k,i}$, that is, the weighted sum of the category-specific mean simulations based on relevant experiences is equal to the mean of simulations based on relevant experiences for aggregate inflation expectations. The parameter ω_k denotes the expenditure weight to category k, which is assumed to be an exogenous parameter. Also, assume that the retrieval fluency is similar for all categories, i.e. $r(k) = r(cat.) \forall k \in (1, K)$.*

Proposition 5 (Covariance between Aggregate and Category-specific expectation)**.** *The covariance between aggregate and category-specific inflation expectations for category k is*

$$Cov(\pi, \pi_k) \quad = \quad r(agg.)r(k)\omega_k \upsilon^2_{\eta_k} > 0$$

The covariance increases in the retrieval fluency, both for aggregate and category specific inflation. The variable $\upsilon^2_{\eta_k}$ denotes the variance of the mean of relevant experience simulations across respondents for category k.

Proof.

$$\begin{aligned}
Cov(\pi_i, \pi_{k,i}) &= \mathbb{E}[\pi_i \pi_{k,i}] - \mathbb{E}[\pi_i]\mathbb{E}[\pi_{k,i}] \\
&= r(a)r(k)\mathbb{E}[s(a)s(k) \mid a \in H_{agg.}, k \in H_k] + (1 - r(a))r(k)\mathbb{E}[s(a)s(k) \mid a \notin H_{agg.}, k \in H_k] \\
&\quad + r(a)(1 - r(k))\mathbb{E}[s(a)s(k) \mid a \in H_{agg.}, k \notin H_k] \\
&\quad + (1 - r(a))(1 - r(k))\mathbb{E}[s(a)s(k) \mid a \notin H_{agg.}, k \notin H_k] - \mathbb{E}[\pi_i]\mathbb{E}[\pi_{k,i}] \\
&= r(a)r(k)\mathbb{E}\eta_i\eta_{i,k} + (1 - r(a))r(k)\mathbb{E}\iota_i\eta_{i,k} + r(a)(1 - r(k))\mathbb{E}\eta_i\iota_{k,i} \\
&\quad + (1 - r(a))(1 - r(k))\mathbb{E}\iota_i\iota_{i,k} - \mathbb{E}[\pi_i]\mathbb{E}[\pi_{k,i}] \\
&= r(a)r(k)(\omega_k\upsilon^2_{\eta_k} + \eta_i\eta_{i,k}) + (1 - r(a))r(k)\mathbb{E}\iota_i\mathbb{E}\eta_{i,k} + r(a)(1 - r(k))\mathbb{E}\eta_i\mathbb{E}\iota_{k,i} \\
&\quad + (1 - r(a))(1 - r(k))\mathbb{E}\iota_i\mathbb{E}\iota_{i,k} - \mathbb{E}[\pi_i]\mathbb{E}[\pi_{k,i}] \\
&= r(a)r(k)\omega_k\upsilon^2_{\eta_k}
\end{aligned}$$

The result follows from only the relevant simulations for aggregate and category specific inflation expectations being correlated for any respondent (see Assumption 2). □

D.5 Aggregation Gap

The model is able to explain the documented gap between aggregate and aggregated inflation expectations. The aggregation gap for respondent i is defined as

$$\Lambda_i = \pi_{agg.,i} - \sum_{k=1}^{K} \omega_k \pi_{k,i}$$

Using (20) it follows that the expected aggregation gap $\mathbb{E}\Lambda_i$ is equal to

$$\mathbb{E}\Lambda_i \quad = \quad [1 - r(agg.)](\iota_{agg.} - \eta_{agg.}) - [1 - r(cat.)] \sum_{k=11}^{11} \omega_k(\iota_k - \eta_k) \tag{27}$$

The expected aggregation gap is positive $\mathbb{E}\Lambda_i > 0$, if

$$\frac{1 - r(agg.)}{1 - r(cat.)}(\iota_{agg.} - \eta_{agg.}) \quad > \quad \sum_{k=1}^{K} \omega_k(\iota_k - \eta_k)$$

where $\frac{1-r(agg.)}{1-r(cat.)} > 1$. Thus, as long as $(\iota_{agg.} - \eta_{agg.}) \geq \sum_{k=1}^{K} \omega_k(\iota_k - \eta_k)$—i.e., the bias between the mean relevant and irrelevant simulations is equal or larger for aggregate than for the weighted category-specific beliefs—there will be a positive aggregation gap. Assuming a similar degree of interference in the intensive margin, this is sufficient for a positive aggregation gap. For simplicity, we look at the squared aggregation gap Λ_i^2 in order to derive properties for the absolute aggregation gap. The absolute aggregation gap $\mid \Lambda_i \mid = \sqrt{\Lambda_i^2}$ is a strictly monotonous transformation of the squared aggregation gap.

The mean squared aggregation gap is equal to[19]

$$\mathbb{E}\Lambda_i^2 = \mathbb{E}\left[\pi_{agg}^2 + \left(\sum_{k=11}^{11} \omega_k \pi_k\right)^2 - 2\pi_{agg} \sum_{k=11}^{11} \omega_k \pi_k\right] \geq 0$$

[19]We may further simplify to

$$\begin{aligned}
\mathbb{E}\Lambda_i^2 \quad &= \quad \mathbb{E}\left[\pi_{agg}^2 + \left(\sum_{k=11}^{11} \omega_k \pi_k\right)^2 - 2\pi_{agg} \sum_{k=11}^{11} \omega_k \pi_k\right] \\
&= \quad \mathbb{E}\pi_{agg}^2 + \sum_{k=1}^{11} \omega_k^2 \mathbb{E}\pi_k^2 - 2\sum_{k=1}^{K} \omega_k \left[Cov(\pi_{agg}, \pi_k) + \mathbb{E}\pi_k \mathbb{E}\pi_{agg}\right] \\
&= \quad Var(\pi_{agg}) + \sum_{k=1}^{11} \omega_k^2 Var(\pi_k) + (\mathbb{E}\pi_{agg})^2 + \sum_{k=1}^{11} \omega_k^2 (\mathbb{E}\pi_k)^2 - 2\sum_{k=1}^{K} \omega_k \left[Cov(\pi_{agg}, \pi_k) + \mathbb{E}\pi_k \mathbb{E}\pi_{agg}\right] \geq 0
\end{aligned}$$

D.5.1 Subjective Uncertainty

Similar to the above results in 4, we can formally express the correlation between measures of the aggregation gap and subjective uncertainty in Proposition 6.

Proposition 6 (Covariance between the Aggregation Gap and Subjective Uncertainty)**.** *The covariance between the aggregation gap and subjective uncertainty is*

$$Cov\left(\Lambda_i, S_i(agg.)\right) = Cov\left(\pi^i_{agg}, S_i(agg.)\right) - \sum_{k=1}^{K} \omega_k Cov\left(\pi^i_k, S_i(agg.)\right)$$

with $S(agg.)$ denoting the uncertainty about aggregate inflation expectations (see Proposition 4).

The covariance between the squared aggregation gap and subjective uncertainty is

$$\begin{aligned} Cov\left(\Lambda_i^2, S_i(agg.)\right) &= Cov\left(\pi^2_{i,agg}, S_i(agg.)\right) + Cov\left(\left(\sum_{k=1}^{11} \omega_k \pi_{k,i}\right)^2, S_i(agg.)\right) - \\ & 2Cov\left(\pi_{i,agg} \sum_{k=1}^{11} \omega_k \pi_{k,i}, S_i(agg.)\right) \end{aligned} \tag{28}$$

Proof. The proposition follows directly from the above definitions and the definition of covariance. □

In the data, we find that both covariances, between the aggregation gap as well as the absolute (or squared) aggregation gap and subjective uncertainty about aggregate inflation expectations, are positive. The model can also generate this predictions under the assumption that relevant (irrelevant) experience draws for aggregate inflation expectations come from the same distribution for all respondents, such that:

$$\mathbb{E}_i[s(e) \mid e \in H_C] = \eta; \quad Var[s(e) \mid e \in H_C] = \sigma^2_\eta \tag{29}$$

$$\mathbb{E}_i[s(e) \mid e \in H_{\bar{C}}] = \iota; \quad Var[s(e) \mid e \in H_{\bar{C}}] = \sigma^2_\iota \tag{30}$$

The assumption says that the means of category and aggregate relevant simulations are equal across all respondents, and so category and aggregate inflation simulation draws are independent within and across respondents. It follows that there is no covariance between and category-specific expectations and the subjective uncertainty.

$$\begin{aligned} Cov\left(\pi^i_k, S_i(agg.)\right) &= 0 \\ Cov\left(\left(\sum_{k=1}^{11} \omega_k \pi_{k,i}\right)^2, S_i(agg.)\right) &= 0 \end{aligned}$$

Relaxing the simplifying assumption, for example through an appropriate correlation structure, would allow to retain both subsequent and previous results, though at cost of tractability and ease of exposition.

Proposition 7 follows directly:

Proposition 7 (Covariance between the Belief and Subjective Uncertainty)**.** *The covariance between belief π_C and subjective uncertainty $S(C)$ is a function of the skewness of $s(e)$, γ_3.*

$$Cov\left(\pi_C, S(C)\right) = \frac{\gamma_3}{D} \tag{31}$$

Assuming that $s(e) \overset{e \in H_C}{\sim} N(\eta, \sigma_\eta^2)$ and that $s(e) \overset{e \in H_{\bar{C}}}{\sim} N(\iota, \sigma_\iota^2)$ it that follows:

$$Cov\left(\pi_C, S(C)\right) \quad \begin{cases} > 0 \quad if \quad r(H_C) > \frac{1}{2} \; and \; \iota > \eta \\ < 0 \quad if \quad r(H_C) < \frac{1}{2} \; and \; \iota > \eta \end{cases}$$

Proof. See Shanmugam (2008) and Zhang (2007) for a general proof of equation (31). □

Based on these results, one can sign the two expressions in Proposition 6. First, it follows that $Cov\left(\Lambda_i, S_i(agg.)\right)$ simplifies to $Cov\left(\Lambda_i, S_i(agg.)\right) = \frac{\gamma_3}{D}$, which fits the data given that the underlying distribution is right-skewed.

Second, by replacing $\pi_{i,agg}^2$ with its first-order Taylor approximation, $\pi_{i,agg}^2 \approx \mathbb{E}\pi_{agg}^2 + \mathbb{E}\pi_{agg}(\pi_{i,agg} - \mathbb{E}\pi_{agg})$, it follows that:

$$Cov\left(\pi_{i,agg}^2, S_i(agg.)\right) \approx 2\mathbb{E}\pi_{agg} Cov\left(\pi_{i,agg}, S_i(agg.)\right) \tag{32}$$

We apply the same first-order Taylor approximation to $\pi_{i,agg}\pi_{k,i}$ such that $\pi_{i,agg}\pi_{k,i} \approx \mathbb{E}\pi_{agg}\mathbb{E}\pi_k + \mathbb{E}\pi_k(\pi_{i,agg} - \mathbb{E}\pi_{agg}) + \mathbb{E}\pi_{agg}(\pi_{i,k} - \mathbb{E}\pi_k)$. It follows that

$$Cov\left(\pi_{i,agg}\sum_{k=1}^{11}\omega_k\pi_{k,i}, S_i(agg.)\right) = \sum_{k=1}^{11}\omega_k\mathbb{E}\pi_k Cov\left(\pi_{i,agg}, S_i(agg.)\right) \tag{33}$$

In particular, this results implies that if mean aggregate inflation expectations exceed the weighted mean of category expectations – the empirically relevant case – that is, $\mathbb{E}\pi_{agg} > \sum_{k=1}^{11}\omega_k\mathbb{E}\pi_k$, then the covariance $Cov\left(\Lambda_i^2, S_i(agg.)\right)$ is positive.

E Survey Details and Questions

This section lists relevant survey questions used within the paper. The survey was administered on the Qualtrics Research Core Platform, and Qualtrics Research Services recruited participants to provide responses. Survey data used in this paper spans the time from July 9, 2020 to August 9, 2022. Participants were asked for their macroeconomic expectations. While the survey also contains other blocks with various questions, these are not reported here, since they are asked after the questions on macroeconomic expectations and thus do not affect the answers.

Invitations went out to residents of the US Respondents were pre-screened for residence status, English language fluency, and age. All respondents who failed to meet the screening criteria were discontinued from the survey. Only respondents who confirmed residence in the US, who professed English language fluency, and who reported to be of ages 18 or above, were brought into to the survey proper. Once respondents met these criteria, we screened responses by removing any participants who took less than five minutes to complete the survey or had at least one gibberish response (e.g., "$sd - \$rt2$").

E.1 Aggregate Expectations

To elicit respondents' expectations about future aggregate inflation and income, we use the following set of questions. Note that we first ask about participants' point estimates and then collect additional data on the individual distribution of expectations. By this approach, we can gain insights into individual uncertainty. Survey participants are shown the following introductory text:

In some of the following questions, we will ask you to think about the percent chance of something happening in the future. Your answers can range from 0 to 100, where 0 means there is absolutely no chance, and 100 means that it is absolutely certain. For example, numbers like: 2 and 5 percent may indicate "almost no chance" 18 percent or so may mean "not much chance" 47 or 52 percent chance may be a "pretty even chance" 83 percent or so may mean a "very good chance" 95 or 98 percent chance may be "almost certain".

The survey then asks the following question on aggregate inflation over a 12 moths horizon:

Q1: Aggregate Inflation (Point Prediction)
The next few questions are about inflation. Over the next 12 months do you think there will be inflation or deflation?

O Inflation
O Deflation (opposite of inflation)

Depending on the answer given on the previous question, the participant is shown the next question:

What do you expect the rate of ***inflation/deflation*** *to be over the next 12 months? Please give your best guess.*
I expect the rate of ***inflation/deflation*** *to be _____ percent over the next 12 months.*

We choose to ask about point estimates in this twofold manner in order to avoid issues about the correct sign of the numerical answer, i.e. that respondents intend to answer −3 percent but just put 3 in the answer field.

We then ask about the distribution of an individuals' inflation expectation:

QDIST: Aggregate Inflation (Distribution)
Now we would like you to think about what may happen to inflation over the next 12 months. We realize that this question may take a little more effort. In your view, what would you say is the percent chance that, over the next 12 months. . .

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

We then proceed with questions about the expected change in personal household income for the 12-month horizon:

QPHI: Personal Household Income (Point Prediction)
In your view, will the total income of all members of your household (including you), after taxes and deductions, increase or decrease over the next 12 months?

O Increase
O Decrease

Again, depending on the answer given on the previous question, the participant is shown the next question:

By how much do you expect total income of all members of your household to increase over the next 12 months? Please give your best guess.
Over the next 12 months, I expect total income of all members of my household to ***increase/ decrease*** *by _____ percent.*

E.2 Category-Specific Inflation Expectations and Spending

To elicit participants' category-specific inflation expectations, as well as their nominal expenditure on each consumption category and the subjective importance of each category to overall consumption, we ask the following questions:

Q2: Category-Specific Inflation
Twelve months from now, what do you think will have happened to the price of the following items?
I expect the price of ...

Motor vehicles and parts (such as cars and SUVs)	*to [increase/decrease]*	*by*	___
Recreational goods and vehicles (such as sports equipment and laptops)	*to [increase/decrease]*	*by*	___
Other durable goods (such as furniture, appliances, jewelry, luggage)	*to [increase/decrease]*	*by*	___
Food and beverages for off-premises consumption (such as food from grocery stores)	*to [increase/decrease]*	*by*	___
Gasoline and other energy goods	*to [increase/decrease]*	*by*	___
Other nondurable goods (such as clothing, medicine and personal care products)	*to [increase/decrease]*	*by*	___
Housing and utilities (such as rent and utility bills)	*to [increase/decrease]*	*by*	___
Transportation services (such as public transit tickets and airfare)	*to [increase/decrease]*	*by*	___
Food services and accommodations (such as restaurants and hotels)	*to [increase/decrease]*	*by*	___
Other services (such as internet/phone service, education, financial services, hairdressers)	*to [increase/decrease]*	*by*	___

Q3: Nominal Expenditure on Categories

In terms of consumption spending, how much money did you spend on each of the following broad consumption categories during the last month? Please indicate an approximate dollar amount in each field.

Motor vehicles and parts (such as cars and SUVs)	_____
Recreational goods and vehicles (such as sports equipment and laptops)	_____
Other durable goods (such as furniture, appliances, jewelry, luggage)	_____
Food and beverages for off-premises consumption (such as food from grocery stores)	_____
Gasoline and other energy goods	_____
Other nondurable goods (such as clothing, medicine and personal care products)	_____
Housing and utilities (such as rent and utility bills)	_____
Health care	_____
Transportation services (such as public transit tickets and airfare)	_____
Food services and accommodations (such as restaurants and hotels)	_____
Other services (such as internet/phone service, education, financial services, hairdressers)	_____

Q4: Subjective Importance of Categories

Which of the following broad consumption categories matter the most to you right now in your daily life? Please move the slider to indicate the importance for each of them, with 0 indicating no importance and 100 indicating highest importance.

Motor vehicles and parts (such as cars and SUVs)	0___|___100
Recreational goods and vehicles (such as sports equipment and laptops)	0___|___100
Other durable goods (such as furniture, appliances, jewelry, luggage)	0___|___100
Food and beverages for off-premises consumption (such as food from grocery stores)	0___|___100
Gasoline and other energy goods	0___|___100
Other nondurable goods (such as clothing, medicine and personal care products)	0___|___100
Housing and utilities (such as rent and utility bills)	0___|___100
Health care	0___|___100
Transportation services (such as public transit tickets and airfare)	0___|___100
Food services and accommodations (such as restaurants and hotels)	0___|___100
Other services (such as internet/phone service, education, financial services, hairdressers)	0___|___100

E.3 Expected Changes in Consumption Spending

We ask respondents about their expected spending in 12 months, relative to last month with the following questions:

Q5: Total Spending

Compared with your spending last month, how do you expect your total spending to change in the next . . .

	Go Down	No Change	Go Up	By %
. . . month?	*O*	*O*	*O*	______
. . . two months?	*O*	*O*	*O*	______
. . . year?	*O*	*O*	*O*	______
. . . two years?	*O*	*O*	*O*	______

Q6: Services Spending

Compared with your spending last month, how do you expect your spending on services — such as medical and dental care, haircuts, and restaurant meals — to change in the next. . .

	Go Down	No Change	Go Up	By %
. . . month?	*O*	*O*	*O*	______
. . . two months?	*O*	*O*	*O*	______
. . . year?	*O*	*O*	*O*	______
. . . two years?	*O*	*O*	*O*	______

Q7: Nondurable Spending

Compared with last month, how do you expect your spending on nondurable goods—such as clothes, medicine, food at grocery stores, or personal care products—to change in the next. . .

	Go Down by	No Change	Go Up	By %
. . . month?	*O*	*O*	*O*	______
. . . two months?	*O*	*O*	*O*	______
. . . year?	*O*	*O*	*O*	______
. . . two years?	*O*	*O*	*O*	______

E.4 Demographics

To check for demographics and to make the survey representative, we checked for certain demographic characteristics. These include age, gender, ethnicity, state of residence, income, the highest educational level and weather the respondent is the main grocery shopper in its household..