Federal Reserve Bank of Cleveland Working Paper Series

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

Alexander M. Dietrich, Edward S. Knotek II, Kristian Ove R. Myrseth, Robert W. Rich, Raphael S. Schoenle, Michael Weber

Working Paper No. 22-20

June 2022

Suggested citation: Dietrich, Alexander M., Edward S. Knotek II, Kristian Ove R. Myrseth, Robert W. Rich, Raphael S. Schoenle, and Michael Weber. 2022. "Greater Than the Sum of the Parts: Aggregate vs. Aggregated Inflation Expectations." Working Paper No. 22-20. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202220.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

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

Alexander M. Dietrich, Edward S. Knotek II, Kristian Ove R. Myrseth, Robert W. Rich, Raphael S. Schoenle, and Michael Weber*

This version: June 2022

—Working Paper—

Abstract

Using novel survey evidence on consumer inflation expectations disaggregated by personal consumption expenditure (PCE) categories, we document the paradox that consumers' aggregate inflation expectations usually exceed any individual category expectation. We explore procedures for aggregating category inflation expectations, and find that the inconsistency between aggregate and aggregated inflation expectations rises with subjective uncertainty and is systematically related to socioeconomic characteristics. Overall, our results are inconsistent with the notion that consumers' aggregate inflation expectations comprise an expenditure-weighted sum of category beliefs. Moreover, aggregated inflation expectations explain a greater share of planned consumer spending than aggregate inflation expectations.

Keywords: Household expectations, Survey, Sectoral expectations, Inflation expectations
JEL-Codes: C83, E31, E52

*Dietrich: University of Tübingen, Email: alexander.dietrich@uni-tuebingen.de; Knotek: Federal Reserve Bank of Cleveland, Email: edward.knotek@clev.frb.org; Myrseth: University of York, Email: kristian.myrseth@york.ac.uk; Rich: Federal Reserve Bank of Cleveland, Email: robert.rich@clev.frb.org; Schoenle: Brandeis University, CEPR and CESifo, Email: schoenle@brandeis.edu; Weber: University of Chicago Booth School of Business, Email: michael.weber@chicagobooth.edu. We thank Hassan Afrouzi, Charles Manski, Andrei Schleifer, Wilbert van der Klaauw, as well as seminar participants at the Bundesbank, Bank of Finland, ICEA, SEA, 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 or the Board of Governors of the Federal Reserve System.

1 Introduction

Although inflation expectations are crucial for a central bank's interest rate policy, consumers struggle to grasp the inflation concept and appear to rely on salient information, such as observed changes in grocery prices, to form their expectations (D'Acunto et al., 2021b). Motivated both by recent empirical findings and a psychological literature on limitations to human judgment (Tversky and Kahneman, 1974), this paper explores how aggregate inflation expectations conventionally elicited compare to more specific expectations for the categories of personal consumption expenditures (PCE). To do so, we use a novel, large scale survey of US households that elicits both aggregate and category-specific inflation expectations, on the basis of which we construct *aggregated* inflation expectations comparable to reported aggregate expectations.

Our paper offers several new insights: First, we document that consumers' aggregate inflation expectations tend to exceed any category expectation; they are more volatile in the time series; and they exhibit more cross-sectional disagreement. Second, the category-based aggregated measures are less overstated and dispersed both in the cross-section and across time compared to aggregate inflation expectations. Moreover, among different aggregation procedures, non-core inflation aggregations—the expenditure-weighted average of gasoline and grocery price expectations—seem to align most closely with aggregate inflation expectations in the cross-section, whereas expenditure-weighted category aggregations differ significantly. Third, the respondent-specific inconsistency between aggregate and aggregated inflation expectations rises significantly with subjective uncertainty about aggregate inflation and correlates in a meaningful way with socioeconomic characteristics. These findings are consistent with the conjecture that respondents find questions about category-specific inflation less complex and easier to understand than questions about aggregate inflation. Fourth, in estimations of the consumer Euler equation, aggregated inflation expectations represent a better predictor of planned consumer spending than do aggregate inflation expectations. Notably, non-core aggregation stands out as the best predictor of planned spending over the next 12 months. We conclude that category-specific elicitations of inflation expectations in household surveys, compared to conventional procedures for eliciting aggregate inflation expectations, hold promise to provide a more informative view on the economy and a tool to improve forecasts of economic behavior.

We collect US household expectations in a nationally representative survey between July 2020 and September 2021, as part of the Federal Reserve Bank of Cleveland's daily survey of consumers (Knotek et al., 2020). Elicitation of aggregate inflation expectations follows the conventional point-estimate approach from the New York Fed's Survey of Consumer Expectations (SCE), whereupon the survey presents a novel elicitation of inflation expectations for each of 11 distinct consumption categories, covering the full PCE range. In so doing, we closely match the SCE format for aggregate inflation expectations. In addition, we also ask survey participants about personal expenditures and the relative importance of the consumption categories.

Our analysis explores a range of different procedures for aggregating category inflation expectations. The procedures can be broadly classified into two groups: (i) plausibly rational aggregations and (ii) behavioral aggregations. The three aggregations within the first group use weights that can be understood as reasonable for a rational agent: self-reported expenditure weights, self-reported importance weights, and the official PCE weights. By contrast, the behavioral aggregations capture weighting schemes that depart from rational expectations in favor of behavioral agents, who might use heuristic judgment and hold internally inconsistent beliefs. Such a behavioral agent, for example, might pick up on salient information and report different beliefs depending on how questions are framed. The fourth aggregation sets equal weights to capture agents who can average price changes, but struggle to differentiate weights. Other behavioral combinations include core and non-core inflation expectations, a max operator that picks the highest category expectation, and a second-max operator that picks the second highest. Aggregate inflation expectations, therefore, would not necessarily correspond to the aggregated inflation expectations based on PCE category inflation expectations

Of the aggregations investigated in a simple regression analysis, the model relying on equal weights for each category explains the largest share of the variance in reported aggregate expectations (highest R^2). Moreover, the model using non-core inflation expectations—an expenditure-weighted average of gasoline and grocery inflation expectations—as a predictor produces the best fit (lowest AIC). This might be because the mean of non-core inflation expectations is very close to the reported aggregate mean.

When considering the aggregation inconsistency—the gap between the reported aggregate expectations and aggregated measures—we find that the cross-section mean differs significantly from zero. Absolute inconsistency values vary meaningfully with socioeconomic characteristics; higher education results in a much closer alignment of the reported aggregate expectations and aggregated measures. Both subjective uncertainty about aggregate inflation expectations and the individual dispersion of category expectations correlate strongly with the absolute aggregation inconsistency. We interpret this as evidence that the more uncertain households are about their aggregate forecast, and the less aligned various expected price changes in the economy are, the more the complexity of the aggregation task bears on consumers' ability to perform the aggregation consistently. This resonates with psychological theory suggesting that reliance on cognitive heuristics in judgment and choice increases with uncertainty and complexity. Moreover, we find that the non-core aggregation offers the strongest predictor of spending plans, also consistently outperforming aggregate inflation expectations. This means that non-core aggregation may not only best capture what consumers report when they respond to the conventional question format for aggregate inflation expectations, but also provides policy-relevant information beyond what is gleaned from aggregate inflation expectations.

Finally, our results have implications for theoretical work on the rationality of aggregate inflation expectations. Prior work by Coibion and Gorodnichenko (2012) examining expectations and realizations has rejected the hypothesis of full information and rational expectations (FIRE). Our work addresses the issue from a different angle, demonstrating the internal inconsistency of expectations at the individual level; the internal inconsistency between reasonable linear aggregation models and aggregate inflation expectations speaks against the full rationality of consumer expectations.

Our work builds on a burgeoning literature on consumer inflation expectations. Within the Survey of Consumer Expectations (SCE), inflation expectations are elicited for several salient consumption categories, such as gasoline, rent, and groceries. Using SCE data, Armantier et al. (2016) find that updating in short-term aggregate expectations is consistent with updating at the category-level. Moreover, Bruin et al. (2011) provide evidence that households rely on salient, extreme prices to form their aggregate inflation expectations, and Bruin et al. (2010) show that financial literacy seems to be a driver of inflation expectations, partially explaining demographic patterns. A growing body of research investigates the impact on aggregate inflation expectations of price changes in individual product categories. For example, D'Acunto et al. (2021b) find that consumers rely on observed changes in grocery prices to form their aggregate inflation expectations and that the relative weight products receive depends on the frequency of purchase, rather than expenditure. Binder and Makridis (2022), Binder (2018), Coibion and Gorodnichenko (2015), and Trehan (2011) show similar results of extrapolation from a single consumption category, gasoline, for aggregate inflation expectations. Notably, neither groceries nor gasoline form part of core inflation, and Arora et al. (2013) find that aggregate inflation expectations react excessively to non-core price changes. Analyzing the implications for monetary policy, Dietrich (2022) uses the same data as this paper to document heterogeneity in category-based expectations formation, with households relatively more attentive to their internal food and energy inflation forecasts. Moreover, past experiences seem to impact expectations about future macroeconomic conditions (Malmendier and Nagel, 2011), and Kuchler and Zafar (2019) find that individuals forming forecasts on housing prices extrapolate from recent locally experienced developments. In this paper, we do not consider how observed price changes impact expected aggregate inflation. Rather, we relate expected inflation for consumption categories to expected aggregate inflation and, crucially, planned consumption. In addition, we take into account the full range of household consumption, instead of only (salient) subsets.

Like our paper, several others have investigated inconsistencies between question types in surveys asking about aggregate inflation expectations. Engelberg et al. (2009) find that in the Survey of Professional Forecasters (SPF), point forecasts in general are consistent with forecasts from probability distributions. In contrast, the share of inconsistent respondents among households is considerably higher, about one third, and the probability of an inconsistent answer rises with lower socioeconomic status (Stanisławska et al., 2021), similar to our findings. Whereas this literature defines consistency according to whether a point estimate falls within a probability distribution or fits a qualitative assessment of future inflation, our paper considers the distance between two point estimates: between the reported aggregate and the *aggregated* inflation expectations.

Methodologically, our paper is related to a literature that elicits past consumption—but not inflation expectations—using disaggregated survey questions. Winter (2004) finds that for non-durable consumption, disaggregated questions yield improved data quality over questions asking about aggregates, and discrepancies vary with socioeconomic characteristics, similar to what we find for the variation in aggregation inconsistency. 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. Moreover, Hurd and Rohwedder (2008, 2012) field surveys to ask households about past spending using disaggregated category questions. In the spirit of this literature, our survey elicits past expenditure with questions disaggregated by PCE consumption categories.

Our paper proceeds as follows: Section 2 outlines a theory of behavioral inflation forecasts. Section 3 describes our novel survey data. Section 4 examines category inflation expectations and compares them to aggregate inflation expectations. Section 5 investigates procedures for aggregating category inflation expectations and the inconsistency between aggregate and aggregated inflation expectations. Section 6 compares how the measures of expectations fare in explaining household spending plans in Euler equation estimations. A final section concludes.

2 Human Forecasts and Inflation Expectations

When household surveys ask respondents to report their inflation expectations, they are in effect asking for forecasts of an uncertain, abstract variable. The canonical work by Tversky and Kahneman (1974) on heuristics and biases, however, shows that the human mind isn't optimally wired for the task; judgments of uncertain events rely on heuristics—simple rules of thumb—which often lead to predictable discrepancies from rational norms. A common manifestation of this is the salience bias, whereby human judgment is biased by salient information. For example, consumers exposed to price spikes in their grocery bundles may report higher inflation expectations (D'Acunto et al., 2021b), or their expectations may reflect their expenditure bundles (Cavallo et al., 2017). A similar phenomenon, driven by the representativeness heuristic of Kahneman and Tversky (1972), is formalized by Gennaioli and Shleifer (2010) and applied by Bordalo et al. (2018) to model credit cycles. Moreover, Bordalo et al. (2022) show that selective, automatic memory can account for both over- and underestimation of novel risk.

An implication is that human judgment also is notoriously inconsistent (Fischhoff and Broomell, 2020). This has been demonstrated, for example, in experimental tests of the conjunction rule, the basic principle that the probability of events "A" and "B" both being true cannot exceed that of "A" alone being true. A classic class-room illustration is the "Linda problem" (Tversky and

Kahneman, 1983), in which participants read a vignette about Linda and judge the likelihood of the more specific description, that she is both a bank teller and active in the feminist movement, to be higher than the more general description, that she is a bank teller. Inflation expectations, which involve probabilistic beliefs and many uncertain factors, could similarly fall prey to the conjunction fallacy or similar phenomena. Confronted with a question about aggregate inflation expectations, for example, an individual might imagine a series of salient (but unlikely) events and exaggerate the likelihood that they all transpire; this, in turn, could yield reported expectations biased upwards.

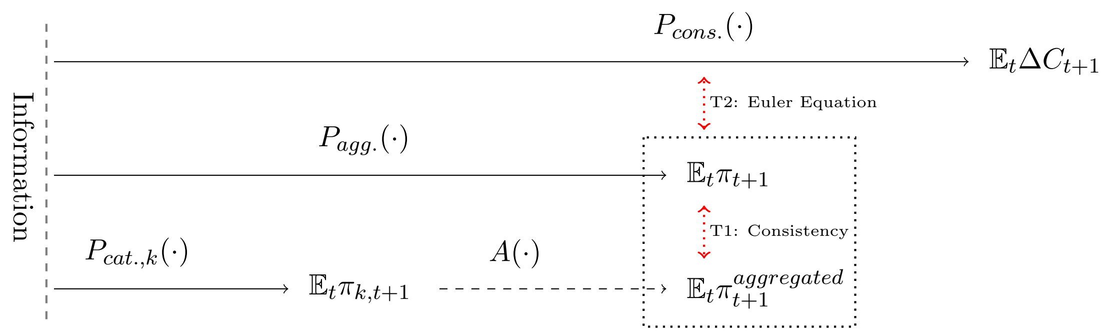
Even experts struggle to incorporate multiple cues into a reliable forecast. Starting with the influential work of Meehl (1954), psychologists discovered that clinical expert forecasts—that is, forecasts based on expert intuition—were surprisingly unreliable across a wide range of domains and were consistently outperformed by rudimentary statistical models. Subsequent work by Dawes (1979) found that linear models with arbitrary weights–including equal weights–outperformed expert human judgment; as long as linear models include the relevant predictor variables, with coefficients set in the correct direction, they prove surprisingly robust (Einhorn and Hogarth, 1975). These findings have held up over time (Dawes et al., 1989), and although they apply to experts, there is little reason to think that lay respondents would perform any better. In fact, professional forecasts of inflation consistently outperform those of lay households (Carroll, 2003; Verbrugge and Zaman, 2021).

By eliciting both category and aggregate inflation forecasts, our analysis adds a unique angle to the study of forecast consistency and aggregation. First, respondents are answering questions about something tangible and concrete, of which they may have better understanding, presumably leaving them less vulnerable to both biased and noisy judgments when providing category inflation forecasts (as opposed to aggregate inflation forecasts). Second, our setup provides the opportunity to combine category forecasts mechanically into "bottom-up," aggregated inflation expectations. Such *aggregated* inflation expectations, compared to explicitly articulated expectations of aggregate inflation, might better represent respondents' effective inflation beliefs–which they may not necessarily articulate explicitly, but act as if they hold. The reason for this is that the mere act of articulating the abstract inflation concept–which most respondents don't really understand–could involve cognitive distortion causing both bias and noise.

2.1 Model of Inflation Expectations

To formalize our thinking about the process by which inflation expectations form, we introduce a simple model. The model posits that an economic agent uses her set of information S_t (i) to form expectations and (ii) to make economic plans. Crucially, we allow her to hold inconsistent expectation formation processes; the beliefs on which she acts—*effective* expectations—need not be consistent with aggregate inflation expectations elicited, nor with the category-based expectations aggregated. The intuition is that explicit elaboration of a question posed in a survey, such as asking

Figure 2.1: Inflation Expectations - Formation and Revelation



Notes: The figure shows the proposed mechanism for the formation of expectations for category and aggregate inflation as well as planned consumption spending. Red arrows denote empirically testable relations: the consistency between aggregate and aggregated inflation expectations and the connection of inflation expectations to spending plans, as suggested by the Euler equation.

about inflation expectations, may trigger processes entirely distinct from those implicit when the agent makes consumption decisions. Put differently, the agent doesn't necessarily ask herself what her one-year inflation forecast is prior to each consumption decision. Yet, she will act *as if* she has *some* belief, and one question we pose in this paper is whether the aggregate inflation expectation is the most accurate representation or whether some aggregation principle applied to the category expectations can offer improvement.

Figure 2.1 shows how we assume that expectations form. $\mathbb{E}_t \, \Delta C_{t+1}$ denotes expected consumption expenditure changes, $\mathbb{E}_t \, \pi_{t+1}$ expected aggregate inflation, and $\mathbb{E}_t \, \pi_{k,t+1}$ expected inflation in category k over the next 12 months. Expectations are formed with the mental processes $P_{cons.}(\cdot)$ for consumption plans, $P_{agg.}(\cdot)$ for aggregate inflation, and $P_{cat.,k}(\cdot)$ for category k inflation expectations. Crucially, those mental processes may be connected, but need not be. Using the category expectations, we can apply a mechanical aggregation process $A(\cdot)$ to form an *aggregated* expectation $\mathbb{E}_{t+1} \pi_{t+1}^{aggregated}$. This could, for example, be a weighting scheme using expenditure shares or one using PCE weights. Note that the process governing the expected change in consumption spending reveals effective inflation expectations; the Euler equation in Section 6 formalizes this.

A strength of our analysis is that we remain agnostic about certain aspects of the expectations formation processes. While we allow internally consistent beliefs, we do not require them. For example, the agent may think about the aggregate inflation concept in terms distinct from how she considers inflation for each separate category, even in response to the same information S_t. Categories, tangible and concrete, may well be more intuitive than the aggregate inflation concept, leading to differential cognitive heuristics being used.

Our objective is to exploit these degrees of freedom and to apply aggregation principles A() to the category inflation expectations in order to gain insight into the process by which aggregate inflation expectations form. To that end, we test the consistency between aggregate and *aggregated* inflation expectations (illustrated by T1 in Figure 2.1) across a range of aggregation techniques. Next, we test whether *aggregated* inflation expectations predict consumption plans better than do aggregate inflation expectations (illustrated by T2 in Figure 2.1), and whether a plausibly rational aggregation technique, such as weighting by self-reported expenditure or official PCE weights, tracks aggregate inflation expectations better than does a more behavioral alternative, such as a heuristic max operator.

Among the aggregation techniques tested, non-core inflation expectations appear to be the least inconsistent with aggregate inflation expectations (Section 5) while also offering the best predictor of planned consumption (Section 6). In other words, aggregated non-core inflation expectations may represent a closer (even if imperfect) approximation of consumers' effective beliefs than do both the alternative aggregation mechanisms and aggregate inflation expectations, even as the aggregation technique most consistent with the latter.

3 Survey

Our survey module is part of a larger daily survey of consumer expectations maintained by the Federal Reserve Bank of Cleveland and administered by Qualtrics Research Services. It includes a nationally representative sample of 17,888 responses, collected between July 9, 2020 and September 9, 2021. Dietrich et al. (2022) and Knotek et al. (2020) provide further information about the survey of which our module is a part. 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).

The survey was run in real time, with a daily sampling size of at least 100 respondents. We required all respondents to be US residents and to speak English as their primary language. Respondents were representative of the US population according to several key demographic and socioeconomic characteristics; respondents had to be male or female with 50 percent probability; approximately one third were 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 required 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 included filters to eliminate respondents who enter gibberish for at least one response, or who complete the survey in less (more) than five (30) minutes, and CAPTCHA tests to reduce the likelihood that bots would interfere.[1]

Table 1 provides a breakdown of our sample, showing that our sampling criteria generated a sample roughly representative of the US population along key dimensions. To improve the fit further, we compute a survey weight for each respondent; we apply iterative proportional fitting to

[1] Qualtrics Research Services provides the filtered data. The daily sample size refers to the number of respondents after filtering. Survey respondents are provided with fair monetary compensation for their time.

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=17,888		

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).

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—that is, the variables in the right-hand column of Table 1.

Within the survey, we asked respondents first about their aggregate inflation expectations over the next 12 months (Q1 in Table 2), using point forecast questions.[2] Our approach to eliciting aggregate inflation forecasts is methodologically similar to that of other influential household surveys, such as the University of Michigan's Surveys of Consumers (SoC) and the New York Fed's Survey of Consumer Expectations (SCE).[3] Subsequently, we elicited inflation expectations for 11 PCE categories (Q2 in Table 2). Table 3 in Section 4 shows both the PCE categories used in the survey and some summary statistics. The PCE disaggregation used in our survey is based on the US national income and product accounts (NIPA) disaggregation, with some small sectors combined in order to reduce the cognitive burden of completing the survey.[4] Dietrich, 2022 provides more

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

[3] The SoC has collected data on household inflation expectations since 1978; the SCE started in 2013. Both ask about aggregate inflation or the expected change in *aggregate* prices directly, at a monthly frequency, and they include some kind of panel structure; while the SoC asks a subset of participants to answer 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 repeated cross-section, but is conducted at a higher, daily frequency.

[4] We use what might be thought of as the third level of disaggregation of PCE spending—the first being between goods and services, and the second durable and nondurable goods and expenditures on services by households and

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 refer to Appendix B.

details on categories. While aggregate forecasts within the SCE are also elicited using a probability distribution question, we choose to rely on point forecasts for both the aggregate and the category expectations.[5] First, this reduces the mental burden on survey participants. Second, Clements (2014) finds that point forecasts offer superior data quality over that obtained from probability distribution questions when one is concerned with the mean of expectations.

nonprofit institutions serving households.

[5]We use a probability distribution question in the survey only to elicit the subjective uncertainty of survey participants about aggregate inflation. Subjective uncertainty measures are a unique feature that only probability distribution formats offer.

Table 3: Summary Statistics

	Cross-Section		Time Series
	Mean	Std. Dev.	Std. Dev.
Aggregate expectation	5.16	7.59	2.86
Category expectations			
Motor vehicles	4.56	6.61	1.89
Recreational goods	3.24	6.52	1.81
Other durable goods	3.21	6.05	1.87
Food and beverages	4.91	6.90	1.94
Gasoline	4.58	7.33	2.31
Other nondurable goods	3.57	5.92	1.56
Housing and utilities	4.84	7.02	1.83
Health care	3.19	7.15	1.72
Transportation services	4.29	6.68	1.68
Food services	4.23	7.05	1.72
Other services	3.93	5.76	1.44

Notes: This table presents summary statistics on the Huber-robust and survey-weighted mean on expectations, the standard deviation in the cross-section, and the standard deviation in the (daily mean) time series.

Besides inflation expectations within these sectors, we also asked how much survey respondents spent within the respective sector during the last month (Q3 in Table 2) and how "important" they consider it for aggregate inflation (Q4 in Table 2). Responses to these questions allow us to compute both expenditure shares per sector (relative to total expenditure) and a measure of perceived relative importance.

Following questions about category expectations and expenditure shares, respondents were asked about their expected spending relative to spending in the month prior, looking ahead 12 months. This question was also repeated for other, more narrowly defined spending categories, such as services spending and expenditures on nondurable consumption goods. Additionally, respondents were asked about their socioeconomic background and consumer habits. For those questions, such as demographic information as well as the exact layout of our inflation questions, please refer to Appendix B.

4 Aggregate vs. Category Inflation Expectations

In this section, we compare the statistical properties of reported aggregate inflation (Q1 in Table 2) and category inflation expectations (Q2). We find that mean expectations about aggregate inflation in the cross-section exceed inflation expectations for every PCE category. In addition, there is both larger disagreement (cross-sectional standard deviation) among consumers and more volatility within the time series for reported aggregate expectations relative to individual categories.

Table 3 shows summary statistics for aggregate inflation expectations and category expectations. The table reports the mean expectation and disagreement among households (cross-section standard deviation) in the first and second columns; the time-series standard deviation, in the third column, represents the volatility over time–that is, the standard deviation of daily mean estimates. Survey participants expect inflation over the next 12 months to be 5.16 percent on average, between July 2020 and September 2021. Still, in the same period, every category inflation rate is expected to be lower: From 3.19 percent for "Health care services" to 4.91 percent for "Food and beverages." Thus, for a representative agent whose views mirror those of the cross-section, expected aggregate inflation exceeds any component. This result is driven by respondents with aggregate expectations outside the range of their individual category expectations. At a micro-level, about 25 percent of respondents state a larger aggregate expectation than for any category. For 15 percent of respondents, the opposite holds true; they assume aggregate expectations below their smallest category expectation. Only around 60 percent of respondents place their aggregate within the range of their category expectations.

The upper row of Figure 4.1 shows the time series, by daily means, for aggregate and mean category inflation expectations during the survey period. The left panel displays category expectations for the durable (red lines) and nondurable (blue lines) consumption goods, while the right panel shows services categories (green lines). All time series displayed are balanced 11-day moving averages.[6] Aggregate inflation expectations reported by consumers are higher than any category expectations for most periods in our sample. Consequently, for a representative agent, there exists no possible linear combination of category expectations with non-negative aggregation weights that maps category expectations into aggregate expectations.

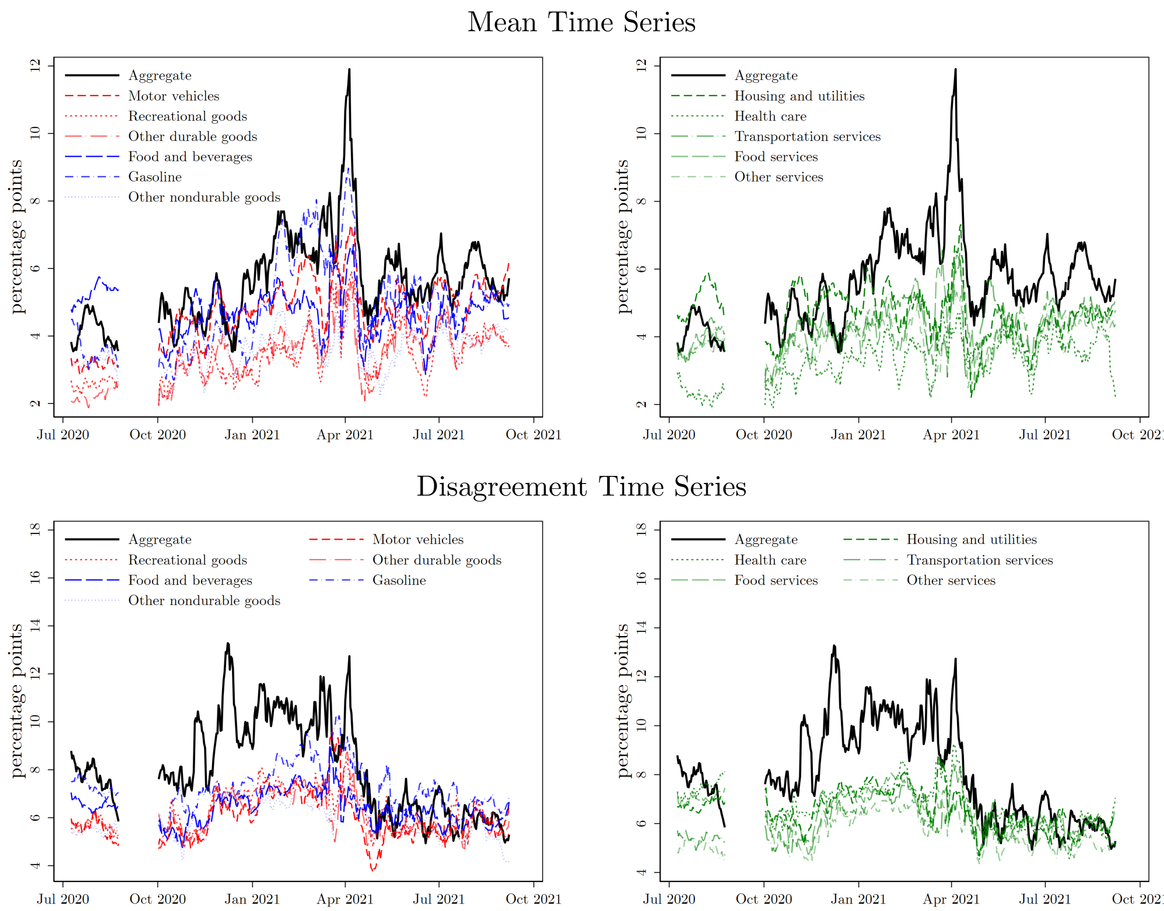
The bottom row of Figure 4.1 shows disagreement among respondents for aggregate inflation expectations (black line) and category expectations, where we measure disagreement as the daily standard deviation of the cross-section. The figures display an 11-day moving average, with durable and nondurable goods sectors in the left panel and services in the right panel. For most of the time surveyed, disagreement is much higher for aggregate expectations than it is for more narrowly defined category expectations (see also Table 3).

Time-series volatility (of expectations) is an important moment in economic analysis. Here, we find that volatility over time is higher for aggregate inflation expectations than it is for individual category expectations (see Column 3 in Table 3).

Tables 9 and 10 in the Appendix reveal demographic heterogeneity in the results; we find that 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 expectations tend to be quite similar across education and income and, where they are not, they do not

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

Figure 4.1: Aggregate vs Category Inflation Expectations



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 balanced moving average. Underlying daily observations are Huber-robust and survey-weighted means.

diverge in a consistent fashion–unlike aggregate inflation expectations, which fall with education and income. Across almost all categories, as well as the reported aggregate, women and respondents who identify themselves as the primary grocery shoppers in their households display higher inflation expectations and greater disagreement. This is generally consistent with demographic patterns found by Bruin et al. (2010). An inconsistent pattern, however, arises with age: For the oldest age group in our sample (older than 55), aggregate inflation expectations are lower than those of younger respondents. For expectations by category, in contrast, the pattern is reversed: older respondents report higher expectations.

5 Aggregate vs. Aggregated Inflation Expectations

Next, we study the relationship between reported aggregate inflation expectations and aggregated measures of the categories that also describe overall inflation. Section 5.1 introduces the aggregation methods, Section 5.2 the statistical properties of aggregated inflation expectations relative to aggregate expectations, and Section 5.3 the inconsistency between aggregate inflation expectations and the aggregated measures. We find that aggregated measures of inflation tend to be closer to zero than aggregate expectations, and disagreement among survey participants is higher for the latter. The statistically significant, positive aggregation inconsistency is particularly noteworthy for expenditure- and PCE-weighted aggregations as it reflects internally inconsistent beliefs about inflation. The inconsistency increases with uncertainty and varies in a meaningful way with socioeconomic and demographic characteristics.

5.1 Aggregated Inflation Expectations

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

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

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

Table 4 summarizes different sets of weights used in this paper, and we start by outlining those that could plausibly describe a rational agent. The first weights category inflation expectations with the self-reported expenditure shares. The second uses weights derived from questions asking respondents to indicate the qualitative "importance" of each category for their consumption. And the third relies on the official monthly BEA nominal expenditure shares used to construct the PCE inflation statistics. The remaining five sets of weights, in contrast, represent some form of "behavioral" expectations formation. The first sets equal weights, reflecting an agent who notices price changes but neglects expenditure shares. The 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. The third is the inverse of the aforementioned, reflecting an agent who pays attention to non-core inflation. The non-core weights are motivated by earlier work, which demonstrates the salience of non-core prices for households, such as D'Acunto et al. (2021b) 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. The fourth and fifth

Table 4: Aggregated Expectations - Weights

$\mathbb{E}_t^i\,\pi_{t+1}^{aggregated}$	Weights ω_k	Notes
Plausibly rational aggregation		
$\mathbb{E}_t^i\,\pi_{t+1}^{exp}$	$\omega_k^i = \frac{C_{k,t}^i}{\sum_{k=1}^N C_{k,t}^i} \quad \forall k$	**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.
$\mathbb{E}_t^i\,\pi_{t+1}^{PCE}$	$\omega_k = \frac{C_{k,t}^{PCE}}{\sum_{k=1}^N C_{k,t}^{PCE}} \quad \forall k \forall i$	**PCE weights**; $C_{k,t}^{PCE}$ denotes monthly PCE expenditure from BEA.
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.

sets 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. Crucially, both measures are irrespective of the particular category or categories that receive the highest or second-highest expectation. These are motivated by Bruin et al. (2011), who find that extreme inflation rates play an important role in household expectations.[7]

5.2 Statistical Properties of Aggregated Inflation Expectations

Table 5 provides summary statistics. The mean aggregate inflation expectation exceeds those of all three plausibly rational aggregations; it matches that of the non-core aggregation, and it is lower than those of both max operators. In the cross-section, the standard deviation of aggregate

[7] In the Appendix, we also use a constrained regression to determine the implied weights that create the closest match between aggregate and aggregated inflation expectations, subject to non-negative weights summing up to unity. Those weights can provide additional insights into the categories on which consumers rely when forming their aggregate expectations. Section A.4 provides further details.

Table 5: Summary Statistics

	Cross-Section		Time Series
	Mean	Std. Dev.	Std. Dev.
Aggregate expectation	5.16	7.59	2.03
Aggregated expectations			
Plausibly rational aggregation			
Expenditure weights	4.50	5.19	1.14
Importance weights	3.97	4.44	1.02
PCE weights	3.93	4.39	1.02
Behavioral aggregation			
Equal weights	3.79	4.25	0.97
Core inflation	4.30	5.29	1.14
Non-core inflation	5.17	6.02	1.33
Max	10.37	7.54	2.34
Second max	6.64	6.96	1.43

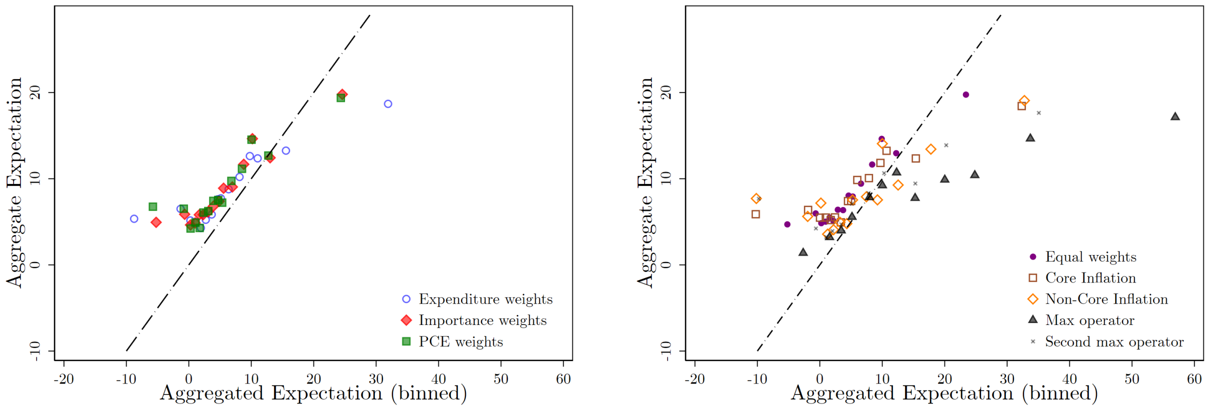
Notes: This table presents summary statistics on the Huber-robust and survey-weighted mean of expectations, the standard deviation in the cross-section, and the time series standard deviation (std. dev. of time series shown in Figure 5.2).

inflation expectations is matched only by that of the max operator. Similarly, in the time-series dimension, aggregate inflation expectations and the max operator yield the two highest standard deviations.

Figure 5.1 compares aggregate inflation expectations against *aggregated* expectations in the cross-section, plotting on the horizontal axis (binned) measures of the latter, with the vertical axis giving the mean of aggregate inflation expectations for each respective bin. Two features stand out. First, almost all observations are above the 45° line, indicating that aggregate inflation expectations tend to be higher than aggregated measures. This, 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 threshold, more extreme aggregated expectations correspond to only slightly more extreme aggregate expectations.

In Table 6, 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 equal weights, showing that they explain the largest share of variation in reported aggregate expectations. The Akaike information criterion (AIC), however, is minimized for non-core expectations, indicating the best model fit of the eight measures of aggregated expectations. Finally, when we run the regression model with all aggregation measures on the right-hand side, we find that both fully saturated model aggregation measures and restricted aggregation expectations turn out to be significant. Both the R^2 and the AIC criterion suggest this fully saturated specification provides

Figure 5.1: 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.

the best fit to aggregate inflation expectations.

Figure 5.2 focuses on the times series properties of aggregate and aggregated expectations, abstracting from cross-sectional variation. Aggregate inflation expectations generally exceed aggregations by expenditure, importance, equal, and PCE weights (top left panel), but are exceeded by the max operator (top right panel); the remaining aggregations appear to cluster around aggregate inflation expectations. The bottom row of Figure 5.2 shows that disagreement, measured as the daily cross-sectional standard deviation of expectations, in aggregate inflation expectations consistently exceeds that in expenditure, importance, equal, and PCE aggregations (bottom left panel), and, until about April 2021, that in core, non-core, and second-max aggregations (bottom right panel)—after which it roughly coincides with disagreement in the latter three aggregations.

5.3 Inconsistency between Aggregate and Aggregated Inflation Expectations

In this section, we examine further the relationship at the individual level between aggregate and aggregated inflation expectations. For this purpose, we define the *aggregation inconsistency* as the difference between the aggregate expectation and any aggregated measure based on the category inflation expectations.

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

Table 6: Aggregate vs. Aggregated Inflation Expectations

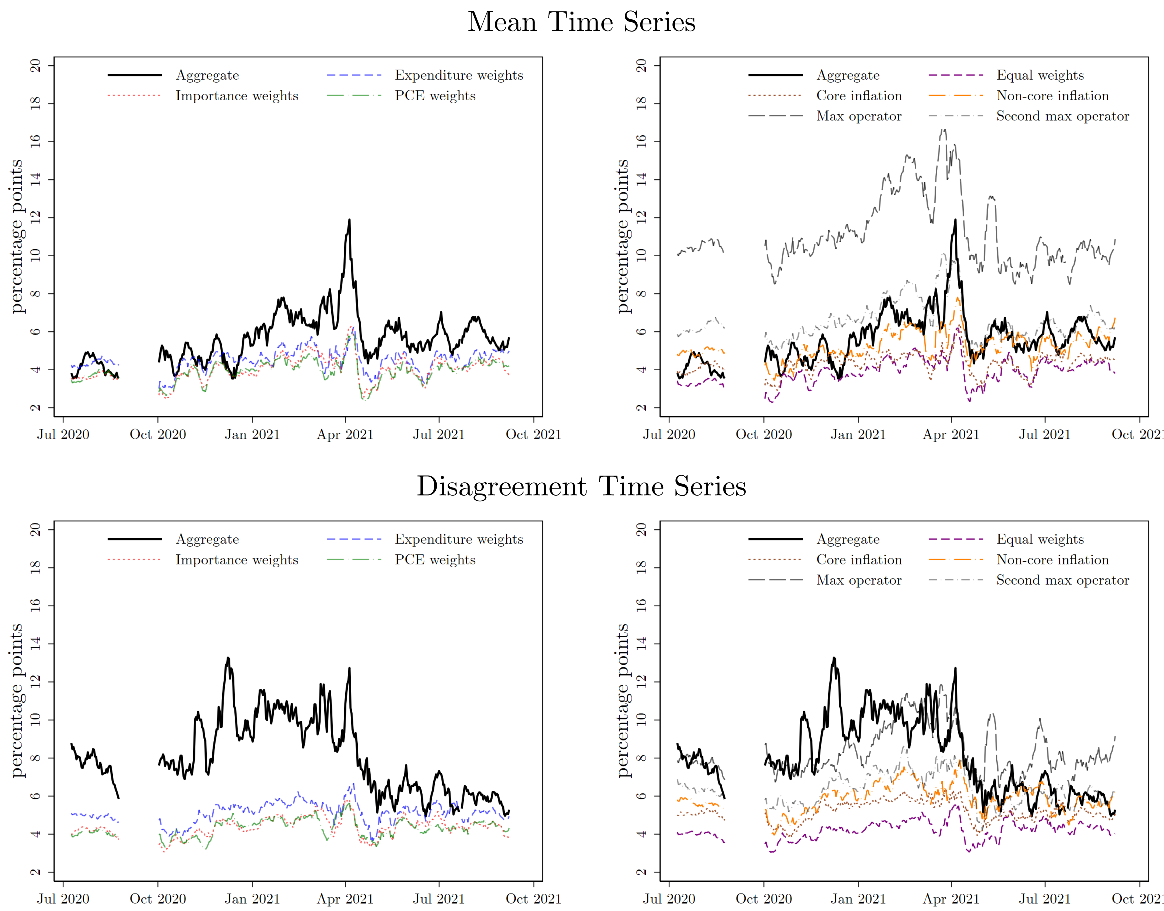
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Expenditure	0.420***								-0.0206
	(36.77)								(-0.42)
Importance		0.605***							-0.0171
		(43.45)							(-0.27)
PCE			0.615***						0.187***
			(44.76)						(3.78)
Equal				0.643***					0.386***
				(45.93)					(5.40)
Core Inflation					0.357***				-0.0350
					(30.99)				(-0.90)
Non-core inflation						0.377***			0.0962***
						(36.05)			(4.50)
Max							0.169***		-0.00731
							(24.09)		(-0.69)
Second max								0.335***	0.0648***
								(35.08)	(3.76)
Constant	3.265***	2.718***	2.682***	2.636***	3.615***	3.253***	3.369***	3.109***	2.272***
	(39.31)	(31.81)	(32.11)	(31.91)	(43.92)	(39.26)	(33.88)	(35.96)	(24.06)
N	15989	15965	15978	15972	15620	14227	16180	15996	13818
R^2	0.176	0.209	0.217	0.223	0.142	0.165	0.0770	0.151	0.246
AIC	108946	108617	108076	107751	106541	95514	112495	108984	92128

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.

Λ_i defines the aggregation inconsistency for survey participant i as the difference between his or her aggregate forecast $\mathbb{E}_t^i\,\pi_{t+1}$ and an aggregated expectation measure $\mathbb{E}_t^i\,\pi_{t+1}^{aggregated}$.

Table 7 displays Huber-robust and survey-weighted estimates, across all individuals in our sample, for the mean aggregation inconsistency and the mean absolute aggregation inconsistency for our range of aggregations. The mean aggregation inconsistency allows us to gauge the direction of the discrepancy, whereas the mean absolute aggregation inconsistency gives us the average discrepancy, irrespective of sign. The plausibly rational aggregations all yield a positive mean aggregation inconsistency, implying that aggregate inflation expectations on average exceed aggregated expectations. This result is noteworthy, especially for expenditure and PCE weights, as it rejects the idea that the reported aggregate simply represents a mental process summing categories by either self-reported expenditure shares or official PCE weights. While for an individual survey participant $\Lambda_i \neq 0$ might be explained by noise in reporting, it cannot be explained as such in the cross-section, as the estimated mean is significantly different from zero. The direction is reversed for the max op-

Figure 5.2: Aggregate vs Aggregated Measures



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 balanced 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.

erators, which is not surprising given that, by construction they select the highest or second-highest category inflation expectation. Nevertheless, the mean aggregation inconsistency for the *second*-max operator is relatively small, comparable in magnitude to that for the aggregation using PCE weights. The smallest inconsistency, however, is obtained for the non-core aggregation. Moreover, the aggregation inconsistency for non-core inflation expectations is much smaller than that for core expectations, indicating that non-core category expectations–gasoline, energy, and groceries–play an important role in aggregate inflation expectations, in line with the recent literature. As for mean absolute inconsistency, the max operator yields the largest by a clear margin, and the equal weights aggregation the smallest.

Table 7: Summary Statistics

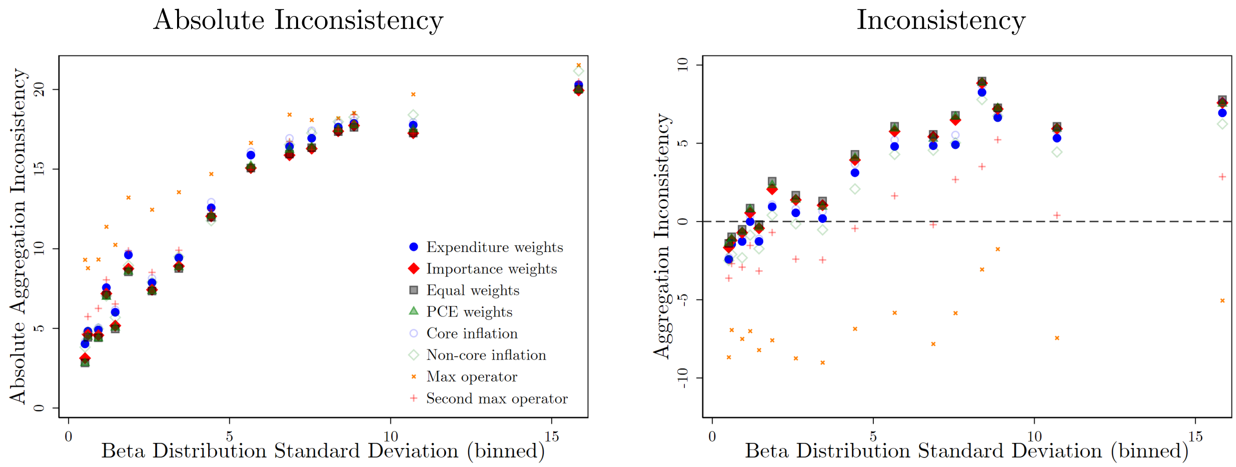
	$\Lambda_i = \beta_0 + \epsilon_i$ $\hat{\beta}_0$	$\lvert\Lambda_i\rvert = \beta_0 + \epsilon_i$ $\hat{\beta}_0$
Plausibly rational aggregation		
Expenditure weights	0.75***	6.24***
	[0.61 0.90]	[6.13 6.35]
Importance weights	1.17***	5.83***
	[1.04 1.31]	[5.73 5.93]
PCE weights	0.51***	6.28***
	[0.37 0.66]	[6.17 6.39]
Behavioral aggregation		
Equal weights	1.35***	5.67***
	[1.21 1.48]	[5.57 5.77]
Core inflation	0.94***	6.41***
	[0.79 1.08]	[6.30 6.53]
Non-core inflation	0.23***	6.20***
	[0.08 0.37]	[6.08 6.31]
Max	-4.75***	9.40***
	[-4.94 -4.57]	[9.24 9.56]
Second max	-0.75***	6.93***
	[-0.91 -0.59]	[6.80 7.05]

Notes: This table presents Huber-robust and survey-weighted estimates for the aggregation mean inconsistency and mean absolute aggregation inconsistency; the numbers in brackets below show the 95% confidence bounds.

5.3.1 Demographics and Aggregation Inconsistency

It is worth considering how demographic and socioeconomic factors relate to the aggregation inconsistency presented in Table 7. Tables 12 and 13 in the Appendix regress aggregation inconsistency and *absolute* aggregation inconsistency, respectively, on an array of demographic characteristics. Women tend to display a larger *absolute* inconsistency than do men, but there is no gender difference in aggregation inconsistency for most aggregation measures (except for the max operator). Younger respondents, relative to older respondents, exhibit a larger *absolute* inconsistency as well as a larger aggregation inconsistency. The highly educated respondents display both a smaller *absolute* inconsistency and a smaller aggregation inconsistency compared to the less educated. In general, the pattern is consistent across most aggregations. This demographic heterogeneity in aggregation inconsistency might offer promising directions for exploring why mean aggregate inflation expectations in major surveys of US consumers, 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 certain demographic segments (such as

Figure 5.3: Aggregation Inconsistency and Uncertainty



Notes: The left panel shows the correlation between the absolute aggregation inconsistency $abs(\Lambda_i^{exp})$ and the individual standard deviation of aggregate inflation expectations obtained via a beta distribution over a probabilistic question; the right panel shows the correlation of aggregation error with the individual beta distribution uncertainty.

the young with less education).

Moreover, the finding that higher education is associated with less *absolute* aggregation inconsistency is consistent with the notion that responses for one of the two inflation expectation measures—aggregate or *aggregated*—become more arbitrary and noisy when respondents experience the inflation questions as more complex or difficult to understand. Indeed, D'Acunto et al. (2019, 2021a) find that cognitive abilities are an important determinant of forecast accuracy. D'Acunto et al. (2021a) show that the responses of lower IQ respondents, for which educational attainment might serve as a proxy, are more likely to be rounded in a survey. Those results are consistent with our interpretation that expectations become more arbitrary and noisy for less-educated respondents, ultimately yielding larger aggregation inconsistencies. Binder (2017) finds similar results for rounding in surveys, and Stanisławska et al. (2021) find similar demographic patterns for the probability of consistent responses to questions eliciting expected changes in inflation numerically and qualitatively.

5.3.2 Uncertainty and Aggregation Inconsistency

One way to probe the implications of question complexity is to consider the relationship between inflation uncertainty and both aggregation inconsistency and *absolute* aggregation inconsistency. Presumably, elevated uncertainty about inflation expectations may indicate heightened perceived complexity. As a proxy for aggregate inflation expectations uncertainty at the individual respondent level, we take the standard deviation of aggregate inflation expectations reported in a density forecast (QDIST, Appendix B). Figure 5.3 shows that the absolute aggregation inconsistency increases in a pronounced fashion with respondents' uncertainty about aggregate inflation. This fact suggests that greater uncertainty about aggregate inflation expectations is associated with noisier and more arbitrary responses.

The right panel of Figure 5.3 shows that aggregation inconsistency yields a similar pattern: aggregation inconsistency increases with uncertainty about aggregate inflation expectations for all measures of *aggregated* inflation expectations, except the max operator, which we may look past as by design it takes the most extreme category inflation expectations. 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; had they not, we would have expected a flat line.

These 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 suggest that lower personal confidence in forecasts, as measured by uncertainty, corresponds to greater inconsistency, as survey responses become more arbitrary—possibly because the inflation concept respondents have in mind is less clear.

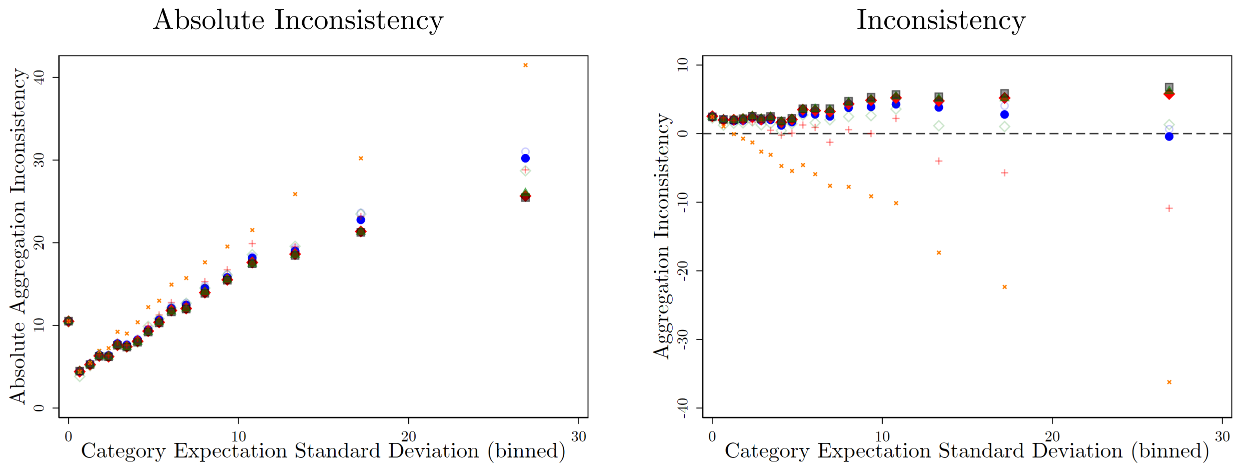
5.3.3 Category Expectation Dispersion and Inconsistency

Another aspect of complexity in inflation expectations pertains to variation between consumption categories. When an individual expresses greater dispersion in category expectations, this may reflect a more complex, differentiated view on the economy, rendering a judgment on future aggregate inflation inherently more difficult. Moreover, the mere mental computation of aggregate expectations also becomes more challenging.

We use the standard deviation across a respondent's category inflation expectations as a proxy for the dispersion of category inflation expectations. Figure 5.4 shows that the *absolute* aggregation inconsistency (left panel)increases strongly with dispersion in category expectations. The picture is less clear, however, for aggregation inconsistency (right panel), but this is partly because higher dispersion implies more extreme category inflation expectations, which would be captured by the max operators. When we abstract from the max operators, we see a slight upward association between aggregation inconsistency and dispersion in category inflation expectations.

Overall, we find that both aggregation inconsistency and *absolute* aggregation inconsistency are positively associated with proxies for the complexity of the aggregate inflation expectations concept. In other words, the more complex the aggregate inflation concept, the greater the divergence between aggregate and *aggregated* inflation expectations.

Figure 5.4: Aggregation Inconsistency and Uncertainty



Notes: The left panel shows the correlation between the absolute aggregation inconsistency $abs(\Lambda_i^{exp})$ and the individual standard deviation of category expectations; the right panel shows the correlation of the aggregation error with the individual standard deviation of category expectations.

6 Economic Implications

Our findings have important implications for the estimation of a central relationship in macroeconomics—the consumption Euler equation. Aggregated inflation expectations contain additional, relevant information about consumption plans relative to conventionally elicited inflation expectations.

To show this result, we assume that consumers follow a standard Euler equation, such as

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

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 for individual i-specific levels of the discount factor β_i, as well as a nominal interest rate $r_{i,t} = -\log(Q_{i,t})$. $\mathbb{E}_t^i$ gives the expectations operator for respondent i. A log-linearized version of equation (2) reads as:

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

where $\pi_t = p_t - p_{t-1}$ denotes the inflation rate. While $\mathbb{E}_t \, c_{i,t+1}$ 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 the expression for the expected change in nominal consumption spending into equation (3) 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 can now estimate this equation in the following form:

$$\mathbb{E}_t^i \, \Delta s_{i,t+1} = \beta_0 + \beta_1 \, \mathbb{E}_t^i \, \pi_{t+1} + D_i + T_t + \epsilon_{i,t} \tag{5}$$

where D_i represents demographic fixed effects[8] 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 faced by households (i.e., specific risk premia). The estimation coefficient β_1 is equal to $1 - \sigma$ in the model in equation (4).

Table 8 shows estimation results, using our individual-level, cross-sectional data, for the full array of inflation expectation measures in the cross-section. Here, we report $1 - \hat{\beta}_1$, 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.

Three results are of note: 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 the 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 other aggregated measure; aggregated measures of inflation expectations are more informative for future spending plans and can thus be said to better represent effective beliefs.

Second, the picture is similar 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, although its performance is matched by the model using self-reported expenditure weights.

Third, non-core inflation expectations, although representing only a small fraction of total spending, seem to provide the best fit in our data with the model in (5). This suggests that spending plans—even in the services category—seem to rely more on expectations for salient non-core product categories than for others.

We turn next to time-series estimations, which confirm the main finding from our cross-sectional estimations—that aggregated measures of inflation expectations are more informative for future spending plans. Because our time-series estimations are based on responses aggregated across

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

Table 8: 1 Year Ahead Spending Plans

	$\hat{\sigma} = 1 - \hat{\beta}_1$	t-stat	R^2	AIC	p-val (LR)
12-months-ahead aggregate spending					
Aggregate	0.968***	5.05	0.06	81615	-
Expenditure	0.910***	6.97	0.07	81527	0.000
Importance	0.801***	10.33	0.08	81090	0.000
PCE	0.837***	10.28	0.08	81104	0.000
Equal	0.788***	10.43	0.08	81318	0.000
Core inflation	0.918***	6.45	0.07	79883	0.000
Non-core inflation	0.885***	8.15	0.08	71907	0.000
Max	0.939***	7.81	0.07	81530	0.000
Second max	0.881***	9.62	0.08	81368	0.000
12-months-ahead nondurable spending					
Aggregate	0.967***	4.16	0.05	37652	-
Expenditure	0.889***	6.12	0.06	37587	0.000
Importance	0.755***	10.24	0.09	37326	0.000
PCE	0.799***	10.23	0.09	37335	0.000
Equal	0.738***	10.21	0.09	37432	0.000
Core inflation	0.907***	5.47	0.06	37055	0.000
Non-core inflation	0.839***	8.91	0.08	33431	0.000
Max	0.920***	6.68	0.06	37578	0.000
Second max	0.859***	8.20	0.08	37482	0.000
12-months-ahead services spending					
Aggregate	0.978***	4.19	0.06	79434	-
Expenditure	0.931***	6.58	0.06	79359	0.000
Importance	0.801***	10.33	0.08	81090	0.000
PCE	0.858***	10.41	0.08	78930	0.000
Equal	0.814***	10.78	0.08	79118	0.000
Core inflation	0.935***	6.32	0.06	77750	0.000
Non-core inflation	0.901***	7.73	0.08	69671	0.000
Max	0.945***	8.24	0.07	79318	0.000
Second max	0.899***	9.48	0.08	79184	0.000

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.

individuals, they have the advantage of averaging out individual-level response noise. At the same time, however, our time series is limited by the length of our survey, which is 42 weeks. Table 15 in the Appendix displays estimations for one-year-ahead aggregate spending plans, nondurable spending plans, and services spending plans. For aggregate spending plans, only two estimations yield a model fit statistically distinct from the aggregate model ($R^2 = 0.21$; AIC = 96.6): the model using expenditure weights ($R^2 = 0.36$; AIC = 88.9) and the second-max operator ($R^2 = 0.47$; AIC = 85.6), both with a positive and significant coefficient for inflation expectations. For nondurable spending plans, only the max ($R^2 = 0.34$; AIC = 17.6) and second-max operator ($R^2 = 0.34$; AIC = 17.9) are statistically different from the aggregate model ($R^2 = 0.19$; AIC = 25.1), but of these, only the max operator has a positive and significant coefficient. For services spending, the pattern is similar; only the max ($R^2 = 0.64$; AIC = 33.2) and second-max operators ($R^2 = 0.55$; AIC = 51.4) are statistically different from the aggregate model ($R^2 = 0.16$; AIC = 71.2), and all three have positive and significant coefficients for inflation expectations.

It is perhaps not surprising that the max operators seem to outperform the other aggregations in our time series. Compared to the cross-section, the time series are likely characterized by different kinds of variation, such as the occurrence of extreme category values over given periods. The max operators by design better capture such variation.

7 Conclusion

We present novel survey evidence on consumer inflation expectations by PCE categories, the aggregations of which we compare to a conventional measure of aggregate inflation expectations.

Four striking facts stand out. The first is that aggregate inflation expectations are higher than inflation expectations for any single category. For the representative agent, this rules out a linear mapping (with non-negative weights) of the category expectations into the aggregate inflation expectations. Moreover, disagreement among respondents over aggregate inflation expectations is higher than that over any category. Second, *aggregated* inflation expectations are less overstated than are the aggregate expectations–the whole is greater than the sum of the parts. *Aggregated* inflation expectations are also less dispersed, and, of the aggregations investigated, non-core inflation expectations seem to align most closely with aggregate expectations in the cross-section. Third, the respondent-specific inconsistency 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, *aggregated* inflation expectations represent better predictors of planned household spending than do aggregate inflation expectations.

The first and second facts are consistent with a psychological interpretation of expectation formation, whereby individuals rely on nonlinear cognitive heuristics to express their explicit aggregate inflation expectations. Following this line of reasoning, the third fact suggests that the heuristics involved in expressing aggregate inflation expectations may differ from the processes underlying *aggregated* inflation expectations.

The fourth fact indicates that *aggregated* category expectations provide the most informative measure of the beliefs on which individuals act—that is, effective beliefs. We explore this point both in the cross-section and in the time-series dimension. Our cross-section estimations have the advantage of a very large number of observations, albeit at the expense of relatively noisy survey measures. Conversely, our time-series estimations rely on response averages, which should cancel noise, but provide a small number of data points. Nevertheless, both sets of models paint a consistent picture: models with *aggregated* inflation expectations yield improved fit over those with aggregate inflation expectations. In the time-series models, this improvement is very pronounced. Moreover, models with plausibly rational aggregations of category expectations—such as self-reported expenditure—consistently outperform those with aggregate inflation expectations, in the cross-section and in some cases also in the time-series dimension.

Effective inflation expectations, it would appear, are *not* best represented by explicit, conventionally reported aggregate inflation expectations. Rather, aggregations of category inflation expectations hold more promise—and the non-core aggregation most of all.

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.

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.

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.

Binder, Carola (2017). "Measuring uncertainty based on rounding: new method and application to inflation expectations". *Journal of Monetary Economics 90*, 1–12. DOI: 10.1016/j.jmoneco.2017.06.001.

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.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). "Diagnostic expectations and credit cycles". *The Journal of Finance 73* (1), 199–227. DOI: 10.1111/jofi.12586.

Bordalo, P., B. Giovanni, K. Coffman, N. Gennaioli, and A. Shleifer (2022). "Imagining the future: memory, simulation and beliefs about COVID". Working paper.

Bruin, Wändi Bruine de, 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.

Bruin, Wändi Bruine de 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.

Carroll, C. D. (2003). "Macroeconomic expectations of households and professional forecasters". *Quarterly Journal of Economics 118* (1), 269–298. DOI: 10.1162/00335530360535207.

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.

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.

——— (2021a). "IQ, expectations, and choice". *Review of Economic Studies*. forthcoming.

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

Dawes, R. M., D. Faust, and P. E. Meehl (1989). "Clinical versus actuarial judgment". *Science 243* (2), 1668–1674. DOI: 10.1126/science.2648573.

Dawes, Robyn M. (1979). "The robust beauty of improper linear models in decission making". *American Psychologist 34* (7). DOI: 10.1037/0003-066X.34.7.571.

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 (2022). "Household inflation expectations and the optimal inflation target measure". Unpublished manuscript.

Dietrich, Alexander, 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*. Forthcoming. DOI: 10.1016/j.jmoneco.2022.02.004.

Einhorn, H. J. and R. M. Hogarth (1975). "Unit weighting schemes for decision making". *Organizational Behavior and Human Performance 13* (2), 171–192. DOI: 10.1016/0030-5073(75)90044-6.

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.

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

Gennaioli, N. and A. Shleifer (2010). "What comes to mind". *Quarterly Journal of Economics 125* (4), 1399–1433. DOI: 10.1162/qjec.2010.125.4.1399.

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.

——— (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.

Kahneman, D. and A. Tversky (1972). "Subjective probability: a judgment of representativeness". *Cognitive Psychology 3*, 430–454. DOI: 10.1016/0010-0285(72)90016-3.

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.

Malmendier, Ulrike and Stefan Nagel (2011). "Depression babies: do macroeconomic experiences affect risk taking?" *The Quarterly Journal of Economics 126* (1), 373–416. DOI: 10.1093/qje/qjq004.

Meehl, Paul E. (1954). *Clinical versus statistical prediction: a theoretical analysis and review of the evidence*. University of Minnesota Press. DOI: 10.1037/11281-000.

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

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, A. and D. Kahneman (1983). "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment". *Psychological Review 90* (4), 293–315. DOI: 10.1037/0033-295X.90.4.293.

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.

Verbrugge, R. and S. Zaman (2021). "Whose inflation expectations best predict inflation?" *Economic Commentary* (2021-19). DOI: 10.26509/frbc-ec-202119.

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.

A Additional Tables

A.1 Demographic Summary Statistics

Table 9: Summary Statistics - Mean Demographics

	Gender		Grocery		Education			Income	
	Female	Male	Yes	No	High	Low	High	Middle	Low
Aggregate expectation	5.63	4.56	5.40	3.54	4.48	5.75	4.41	4.89	5.47
Category expectations									
Motor vehicles	4.48	4.53	4.71	3.61	4.65	4.41	4.61	4.57	4.48
Recreational goods	3.54	3.01	3.37	2.42	3.32	3.17	3.36	3.28	3.11
Other durable goods	3.29	3.18	3.33	2.52	3.24	3.20	3.33	3.33	3.09
Food and beverages	5.36	4.42	5.00	4.22	4.74	4.98	4.81	5.03	4.79
Gasoline	4.66	4.50	4.66	4.05	4.41	4.71	4.35	4.84	4.54
Other nondurable	3.77	3.40	3.66	3.01	3.55	3.62	3.73	3.61	3.42
Housing and util.	5.19	4.52	4.92	4.36	5.04	4.68	4.91	5.27	4.44
Health care	3.16	3.22	3.30	2.46	3.23	3.16	3.47	3.25	2.95
Transportation	4.58	3.96	4.42	3.35	4.15	4.36	4.02	4.39	4.31
Food services	4.39	4.08	4.32	3.66	4.29	4.19	4.41	4.27	4.06
Other services	4.24	3.65	4.08	3.31	3.86	3.99	3.92	4.09	3.86
Aggregated expectations									
Plausibly rational aggregation									
Expenditure weights	4.89	4.19	4.61	3.90	4.47	4.61	4.37	4.72	4.49
Importance weights	4.27	3.72	4.06	3.43	3.90	4.05	3.93	4.12	3.88
PCE weights	4.03	3.62	3.87	3.24	3.72	3.88	3.77	3.92	3.71
Behavioral aggregation									
Equal weights	4.26	3.65	4.02	3.40	3.86	4.00	3.86	4.11	3.84
Core inflation	4.67	4.00	4.42	3.58	4.30	4.36	4.22	4.51	4.26
Non-core inflation	5.71	4.75	5.25	4.68	4.87	5.53	4.86	5.27	5.39
Max	11.58	9.80	10.44	9.97	10.30	11.01	10.10	10.53	11.05
Second max	6.81	5.97	6.72	6.17	6.31	6.43	6.09	6.83	6.44

Notes: This table presents summary statistics on the Huber-robust and survey-weighted mean on expectations across demographics.

Table 10: Summary Statistics - Standard Deviation Demographics

	Gender		Grocery		Education		Income		
	Female	Male	Yes	No	High	Low	High	Middle	Low
Aggregate expectation	9.39	5.79	7.64	5.97	5.66	9.33	5.67	6.22	8.32
Category expectations									
Motor vehicles	7.01	5.56	6.65	5.66	5.68	6.78	6.22	5.72	7.03
Recreational goods	7.18	5.05	6.93	4.94	5.02	7.42	5.75	5.93	8.00
Other durable goods	7.69	4.99	6.81	4.90	5.05	6.90	5.70	4.99	7.23
Food and beverages	7.19	5.88	6.95	5.86	5.84	7.13	6.47	5.93	7.38
Gasoline	7.54	7.11	7.33	7.30	7.22	7.41	7.11	7.32	7.50
Other nondurable	6.95	5.66	6.62	5.04	5.72	6.77	5.53	5.83	6.27
Housing and util.	7.45	6.57	7.05	6.83	6.76	7.20	6.59	6.85	7.45
Health care	8.18	6.05	7.79	6.27	7.04	7.86	7.64	7.04	7.32
Transportation	7.01	5.63	6.70	4.86	5.72	6.88	5.56	5.78	7.18
Food services	7.39	6.69	7.06	6.93	6.91	7.14	6.72	6.88	7.40
Other services	6.73	4.69	6.45	4.68	4.74	6.61	5.42	5.56	6.88
Aggregated expectations									
Plausibly rational aggregation									
Expenditure weights	5.89	4.62	5.27	4.64	4.76	5.78	4.63	5.01	6.02
Importance weights	5.01	3.96	4.47	4.25	4.06	4.80	4.03	4.31	4.88
PCE weights	4.78	3.85	4.26	3.94	3.90	4.58	3.87	4.11	4.66
Behavioral aggregation									
Equal weights	4.96	3.87	4.45	4.06	3.99	4.75	3.92	4.23	4.90
Core inflation	6.13	4.54	5.38	4.64	4.81	5.92	4.66	5.18	6.14
Non-core inflation	6.78	5.56	6.12	5.60	5.59	6.75	5.47	5.72	7.01
Max	8.72	7.16	7.53	7.57	7.29	8.55	7.23	7.52	8.66
Second max	6.48	5.83	6.99	6.80	5.89	6.37	5.81	6.83	6.47

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

Table 11: 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	5.61	6.44	5.46	4.14	10.75	10.63	7.48	3.77
Category expectations								
Motor vehicles	3.97	5.28	4.35	4.93	7.15	6.85	6.43	5.15
Recreational goods	2.06	3.50	3.61	4.23	8.14	7.89	5.92	5.23
Other durable goods	2.25	3.59	3.50	3.89	7.20	7.26	5.94	4.49
Food and beverages	3.58	4.74	5.51	5.79	8.05	7.23	6.52	4.50
Gasoline	3.41	4.60	4.79	6.05	8.12	7.47	7.16	7.04
Other nondurable	2.52	3.81	4.04	4.29	7.23	7.01	6.35	4.23
Housing and util.	3.64	4.62	5.37	5.83	7.48	7.29	6.74	5.50
Health care	1.81	3.49	3.59	4.55	8.10	7.82	6.95	6.60
Transportation	3.41	4.24	4.39	5.16	7.85	6.93	6.58	5.08
Food services	2.63	4.30	4.71	5.81	8.05	7.18	6.77	5.67
Other services	3.23	3.82	4.29	4.48	7.04	6.81	5.42	4.16
Aggregated expectations								
Plausibly rational aggregation								
Expenditure weights	3.43	4.55	5.02	5.49	5.35	5.47	5.27	4.40
Importance weights	2.67	4.15	4.35	5.32	3.89	4.71	4.49	4.20
PCE weights	2.74	3.99	4.29	5.18	4.06	4.80	4.45	3.98
Behavioral aggregation								
Equal weights	2.58	3.94	4.08	5.11	3.76	4.48	4.34	4.04
Core inflation	3.29	4.48	4.78	5.16	5.60	5.71	5.33	4.44
Non-core inflation	3.93	5.03	5.38	6.12	6.30	6.78	5.59	5.06
Max	10.45	10.87	9.82	11.32	8.36	8.46	7.43	7.52
Second max	5.99	6.41	6.22	7.29	7.37	6.22	6.82	5.64

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

A.2 Demographic Effects: Aggregation Inconsistency

Table 12: Demographics and Aggregation Inconsistency

	(1) Expenditure	(2) Importance	(3) PCE	(4) Equal	(5) Core	(6) Non-core	(7) 1^{st}max	(8) 2^{nd} max
Female	-0.185	0.0142	-0.0972	0.0472	-0.108	0.125	-0.802***	-0.341*
	(-1.17)	(0.10)	(-0.68)	(0.33)	(-0.66)	(0.76)	(-3.98)	(-2.00)
Grocery Shopper	0.585**	0.615**	0.591**	0.517**	0.542*	0.408	1.166***	0.887***
	(2.72)	(3.02)	(2.96)	(2.61)	(2.47)	(1.81)	(4.12)	(3.76)
35 to 44 years	0.289	0.113	0.173	0.140	0.226	0.136	0.556*	0.452*
	(1.33)	(0.56)	(0.88)	(0.72)	(1.00)	(0.58)	(2.08)	(1.97)
45 to 54 years	-1.102***	-1.135***	-1.072***	-1.053***	-1.071***	-1.205***	0.171	-0.410
	(-4.77)	(-5.23)	(-5.07)	(-5.01)	(-4.44)	(-4.99)	(0.58)	(-1.66)
above 55 years	-2.112***	-2.372***	-2.162***	-2.272***	-2.061***	-2.458***	-2.212***	-2.024***
	(-12.05)	(-14.47)	(-13.71)	(-14.39)	(-11.36)	(-12.92)	(-9.79)	(-10.63)
High Educated	-0.705***	-0.760***	-0.828***	-0.780***	-0.690***	-0.597***	-0.789***	-0.885***
	(-4.28)	(-4.94)	(-5.60)	(-5.28)	(-4.02)	(-3.55)	(-3.68)	(-4.91)
Middle Income	-0.346	-0.0456	-0.0706	-0.0114	-0.248	-0.475*	0.0974	0.0439
	(-1.83)	(-0.25)	(-0.41)	(-0.07)	(-1.27)	(-2.49)	(0.40)	(0.21)
High Income	-0.203	-0.226	-0.243	-0.184	-0.164	-0.232	0.253	0.126
	(-0.98)	(-1.17)	(-1.30)	(-0.99)	(-0.76)	(-1.10)	(0.94)	(0.55)
Constant	1.803***	2.122***	2.175***	2.293***	1.931***	1.569***	-4.395***	-0.250
	(6.89)	(8.59)	(8.95)	(9.50)	(7.21)	(5.54)	(-13.11)	(-0.87)
N	16248	16109	16086	16114	15857	14516	16796	16404
r2	0.0177	0.0239	0.0225	0.0235	0.0160	0.0227	0.0137	0.0153

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 aggregation error on several demographic characteristics. The headers for each column represent the aggregation mechanism. For details, see Table 4.

Table 13: Demographics and Absolute Aggregation Inconsistency

	(1) Expenditure	(2) Importance	(3) PCE	(4) Equal	(5) Core	(6) Non-core	(7) 1^{st}max	(8) 2^{nd} max
Female	0.761***	0.824***	0.878***	0.863***	0.770***	0.808***	0.950***	0.579***
	(6.17)	(7.34)	(7.95)	(7.83)	(5.95)	(6.28)	(5.22)	(4.26)
Grocery Shopper	0.295	0.223	0.102	0.130	0.435*	-0.0269	-0.435	0.0816
	(1.78)	(1.46)	(0.68)	(0.86)	(2.51)	(-0.16)	(-1.72)	(0.44)
35 to 44 years	-0.190	-0.0341	0.0864	0.0242	-0.207	-0.0155	-0.269	-0.562**
	(-1.12)	(-0.22)	(0.56)	(0.16)	(-1.15)	(-0.08)	(-1.10)	(-2.98)
45 to 54 years	-1.488***	-1.346***	-1.245***	-1.294***	-1.585***	-1.773***	-1.675***	-1.774***
	(-8.37)	(-8.35)	(-7.74)	(-8.09)	(-8.43)	(-9.48)	(-6.43)	(-9.10)
above 55 years	-2.495***	-2.307***	-2.317***	-2.312***	-2.643***	-2.239***	-1.654***	-2.530***
	(-18.35)	(-18.63)	(-19.15)	(-19.07)	(-18.57)	(-15.11)	(-8.23)	(-16.87)
High Educated	-1.060***	-0.959***	-1.006***	-0.955***	-1.154***	-0.812***	-0.738***	-1.050***
	(-8.14)	(-8.12)	(-8.71)	(-8.30)	(-8.40)	(-6.03)	(-3.80)	(-7.26)
Middle Income	-0.248	-0.272*	-0.335*	-0.346**	-0.215	-0.516***	-0.416*	-0.350*
	(-1.68)	(-2.03)	(-2.52)	(-2.62)	(-1.39)	(-3.46)	(-1.96)	(-2.18)
High Income	-0.302	-0.399**	-0.453**	-0.512***	-0.322	-0.377*	-0.255	-0.162
	(-1.84)	(-2.65)	(-3.10)	(-3.49)	(-1.85)	(-2.21)	(-1.01)	(-0.87)
Constant	7.367***	6.874***	6.855***	6.858***	7.540***	7.640***	10.96***	8.429***
	(36.30)	(37.06)	(37.61)	(37.12)	(35.59)	(35.40)	(36.31)	(37.71)
N	16188	16047	16078	16088	15854	14503	17118	16580
r2	0.0494	0.0539	0.0577	0.0571	0.0528	0.0432	0.0130	0.0370

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 error on several demographic characteristics. The headers for each column represent the aggregation mechanism. For details, see Table 4.

A.3 Category Expectations

Table 14: Categories with 1^{st} and 2^{nd} highest expectation

Category	1st max	2nd max
Motor vehicles	33.4%	36.1%
Recreational goods	28.8%	35.1%
Other durable goods	28.1%	35.1%
Food and beverages	35.6%	39.8%
Gasoline	39.2%	36.4%
Other nondurable goods	28.0%	36.5%
Housing and utilities	37.7%	38.1%
Health care	30.5%	35.8%
Transportation services	31.7%	38.1%
Food services	32.9%	39.0%
Other services	28.8%	36.8%

Notes: The table shows the frequency for each category of being a survey participant's largest or second-largest expectation in the cross-section. Note that numbers need not add up to 1 as a respondent might have the same expectation for multiple categories.

A.4 Implied Aggregation Weights

The data set collected allows us to investigate the data-implied weights people use internally to sum up categorical inflation expectations toward an aggregate expectation. This section will outline the approach and show results.

We can estimate the data implied aggregation weights $\hat{\omega}_k$ by the following equation:

$$\mathbb{E}_t^i \, \pi_{t+1} = \sum_{k=1}^{11} w_i \hat{\omega}_k \, \mathbb{E}_t^i(\pi_{k,t+1}) + \epsilon_i \tag{6}$$

subject to a set of 12 constraints:

$$\omega_k \geq 0 \quad \forall k \in (1, 11) \tag{7}$$

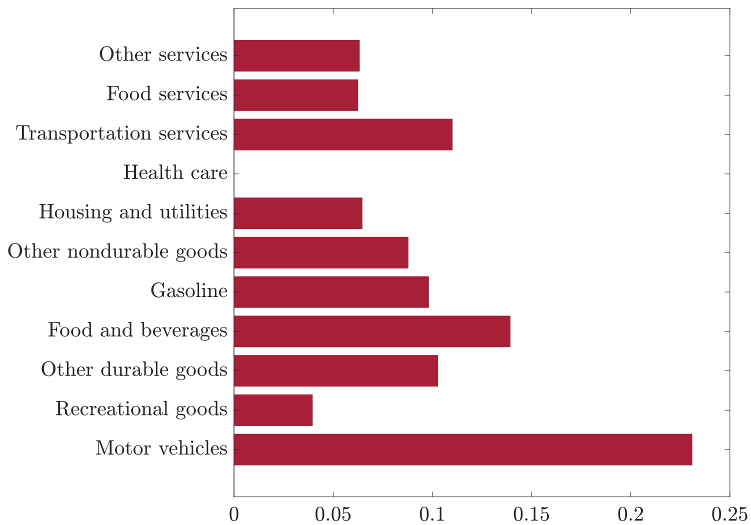
$$\sum_{k=1}^{11} \omega_k = 1 \tag{8}$$

The imposed constraints ensure that all weights are non-negative (7) and that the sum of weights is equal to one (8). This produces results that are directly comparable to PCE weights.

Before estimation, respondents receive survey weights w_i based on demographic characteristics, such as age, gender, census region, race, income, and education (see Table 9).

We then use a weighted least squares approach minimizing the weighted mean squared error subject to (7) and (8).[9] Estimation results $\hat{\omega}_k$ may be interpreted as the mean of weights used by respondents to sum up categorical expectations. Figure A.1 displays the data-implied weights.

Figure A.1: Data Implied Weights



Notes: The figure shows implied weights for each category.

[9] In matrix notation, the minimization problem can be expressed as:

$$\min_{\beta} WMSE = (\boldsymbol{E(\pi)} - \boldsymbol{E(\pi_{Cat.})\omega})'\boldsymbol{W}(\boldsymbol{E(\pi)} - \boldsymbol{E(\pi_{Cat.})\omega}) \quad s.t. \quad \begin{Bmatrix} \boldsymbol{\omega} \geq \boldsymbol{A} \\ \boldsymbol{B\omega} = 1 \end{Bmatrix} \tag{9}$$

$\boldsymbol{E(\pi)}$ is a (j x 1) vector of aggregate inflation expectations, $\boldsymbol{E(\pi_{Cat.})}$ the (j x 11) matrix of category expectations. $\boldsymbol{\omega}$ is the (11 x 1) vector of aggregation weights ω_k. $\hat{\boldsymbol{\beta}}$ gives the (11 x 1) vector that minimizes expression (9). $\boldsymbol{W}$ is the (j x n) diagonal matrix of survey weights. $\boldsymbol{A}$ is a (11 x 1) vector of zeros, $\boldsymbol{B}$ is a (1 x 11) row vector of ones. j gives the number of respondents.

A.5 Time Series - Spending Plans

Table 15: Time Series: 1 Year Ahead Spending Plans

	$\hat{\sigma} = 1 - \hat{\beta}_1$	t-stat	R^2	AIC	p-val (LR)
12-months-ahead aggregate spending					
Aggregate	0.731**	3.35	0.208	96.62	-
Expenditure	0.366***	4.83	0.358	88.94	0.022
Importance	0.342***	3.63	0.356	92.50	0.128
PCE	0.431***	4.17	0.27	98.35	1.000
Equal	0.351***	3.74	0.318	98.29	1.000
Core inflation	0.468***	3.50	0.26	96.50	0.942
Non-core inflation	0.618**	2.84	0.20	101.42	1.000
Max	0.759***	4.44	0.420	91.23	0.067
Second max	0.485***	4.61	0.465	85.56	0.004
12-months-ahead nondurable spending					
Aggregate	1.027	-0.18	0.19	25.14	-
Expenditure	1.050	-0.21	0.19	25.02	0.940
Importance	0.745	1.02	0.23	24.18	0.620
PCE	0.508*	1.61	0.31	23.14	0.367
Equal	0.751	1.06	0.23	24.13	0.604
Core inflation	0.968	0.13	0.19	25.11	0.987
Non-core inflation	1.026	-0.11	0.19	25.12	0.989
Max	0.813**	2.42	0.34	17.58	0.023
Second max	0.594	1.88	0.34	17.92	0.027
12-months-ahead services spending					
Aggregate	0.750**	2.81	0.162	71.19	-
Expenditure	0.576**	3.01	0.174	77.60	1.000
Importance	0.521***	4.83	0.277	74.54	1.000
PCE	0.500**	3.49	0.254	77.20	1.000
Equal	0.566***	3.71	0.224	77.13	1.000
Core inflation	0.707*	1.72	0.08	82.26	1.000
Non-core inflation	0.569***	3.72	0.26	76.17	1.000
Max	0.815***	11.48	0.642	33.22	0.000
Second max	0.608***	11.04	0.553	51.37	0.000

Notes: Estimated Euler equations, based on weekly time series data, 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 Huber-robust weights to ensure that sample is independent of outliers, respectively.

B Survey Appendix

This section lists relevant survey questions used within the paper.

B.1 Survey Overview

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 September 9, 2021. Participants were asked for their expectations and behavior regarding COVID-19. While the survey also contains other blocks with various questions, these are not reported here, since they are asked after the questions on COVID-19 and thus do not affect the answers.

B.2 Sample

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., "$sd - \$rt2$").

B.3 Aggregate Expectations

To learn about respondents' expectations of future 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".

Q1: 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: 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 start 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 Positive
O Negative

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.*

B.4 Category Expectations and Weights

To elicit participants' category-specific inflation expectations and expenditure weights, we ask the following questions:

Q2: Importance weights

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

Q3: Expenditure weights

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: Category 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	____

B.5 Expected Spending

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

Q4: 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	____

Q5: 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	____

Q6: 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	______

B.6 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, the highest educational level, personal income, and the personal savings rate.