The Dissertation Committee for Gwangmin Kim
certifies that this is the approved version of the following dissertation:

Essays in Empirical Macroeconomics

Committee:

Oliver Coibion, Supervisor

Stefano Eusepi

Saroj Bhattarai

Andy Glover

Essays in Empirical Macroeconomics

by

Gwangmin Kim

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2022

Acknowledgments

I thank Olivier Coibion, Saroj Bhattarai, and Stefano Eusepi for their invaluable long-standing support. Without their advice and guidance, I am sure that it would have been impossible to complete my dissertation. In addition, I am grateful to my co-author, Carola Conces Binder, for her excellent work and collaboration with me for the main chapter of my dissertation. I am also benefited from conversations with Andy Glover, Viacheslav Sheremirov, and Cooper Howes as well as many seminar participants from the University of Texas, the Federal Reserve Bank of Kansas City, and the Federal Reserve Bank of Boston. I also thank Jolie Flink for her excellent administrative help. Finally, of course, I express my deepest gratitude to my family and friends for the constant support and encouragement.

Essays in Empirical Macroeconomics

Publication No. ____________________

Gwangmin Kim, Ph.D.
The University of Texas at Austin, 2022

Supervisor: Oliver Coibion

This dissertation is going to empirically study household inflation expectations and inflation. Inflation expectations and inflation play a central role in economic dynamics. For example, when households expect future price increases, households will try to purchase goods now than later at higher prices which is eventually going to push prices even higher.

Chapter 1 contributes to the literature on inflation expectations by showing a channel that can significantly bias the central banks' aggregate inflation expectations measures. This chapter is joint work with Carola Binder. We show that when inflation expectations surveys rely on repeat survey participants, survey participation itself may affect future responses. Because the central bank's survey asks about future prices and inflation, it prompts information acquisition between survey waves for survey respondents. These "Learning-through-Survey" effects are particularly large for household inflation expectations. For example, after participating twelve consecutive times in

the SCE, respondents end up with a 2.6 percentage point lower inflation forecast and 34% lower inflation uncertainty on average than in the first interview, with most of the decline happening in the first two months of participation. Consequently, repeat participants may be more informed, and not be representative of the broader population of the economy.

Chapter 2 estimates three components of household inflation expectations of the SCE using a dynamic factor model: Common, Learning, and Long-run factor. Using the estimated common inflation expectation factor shared by all survey participants, I recover the household inflation expectations less the learning effect of the SCE without discarding repeat survey participants' data which could have been wasteful otherwise. It successfully corrects for the bias due to the learning effects of repeat survey participants and is significantly less noisy than the raw data. In addition, the estimated learning factor and long-run factor of inflation expectations suggest that inflation expectations of households are largely influenced by news coverage on inflation and oil prices.

Finally, Chapter 3 studies the product life cycle effects on prices and inflation inequality in the U.S. The annual inflation rate of lower-income households has been higher than that of higher-income households in general, a finding termed in extensive literature as "inflation inequality." Using barcode-level retail sales data and household spending data in the U.S, I show that the product life cycle channel can account for a significant portion of this inflation inequality among households. The prices of new products are initially high but steadily decrease after then as it goes out of fashion. Because rich households tend to be early adopters preferring new goods to old goods, those rich early adopters experience a sharp price decrease or lower inflation than poor late-adopters who buy goods when the price decreasing phase has stopped or got less steep.

Table of Contents

List of Tables

List of Figures

Chapter 1

Learning-through-Survey in Inflation Expectations

1.1 Introduction[1]

Inflation expectations are believed to play a central role in economic dynamics. Federal Reserve Chair Jerome Powell testified to Congress in February 2019 that "Inflation expectations are the most important driver in actual inflation" (Powell, 2019). In addition, survey-based inflation expectation measures are increasingly used in economic research in various ways: for estimating the intertemporal elasticity of substitution (Crump et al., 2015), studying inflation expectations of firms (Coibion et al., 2018), and estimating expectations-augmented Phillips curve (Coibion et al., 2019).

Therefore, accurately measuring inflation expectations is crucial for monetary policymaking and economic research. For this reason, the Federal Reserve Bank of New York (FRBNY) began conducting the Survey of Con-

[1]Some parts of Chapter 1 are published in the following citation:
Kim, GwangMin, and Carola Binder. "Learning-through-survey in Inflation Expectations." *American Economic Journal: Macroeconomics*, Conditionally Accepted (2022).
Below briefly describes the nature of authors contribution:
Gwangmin Kim: Conceptualization, Writing - initial draft & editing, Data acquisition, Methodology, Formal statistical analysis, and Reviews. Carola Binder: Conceptualization, Writing - review & editing, Methodology, and Formal statistical analysis.

sumer Expectations (SCE) monthly in 2013. Other central banks, like the European Central Bank (ECB) and the Bank of Canada, are also introducing new household surveys. For example, the ECB launched the Consumer Expectations Survey in January 2020.[2] Globally, dozens of countries run household inflation expectation surveys on a regular basis.[3]

FRBNY SCE respondents can participate in the survey for up to twelve months in a row. A long panel dimension is usually thought to be a desirable feature for a survey, since measuring the same person over time allows researchers to control for unobservable individual-specific characteristics. However, reliance on repeat participants—the SCE includes about 150 new participants out of 1300 in each wave—could pose problems if the act of participating in the survey affects the subsequent responses of these participants. These so-called *learning-through-survey* or *panel conditioning* effects are small in some surveys.[4]

However, we show that this is decidedly *not* the case in surveys of household inflation expectations. After being asked about their inflation ex-

[2]The ECB Consumer Expectations Survey is in pilot phase in six countries: Belgium, France, Germany, Italy, the Netherlands, and Spain. More countries may be added if the pilot proves successful. `https://www.ecb.europa.eu/stats/ecb_surveys/consumer_exp_survey/html/index.en.html` provides further details about this new survey.

[3]See Arioli et al. (2017) and Appendix Table 1 of Coibion et al. (2019) for a list of countries running inflation expectation surveys targeting households. Norway, which is not on their list, also has run an inflation expectation survey since 2002.

[4]For example, Halpern-Manners et al. (2017) find that only about 12% of the selected core items of the General Social Survey display panel conditioning effects at a 5% significance level. They report that the responses of different survey cohorts do not appear to differ in predictable or meaningful ways in most cases.

pectations, individuals significantly (and predictably) revise their expectations in subsequent surveys. For example, after participating twelve consecutive times in the SCE, respondents end up with a 2.6 percentage point lower inflation forecast and 34% lower inflation uncertainty on average than in the first interview, with most of the decline happening in the first two months of participation. Results are similar for longer-run inflation expectations. These effects are so large that repeat participants can no longer be considered representative of the general population.

The extensive panel component of the SCE makes it an ideal setting for studying panel conditioning effects. We note that Armantier et al. (2017) briefly examine panel conditioning effects in the SCE by comparing the median absolute change in the density mean of respondents' inflation expectations across different tenure groups.[5] They find that after the first month of participation, the density mean of a respondent's inflation forecast remains relatively stable.

We provide evidence, however, that the learning-through-survey effects are larger and more economically meaningful than previously recognized. Our approach is to use panel regressions with time and respondent fixed effects and tenure dummy variables to detect conditioning effects that may occur over multiple survey waves, without imposing parametric assumptions on how effects

[5]Throughout this paper, "tenure" refers to the total number of past survey experience of respondents, including the current survey wave. For example, a SCE respondent surveyed each month starting in January will have a tenure of 3 in March.

depend on tenure. This is a novel methodological contribution to a relatively large literature on panel conditioning.[6] The coefficients on the tenure dummies provide non-parametric estimates of how inflation expectations change with survey tenure. Since consumers generally overestimate future inflation,[7] they lower their forecasts and make smaller forecast errors as they remain in the survey and acquire more information.[8]

Furthermore, we characterize which individuals are most sensitive to having their beliefs change through participation in the survey. Respondents who report higher uncertainty about inflation at the time of their first survey tend to make larger revisions to their inflation expectations in subsequent surveys. In addition, more educated and higher-income individuals and retirees, who are generally more informed about inflation prior to the survey, display significantly smaller learning effects throughout the survey waves.

These heterogeneity results are consistent with models that emphasize the endogenous nature of information rigidities. Under the rational inattention

[6] Previous studies conduct t-tests for the difference in mean responses between two cohorts of respondents who first entered the survey sample at two different dates. For example, Halpern-Manners et al. (2017) compare responses on the 2008 General Social Survey for respondents who took the survey in 2006 and 2008 versus respondents who took the survey in 2008 and 2010. We pool information from *all* dates of the SCE rather than from a single survey date, and observe how responses change not only from a respondent's first to second round of participation, but also from her second to third round of participation and so on."

[7] Although Consumer Price Index inflation has been recently low and stable—at around 1.5% from 2013 to 2018—the inflation expectation of consumers was consistently above 2.5% during the same period.

[8] One difference between our analysis and that of Armantier et al. (2017) is that we identify the average change, not the average absolute change, in inflation forecasts. This is appropriate in this context because forecasts have positive bias, so errors have non-zero mean.

model, economic agents have a limited cognitive ability to process information, and must choose how to allocate attention.[9] Though the inflation rate is an important aggregate variable, it may be optimal for households to pay greater attention to tracking other variables, like their own income, that are more relevant to their consumption decisions (Carroll et al., 2020). Indeed, Karahan et al. (2017) find that the income expectations of consumers tend to be accurate. Similarly, when firm-specific conditions are more important than aggregate conditions, firms pay more attention to idiosyncratic variables and devote few resources to collecting and processing information about inflation (Mackowiak and Wiederholt, 2009). Consumers' relatively greater attentiveness to their own income than to inflation is consistent with our result that the learning effect is smaller for personal earnings and household income growth expectations than for inflation expectations.

As we demonstrate, our results should be kept in mind by users and developers of new household surveys, as they affect empirical estimates and interpretations using the survey data in some contexts. We illustrate this point using two application cases: the oil price collapse in 2014 and the estimation of elasticity of intertemporal substitution by Crump et al. (2015). Household inflation expectations are known to be sensitive to gas prices (Coibion and Gorodnichenko, 2015a). However, this stylized fact tends to be significantly weaker for repeat participants of the SCE. During a period of sharp decline in

[9]Related models allow for different types of information rigidities such as infrequent updating (Mankiw and Reis, 2002; Reis, 2006), signal-extraction problems (Sims, 2003), or model complexity restrictions (Gabaix, 2014).

oil and gas prices, we find that inflation expectations of new participants are more influenced by gas prices when compared to repeat survey participants. Also, we show that estimates of the elasticity of intertemporal substitution following the methodology of Crump et al. (2015) are lower for new participants. We suggest that users of survey microdata should check whether their estimates are robust to using subsamples of shorter-tenured and longer-tenured respondents when they use panel data.

Finally, we provide evidence of learning-through-survey effects in other contexts using the Michigan Survey of Consumers (MSC) and a new firm survey to show that learning-through-survey effects are not confined to a specific period or only to household surveys. Learning-through-survey effects tend to be smaller when there is a longer period between baseline and follow-up surveys, as in the MSC. This difference in size of learning effects is consistent with recent evidence from randomized information treatments that finds providing information about inflation to households has large contemporaneous effects on their expectations but that these effects fade very rapidly (Coibion et al., 2020).

The paper is structured as follows. Section 1.2 provides information about the dataset and presents our estimates the learning-through-survey effects in the FRBNY SCE. Section 1.3 provides implications of the learning-through-survey effect for interpretation of survey data using two application cases: the oil price collapse in 2014 and estimation of elasticity of intertemporal substitution. Section 1.4 documents the effect in two other survey datasets.

Section 1.5 discusses an alternative explanation for the tenure effects and implications for future research.

1.2 Learning Effects in the Survey of Consumer Expectations

1.2.1 Data

The Federal Reserve Bank of New York's (FRBNY) Survey of Consumer Expectations (SCE) is an online survey that began 2013. The SCE is monthly and nationally-representative with a rotating panel structure, tracking each respondent up to 12 times consecutively. Each month, the SCE has a sample size of approximately 1,300, and the number of new participants is about 150.

In addition to inflation point forecasts, the FRBNY elicits the respondent's histogram or density forecasts for inflation by asking the respondent to assign probabilities that future inflation will fall into various bins, summing to 100%. Hence, for inflation uncertainty, we use the interquartile range (IQR) estimated from each individual's probabilistic forecast.[10] The exact phrasing of survey questions is available in the appendix A.3.1.

[10]The FRBNY provides estimates of the mean, median, and IQR of each density forecast. These estimates are obtained by fitting parametric (beta) distributions to the density forecasts. See FRBNY SCE documentation for details.

1.2.2 Identification of Tenure Effects

We begin by documenting the presence of tenure effects in the FRBNY SCE data, using linear panel fixed effects regressions of the form:

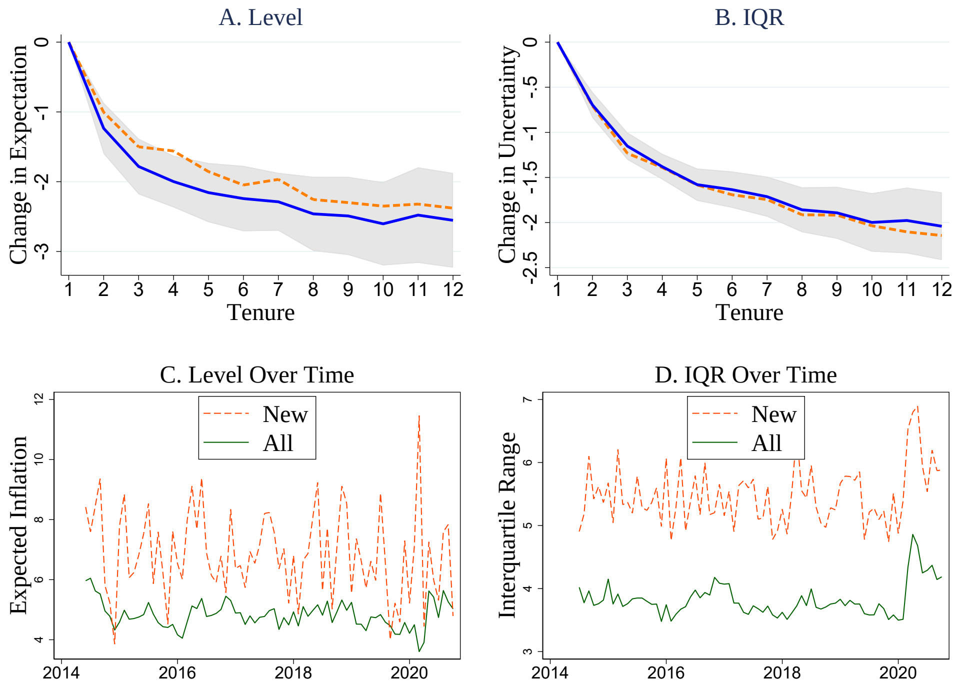
$$y_{its} = \sum_{s=2}^{12} \beta_s \tau_s + \alpha_i + \gamma_t + \varepsilon_{it}, \tag{1.1}$$

where the dependent variable y_{its} is the inflation expectation or inflation uncertainty of respondent i with survey experience (or tenure) s at time t, τ_s is an indicator variable for tenure s, α_i and γ_t are individual- and time- fixed effects to control for unobserved heterogeneity, and ε_{it} is an error term. The regression coefficients on the tenure dummies, $\{\beta_s\}_{s=2}^{12}$, measure the average learning-through-survey effects on the dependent variable. To make the regression coefficients more robust to outliers, we remove the top and bottom 5% of each dependent variable for each tenure group and period.[11]

One identification issue is that sample selection may occur due to panel attrition. For example, more educated and higher income respondents tend to stay in a survey for more waves, and attrition may also depend on unobservable characteristics. To prevent confounding panel conditioning effects with attrition effects, we restrict our sample to consist of "non-attriters," or respondents who eventually participate in the survey for the maximum number of times, following Halpern-Manners et al. (2017).

[11] We obtain qualitatively and quantitatively similar results for different thresholds. Appendix Figure A1 reproduces the results in Figure 1.1 for lower and higher thresholds.

Figure 1.1: Average Learning-Through-Survey Effects on Inflation Expectations in the SCE



Note: Panels A and B show the change in responses of survey participants compared to their initial responses, in percentage points, estimated from regression (1). For Panel A, the dependent variable is the inflation point forecast, and for Panel B, the dependent variable is the interquartile range of the density forecast. The solid blue (dashed orange) lines correspond to one-year (three-year) ahead inflation forecasts. The gray area shows a 95% confidence interval for the solid blue line with Driscoll-Kraay standard errors of lag one. Survey tenure is shown on the x-axis. We restrict samples to respondents who eventually participate in the survey for twelve waves (non-attriters) and winsorize the top and bottom 5% of each dependent variable for each tenure group and period. A full regression table is in Appendix Table A.1. Panels C and D show the mean inflation point forecast and mean inflation density interquartile range over time for new respondents and all respondents, in this case without the non-attrition restriction. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to October 2020 with monthly frequency.

Panel A of Figure 1.1 shows that the estimated average learning-through-survey effect is large and statistically significant for both one-year-ahead and three-year-ahead inflation expectations. Respondents revise their one-year-ahead inflation expectations downward by 1.2 percentage points immediately after the first interview. Respondents with tenure 12 have expectations that are 2.6 percentage points lower than those of new respondents. Three-year-ahead inflation expectations display similar, if slightly smaller, tenure effects. Given this similarity, we primarily focus on one-year-ahead inflation expectations for the remainder of the paper. A full regression table is in Appendix Table A.1.

While the results in Panel A correspond to respondents' point forecasts of inflation, the FRBNY frequently reports on the density mean forecasts. Panel conditioning effects for the density mean forecasts are reported in Columns 3 and 4 of Table A.1. Though statistically significant, they are smaller in magnitude than the effects for the point forecasts: repeat respondents have density means that are around half a percentage point lower than those of new respondents. These smaller tenure effects may be related to the guidance provided by the bin intervals when density forecasts are solicited.

Recall that after the respondent provides a point forecast, her density forecast is solicited, with upper and lower bins corresponding to inflation above 12% and deflation below -12%. The bins near zero are narrower than those above 4% or below -4%. From these bins, the respondent may infer that most of the probability should be placed in $[-4\%, 4\%]$, or at least in $[-12\%, 12\%]$. (Almost a third of first-time respondents provide point forecasts outside of $[-12\%, 12\%]$ and over half provide point forecasts outside of $[-4\%, 4\%]$.) Any learning that occurs from the bin endpoints constitutes "learning-through-survey" in the very first round of the survey, resulting in smaller observed tenure effects in subsequent rounds. We can see some evidence that this occurs, because density forecast means are notably lower than point forecasts. Moreover, new respondents have density forecast means that are significantly further from their point forecasts compared to respondents of higher tenure.

Even though the tenure effects are smaller in magnitude for density mean forecasts than for point forecasts, sizeable tenure effects appear for other features of the density forecasts, such as the interquartile range (IQR). The IQR decreases by about 0.7 percentage points after the first round of the survey, as shown in Panel B of Figure 1.1.

As consumers' uncertainty declines with tenure, their forecast errors also decline, as Appendix Table A.2 shows. Since consumer inflation expectations are typically biased upward, the downward revisions in expectations improve forecast accuracy. The mean absolute forecast error for respondents of tenure 2 is 2.0 percentage points lower, and for respondents of tenure 12 is 4.3 percentage points lower, than that of new participants. The same table shows that higher-tenure respondents make much less frequent forecast revisions.[12] Dräger and Lamla (2017) similarly document that the probability of updating inflation expectations increases when individuals had higher forecast

[12]Note that the frequency of forecast revisions is often used as a proxy for the information rigidity parameter in sticky information models (Coibion and Gorodnichenko, 2015a; Binder, 2017a)

errors in the past. That is, repeat participants achieve lower forecast errors and are less likely to replace their current forecasts, saving cognitive effort in processing information.

Panels C and D of Figure 1.1 show time series plots of the mean inflation point forecasts and density IQR for new respondents compared to all respondents. Here we include both attriters and non-attriters, since policymakers typically monitor the aggregate time series without imposing a non-attrition restriction. The time series corresponding to new respondents are higher and more volatile. These graphs also reveal that tenure effects vary over time. Notably, in March 2020, at the start of the Covid-19 pandemic, new respondents had a mean inflation forecast of 11.5%, compared to 3.6% for the average respondent of any tenure and 2.4% for respondents with tenure greater than one. The relatively less-informed new respondents may have assumed the pandemic would be inflationary (see Binder (2020)).

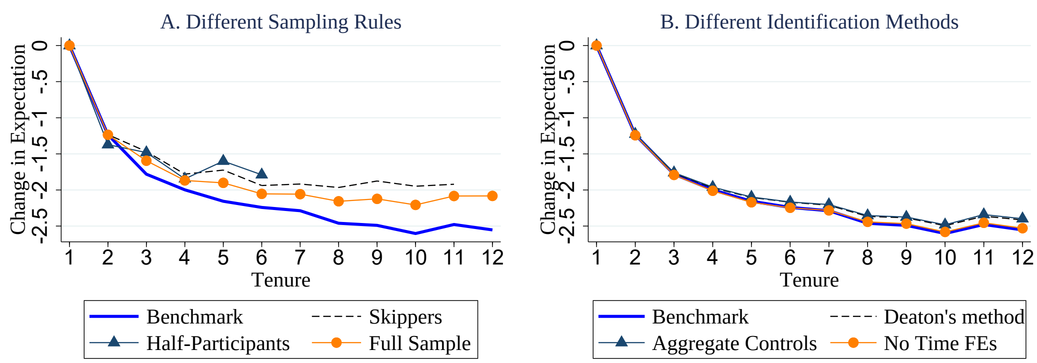
Our time sample also includes a disinflationary episode in 2015. We estimate the tenure effects for 2015 only, and find that the effects are larger in magnitude during this episode (see Column 7 of Table A.1). This is consistent with the learning-through-survey hypothesis, if repeat respondents are more aware of declining inflation than are new respondents. We also construct a time series of the number of articles containing the word "inflation" in the New York Times each month. This series is positively correlated with the inflation expectations of repeat respondents (with correlation coefficient of 0.55), but much less correlated with the inflation expectations of new respondents (with correlation coefficient of 0.38). This could indicate that repeat respondents, primed by their earlier survey participation, are more likely to notice news coverage of inflation and incorporate it into their inflation expectations.

1.2.3 Robustness to Estimation Strategy

A potential concern to our non-attrition restriction is that the tenure effect we identify may only exist for "ready-to-learn" survey respondents—that is, for respondents who are committed to the survey and thereby have more willingness to learn about the economy. However, quantitatively and qualitatively similar results are obtained for different sampling rules. Focusing on one-year-ahead inflation expectations, Panel A of Figure 1.2 reproduces our baseline results from Panel A of Figure 1.1 under various sampling rules. Even if we only include respondents who skip a survey at least one time ("skippers"), who participate in the SCE less than six times in total ("half-participants"), or include the full sample ("full sample"), the estimated tenure effects are similar to those from our "non-attriters" sample. Therefore, throughout this paper, we keep "non-attriters" as our baseline sample.

Another potential identification issue is known as the Age-Period-Cohort (APC) problem, which in its original formulation refers to the problem of separating the independent effects of age, time period, and cohort due to exact linear dependence (Hobcraft et al., 1985; Deaton and Paxson, 1994). In our context, survey experience dummy variables, monthly time-fixed effects, and

Figure 1.2: Baseline Results under Different Sampling Rules and Identification Methods



Note: Panel A reproduces the results from Panel A of Figure 1.1 under different sampling rules. "Benchmark" corresponds to the baseline results using non-attriters only. "Skippers" are respondents who skip a survey at least once. "Half-participants" participate in the SCE no more than six times. "Full sample" corresponds to the case when we do not make restrictions based on total survey participation.
Panel B reproduces the results of one-year-ahead inflation forecasts in Panel A of Figure 1.1 using different identification methods. "Benchmark" corresponds to the baseline results: linear panel fixed effects regression with quarterly time fixed effects and individual fixed effects. "Deaton's method" uses normalization of monthly time fixed effects following Deaton and Paxson (1994). "Aggregate Controls" replaces time-fixed effects with macroeconomic aggregate variables: monthly CPI inflation rates, the aggregate median of MSC one-year-ahead inflation forecasts, unemployment rate, monthly growth rate of the industrial production index, and log of average WTI oil prices. "No Time FEs" corresponds to the case when neither time-fixed effects nor aggregate control variables are used. We winsorize the top and bottom 5% of dependent variables for each tenure group and period.

individual-fixed effects correspond to age, period, and cohort, respectively.[13]

[13] This correspondence is not obvious. Note that first, there is a correspondence between a cohort dummy variable and (a sum of) individual-fixed effects. To see this correspondence, imagine that respondents are "born" when they enter into a survey. Then, a sum of individual dummy variables of respondents who are "born" in period t will be identical to a cohort dummy variable for the respondents who are "born" in period t. Second, note that when there is no panel attrition, # of Survey Experiences (Age) = Current Period (Period) - Survey Entrance Period (Cohort, the date of birth) holds. That is, survey experience dummies (Age), time dummies (Period), and individual dummies (Cohort) are going to be co-linear if they are all used in same time frequency in a linear panel regression.

One simple solution to this APC problem is to replace monthly time-fixed effects with quarterly ones, which we do for the remainder of the paper. Other solutions may include normalization of the parameters (Deaton and Paxson, 1994), replacing the time-fixed effects with aggregate variables (Heckman and Robb, 1985), or omitting time-fixed effects altogether. Panel B of Figure 1.2 reproduces our baseline results in Panel A of Figure 1.1 with each of these alternatives. Our results remain quantitatively and qualitatively similar.[14]

Since central banks often monitor median rather than mean inflation expectations, we also estimate tenure effects on medians using a fixed effects panel quantile regression (Machado and Silva, 2019). The estimated median effects are very close to the mean effects from our baseline regression.

1.2.4 Tenure Effects for Other Survey Measures

Since the SCE includes a variety of other questions about expectations, we estimate analogous regressions for additional outcome variables. In the first column of Table A.3, the dependent variable is the respondent's reported percent change that the unemployment rate will be higher in 12 months. Since unemployment fell steadily from June 2013 to February 2020, we restrict the sample to June 2013 to February 2019, so that lower responses are more accurate. The coefficient estimates indicate that responses indeed become more

[14] One possible reason for this robustness could be the fact that inflation rates have been very stable in recent years. Thereby, the effects of time-fixed effects could have been weak during our sample periods; the overall R^2 is virtually identical when we drop the quarterly time-fixed effects entirely.

accurate with higher survey tenure.

In columns 2 through 7, the dependent variables are expected price changes for gas, food, medical expenses, college education, rent, and gold over the next 12 months. These questions are only asked of respondents with tenure of at least 2. For all price categories except gold, expectations are 1.2 to 2.7 percentage points lower for respondents with tenure 12. The tenure effects for gold price expectations are smaller, with a coefficient estimate of 0.9 percentage points for tenure 12. Respondents of tenure 2 or greater are also asked for a density forecast of national house price growth over the next 12 months. Column 8 of Table A.3 shows that the density forecast interquartile range shrinks with tenure, and is about 1.3 percentage points smaller for tenure 12 respondents compared to tenure 2 respondents.

We also estimate analogous regressions for respondents' nominal personal earnings growth expectations and household income expectations. The estimated effects for personal earnings and income expectations are much smaller when compared to those for inflation expectations. See Columns 9 and 10 of Table A.3.[15]

The larger tenure effects for inflation expectations compared to income

[15] One may point out that inflation rates could be naturally harder to forecast than respondents' own income path. That is, for example, one can argue that 0.5 percentage points of learning effects in income expectations should not be treated equally to the same magnitude in inflation expectations. Reflecting this argument that different forecasts may have different scales, we normalize the estimated learning effects by standard deviation or mean of each forecast. However, we still find that the estimated learning effects for inflation expectations are more than 50 percent larger than those for earnings and income expectations.

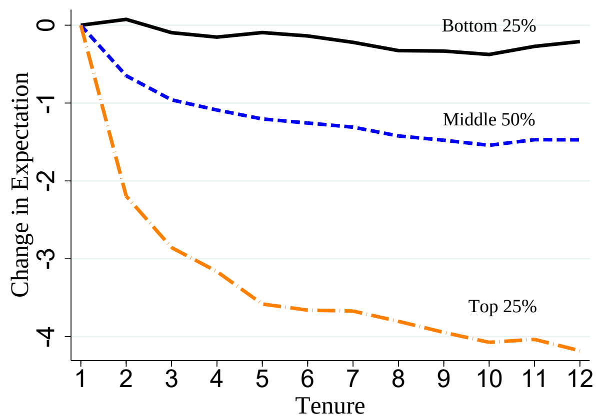
expectations are in line with rational inattention theory, which suggests that households may selectively pay attention to economic variables. Households should be highly attentive to their own income process even prior to participating in the survey, because it is so relevant to their consumption decisions. However, especially in low-inflation environments, consumers with limited information-processing capacity may pay little attention to inflation. Thus, households may not have a good understanding of the nation-wide average price process before taking the survey. For example, Carroll et al. (2020) show that consumers tend to underreact to aggregate macroeconomic shocks. Consumers may neglect aggregate variables in their consumption decisions because aggregate shocks consists only a small proportion of the uncertainty that consumers face, compared to highly idiosyncratic variables like their own income.[16] Therefore, questions that ask for the respondents' beliefs about inflation are more likely to prompt additional attention to inflation, prompting participants to collect more information about inflation.

1.2.5 Heterogeneity in Tenure Effects

While the equation (1) estimates tenure effects for the average respondent, these effects may be heterogeneous depending on households' initial expectations and uncertainty. For example, households who enter the survey

[16]In section II.B of this paper, we provide more direct evidence consistent with Carroll et al. (2020); consumption expectations of new survey participants respond more sluggishly to inflation expectations, while repeat participants more promptly reflect inflation expectations to consumption expectations.

Figure 1.3: Learning Effects on Inflation Expectations by Initial Inflation Uncertainty



Note: The figure plots the learning-through-survey effects by initial inflation uncertainty ($IQR_i \in \{U_H, U_M, U_L\}$), which are estimated from the equation (2). For example, the top 25% line (long dashed orange line) corresponds to the case when a respondent had a high level of inflation uncertainty in the first interview, assuming $\alpha_i = 0, \gamma_t = 0$. The y-axis shows the change in one-year-ahead inflation expectation of respondents compared to their initial responses, in percentage points. Tenure is shown on the x-axis and corresponds to the total number of survey experiences of each respondent (including the current survey wave). Sample is restricted to non-attriters. We truncate the top and bottom 5% of the dependent variable for each tenure group and period.

with high uncertainty may be more susceptible to learning-through-survey effects since their priors are weaker. To allow for such heterogeneity, we extend equation (1) by including interaction terms of tenure dummies with initial inflation uncertainty of respondents from their first survey:

$$\pi^e_{its} = \sum_{s=2}^{12} \left\{ \beta_{1,s} + \beta_{2,s} IQR_i + \beta_{3,s} IQR_i^2 \right\} \tau_s + \alpha_i + \gamma_t + \varepsilon_{it}, \tag{1.2}$$

where π^e_{its} denotes one-year-ahead expected inflation of individual i wof tenure s in period t, and IQR_i is the interviewee's *initial* IQR reported in the first survey. We include the squared term IQR_i^2 to control for possible non-linearity. All other terms are defined as in regression (1). The coefficients on the interaction terms are highly statistically significant.

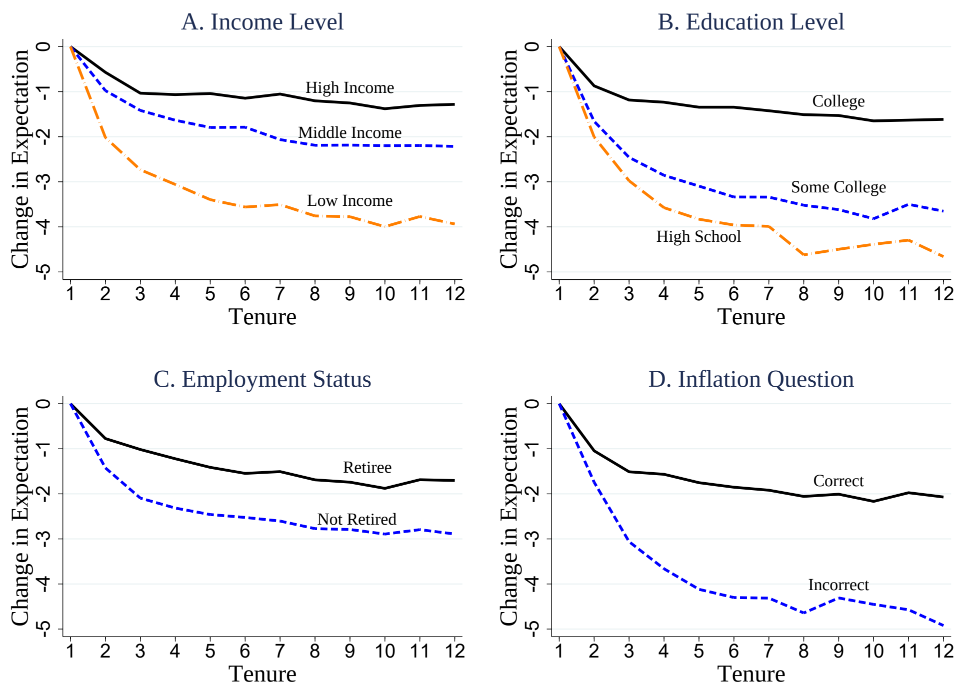
In Figure 1.3, we plot the estimated tenure effects for values of IQR_i corresponding to the 25th, 50th, and 75th percentiles, which we denote U_L, U_M, and U_H. These values are 1.5%, 3.0%, and 6.4%, respectively. Specifically, using regression (2), assuming $\alpha_i = 0, \gamma_t = 0$, we plot

$$\{\frac{\partial \pi^e_{its}}{\partial \tau_s} \mid \; IQR_i \in \{U_H, U_M, U_L\}\}_{s=2}^{12}$$

Respondents who initially entered the survey with a high level of inflation uncertainty learn more about inflation. Figure 1.3 shows that respondents whose initial uncertainty over the future inflation rate was in the bottom quartile of the distribution display a small learning effect. However, if the respondents were initially in the top quartile of inflation uncertainty, then the effect is large: right after the first survey, inflation expectations decrease by 2.2 percentage points on average.

We find further evidence of the heterogeneity of the learning effect when including demographic variables and measures of respondents' understanding of inflation as interactions. For a categorical variable D_i describing a characteristic of respondent i, we estimate the below regression and calculate $\beta_{1,s} + \beta_{2,s} D_i$.

Figure 1.4: Learning Effects by Demographics and Understanding of Inflation



Note: Each panel plots learning-through-survey effects on one-year-ahead inflation expectation by demographic variables and inflation understanding: income level, education level, retiree status, and whether respondents gave a correct answer to a question asking about inflation. Estimates are obtained from regression equation (3) using the indicated dummy variables as the interaction term. Sample is restricted to non-attriters. We winsorize the top and bottom 5% of the dependent variable for each tenure group and period.

$$\pi^e_{its} = \sum_{s=2}^{12} \left\{ \beta_{1,s} + \beta_{2,s} D_i \right\} \tau_s + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.3}$$

We use three demographic variables: income level (less than \$50k, \$50k to \$100k, more than \$100k), education level (college, some college, or high school), and retiree status. In addition, in the SCE, new respondents are required to answer a set of questions measuring their numeracy and financial literacy. A question designed to measure understanding of inflation asks, "Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After one year, how much would you be able to buy with the money in this account?" Respondents can choose "More than today," "Exactly the same," or "Less than today." We measure respondents' understanding of inflation by whether respondents gave a correct answer to this question. Only 52% answered this question correctly.

Consistent with the previous results from Figure 1.3, Figure 1.4 shows that respondents who are generally more informed about inflation prior to the survey display significantly smaller learning effects throughout the survey waves. Panels A and B show that the estimated learning effects are substantially smaller for higher-income and more educated individuals, while Panels C and D show that retirees and respondents who gave a correct answer the question measuring understanding of inflation display relatively smaller learning effects.[17] The coefficients on the interaction terms for income, education, retiree status, and inflation understanding are all highly statistically significant. We also run the same regression jointly including all indicator variables for demographics and inflation understanding; the overall results are similar to those in Figure 1.4, though the interaction terms with inflation understanding loses statistical significance..

[17]Aguiar and Hurst (2007) find that older households invest more in shopping time and pay the lowest prices compared to other households.

In summary, tenure effects are a robust feature of the SCE data. On average, consumers with more past survey experience have inflation expectations that are lower and more accurate. They also tend to have lower uncertainty about inflation and are less likely to update their forecasts in subsequent surveys. Further, survey participants who are generally more informed about inflation prior to the survey display smaller tenure effects.

1.3 Implications for Interpretation of Survey Data

Inflation expectation surveys conducted by the central banks are generally intended to be used for two major purposes: i) monitoring inflation expectations through an aggregate index and ii) researching consumer expectations and behavior using the underlying micro data. The results from the previous section imply that aggregate measures of inflation expectations and of inflation uncertainty would be higher if only new participants—those not subject to tenure effects—were included. This was shown in Panels C and D of Figure 1.1.

Using an episode of oil price collapse in 2014, we show how the tenure effects can potentially impede central banks' monitoring of inflation expectations. Then, by revisiting estimation of elasticity of intertemporal substitution by Crump et al. (2015), we show how the learning effect can influence micro estimates and provide useful insights for studies using survey micro data.

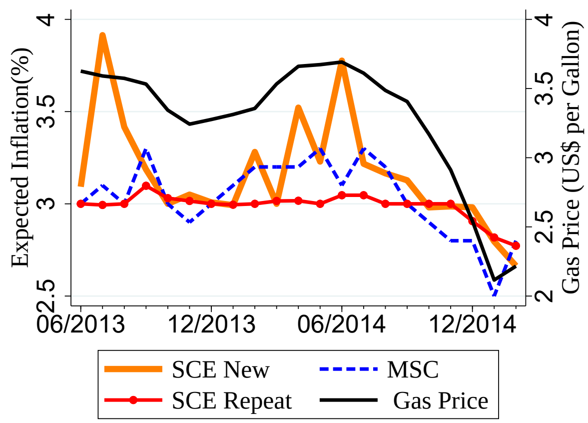
1.3.1 Oil Price Decline in 2014

In order to monitor the inflation expectations of U.S. consumers, the FRBNY conducts the SCE data each month and reports the sample median of the density mean inflation expectation. However, the tenure effects we documented suggest that repeat participants' prior survey participation may have prompted them to seek information about or otherwise reflect on inflation. Exploiting the episode of a sharp drop in oil prices during 2014, we show that the dynamics of inflation expectations of new participants can be significantly different than those of the repeat survey participants.

In particular, household inflation expectations are sensitive to oil and gas prices. Coibion and Gorodnichenko (2015b) find that the increase in inflation expectations of households during the Great Recession can be attributed to the rise in the oil price, since the price of gasoline is one of the most salient prices for consumers. However, this stylized fact tends to be significantly weaker for repeat participants of the FRBNY SCE compared to new participants.

First, Figure 1.5 shows that crude oil prices plunged by half in only six months in 2014, from \$103.59 per barrel in July 2014 to \$50.58 in February 2015, as innovation in Hydraulic Fracturing technology boosted oil production in the U.S. During this period, other macroeconomic conditions were fairly stable; the seasonally-adjusted industrial production index decreased by 0.17

Figure 1.5: Inflation Expectation of New and Repeat Participants



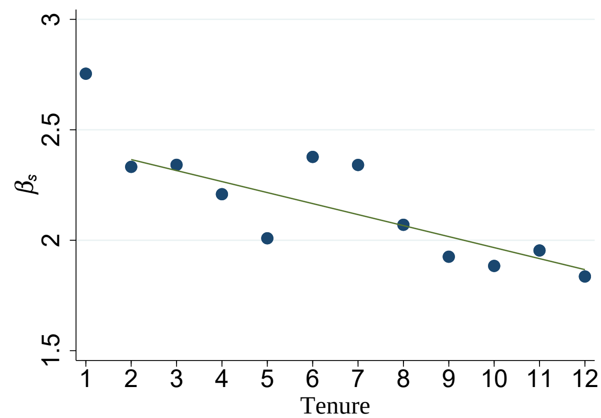
Note: The monthly average nominal WTI crude oil price per barrel in US$ is on the right y-axis (thin solid black line; "Oil Price"). For the left y-axis, one-year-ahead median density mean inflation expectations of new survey participants of the SCE (thick solid orange line; "SCE New"), repeat survey participants of the SCE (connected red line; "SCE Repeat"), and the median inflation expectations of Michigan Survey of Consumers (dashed blue line; "MSC") are presented in percentage points. Data is from the FRBNY Survey of Consumer Expectations and Federal Reserve Economic Data.

percent and the unemployment rate decreased by 0.7 percentage points.[18]

Figure 1.5 also compares the median density mean inflation expectations of new SCE participants with those of repeat participants, whose expectations have been subject to the tenure effects. The inflation expectations of repeat participants are relatively steady from June 2013 to February 2015, only declining by 0.27 percentage points from July 2014 to February 2015 as oil prices fell. By contrast, the inflation expectations of new participants gen-

[18]See Baffes et al. (2015) for more discussion on the causes of the oil price decline, including weakening global demand, a significant shift in OPEC policy, geopolitical shifts, and U.S. dollar appreciation.

Figure 1.6: Responses to Gas Prices by Survey Tenure



Note: The regression coefficients $\{\beta_s\}_{s=1}^{12}$ obtained from our benchmark regression (4), which measures the response of density mean inflation expectation to the increase in log gas prices ($= \partial \pi^e / \partial log(Gas)$), are presented in the figure by each tenure group, s. Tenure is shown on the x-axis and corresponds to the total number of survey experiences of each respondent, including the current survey wave. A linear fitted line is presented for repeat participants (tenure>1). Data is from the FRBNY Survey of Consumer Expectations, July 2014 to February 2015 February. A full regression table is available in Appendix Table A.4.

erally track high-frequency fluctuations in oil prices, which is consistent with what Coibion and Gorodnichenko (2015b) have found. Inflation expectations from the Michigan Survey of Consumers (MSC) show a similar pattern to the inflation expectations of new participants in the SCE.[19]

We quantitatively evaluate the differences in responses to gas prices between new and repeat participants using the following panel linear regression:

$$\pi^e_{its} = \sum_{s=1}^{12} \beta_s(\tau_s \times log(Gas_t)) + \alpha_i + \gamma_t + \varepsilon_{it} \tag{1.4}$$

[19]MSC has a rotating panel component, but respondents are surveyed at most twice, with six months between interviews. Tenure effects on the MSC are discussed in a later section.

where $log(Gas_t)$ is the log of the monthly gas price, and τ_s is a tenure dummy variable for s number of total survey experience. π^e_{its} denotes one-year-ahead density mean inflation expectations of an individual i whose total number of survey experience is s at period t. α_i and γ_t are individual- and quarterly time-fixed effects. ε_{it} is an error term. The sample period is from July 2014 to February 2015, the period when oil and gas prices plunged. We restrict samples to respondents who participate in the survey for the maximum number of times, as we did in our main result section.

Figure 1.6 visually shows the estimated regression coefficients $\{\beta_s\}_{s=1}^{12}$ by tenure group s. Clearly, new survey participants display the largest regression coefficients indicating the strongest response of inflation expectations to gas prices. For our benchmark regression, we find that the inflation expectations of new survey participants respond about 50 percent more strongly to gas prices on average when compared to the most experienced participants.[20] This is consistent with our previous finding from the aggregate times-series data. Table A.4 presents various regression specifications, including those with truncation of extreme expectations, full sample periods, and point inflation expectations. Qualitative features of our results are not changed. Rather, our benchmark regression specification tends to be conservative when compared to the results from other specifications. When we truncate 10 percent

[20]We have conducted an F-test for H0: $\beta = \beta_s \ \forall s$ over extended sample periods (from March 2014 to June 2015) and full sample periods (from June 2013 to October 2020) with Driscoll-Kraay standard errors of lag one. For both cases, we could reject the null hypothesis that the regression coefficients across survey tenure are equal to each other at 1% significance level.

of extreme expectations, new participants respond almost twice as much as repeat participants to gas prices.

Our results are also consistent with those of Verbrugge and Binder (2016), who partitioned the MSC respondents into those with low and high inflation uncertainty, using the methodology in Binder (2017b). They show that the inflation expectations of less-uncertain consumers are more stable than those of more-uncertain consumers. In particular, the expectations of less-uncertain consumers did not respond strongly to the oil price decline in 2014.

This evidence suggests that the tenure effects we documented are not constant over time, and thus cannot be removed simply by taking a first difference. The expectations of repeat and new participants can exhibit different dynamics in response to economic shocks. In such a case, repeat participants cannot be viewed as representative of the broader population who potentially lack any past survey experience. In this example, if the central bank were only given inflation expectations of repeat participants, they would conclude that the inflation expectations of consumers do not respond to the plunging oil prices and miss some of the timely high-frequency information from survey expectations.

1.3.2 Estimating Elasticity of Intertemporal Substitution

The survey micro data on consumer expectations has begun to be used in a variety of applications, and holds great potential for use in many more.

The tenure effects that we have documented may affect the estimates and interpretation of such studies. For example, Crump et al. (2015) use the SCE data to estimate the elasticity of intertemporal substitution (EIS). More precisely, they estimate the response of expected consumption growth to changes in expected inflation rates. We revisit this analysis by allowing estimates to vary by respondents' survey experience. Among regression specifications of Crump et al. (2015), for simplicity we focus on the following panel linear regression model with fixed effects:[21]

$$ExpCG^i_{t,t+12} = -\sigma ExpInf^i_{t,t+12} + \gamma ExpIG^i_{t,t+12} + \alpha_i + \beta_t + \varepsilon_{i,t}, \tag{1.5}$$

where $ExpInf^i_{t,t+12}$ is a 12-month ahead density-implied mean inflation expectation of household i at period t, and $ExpCG^i_{t,t+12}$ is expected real consumption growth over the next 12 months by household i at period t, which is calculated as, $ExpCG^i_{t,t+12} \equiv ExpSG^i_{t,t+12} - ExpInf^i_{t,t+12}$, when $ExpSG^i_{t,t+12}$ is a point forecast for nominal spending growth of the household over the next 12 months. Similarly to the calculation of $ExpCG^i_{t,t+12}$, expected real household income growth, $ExpIG^i_{t,t+12}$, is the difference between point forecast for household nominal income growth and $ExpInf^i_{t,t+12}$. α_i and β_t are individual- and time-fixed effects.

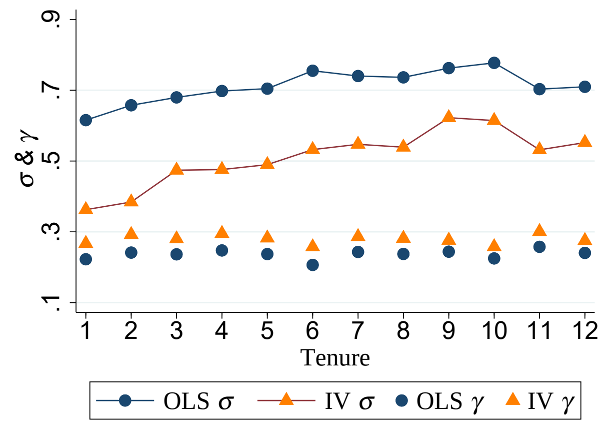
[21]The most recent version of Crump et al. (2015) uses a panel linear regression model *without* fixed effects as their baseline since the SCE data allows many control variables. However, Crump et al. (2015) also show results based on a model with fixed effects and emphasize that their main results remain similar. See section 6.4 and table 9 of Crump et al. (2015).

The above expression represents the first-order approximation of a usual consumption Euler equation where σ is the elasticity of intertemporal substitution (EIS) and γ measures "excess sensitivity" of consumption growth to anticipated income changes. The literature commonly finds that expected/predictable income growth has a significant effect on consumption growth. Inclusion of γ in the regression model therefore reflects a possible deviation from the permanent income hypothesis.

First, we estimate the above regression model as-is and find that our estimates on σ and γ are indeed very similar to those of Crump et al. (2015); $\hat{\sigma}$ of Crump et al. (2015) under fixed effects is 0.71 and $\hat{\gamma}$ is 0.20 while our estimates are 0.70 and 0.24.[22] Next, we replace σ and γ with $\sum_{s=1}^{12} \sigma_s \tau_s$ and $\sum_{s=1}^{12} \gamma_s \tau_s$ and re-estimate the regression, where τ_s is an indicator variable for tenure s. This modification allows the regression coefficients σ and γ to vary by survey experience of respondents non-parametrically. Also following Crump et al. (2015), we use an instrumental variable (IV) strategy that uses the point inflation expectation as an instrument of density-implied mean inflation expectation, and again allow coefficients to vary by tenure. A full regression table is available in Table A.5 of the appendix.

[22]Our estimates are slightly different from those of Crump et al. (2015) since we use the same sampling rule as in the rest of this paper, truncating the top and bottom 5% of all point forecasts for each tenure group and period. We restrict samples to respondents who participate in the survey for the maximum number of times in order to minimize the effects of panel attrition. Finally, we use quarterly time-fixed effects instead of monthly time-fixed effects. While our sampling rule is different from that of Crump et al. (2015), as mentioned in the main text, our baseline estimates are very similar to their results, suggesting sampling rule did not drive our results in this section.

Figure 1.7: Estimates of EIS and Excess Sensitivity of Consumption by Survey Tenure



Note: We run a linear panel regression as in Crump et al. (2015), allowing regression coefficients to vary by survey experience : $ExpCG^i_{t,t+12} = -\sum_{s=1}^{12} \tau_s \sigma_s ExpInf^i_{t,t+12} + \sum_{s=1}^{12} \tau_s \gamma_s ExpIG^i_{t,t+12} + \alpha_i + \beta_t + \varepsilon_{i,t}$. The estimated regression coefficients $\{\hat{\sigma}_s\}_{s=1}^{12}$ (EIS) and $\{\hat{\gamma}_s\}_{s=1}^{12}$ (Excess Sensitivity) are presented in the figure. For the case of IV, the point inflation expectation is used as an instrument of density-implied mean inflation expectation. The sample is restricted to non-attriters. We truncate the top and bottom 5% of all point forecasts for each tenure group and period. Data is from the FRBNY Survey of Consumer Expectations, June 2013 to October 2020.

Figure 1.7 shows that the estimated EIS, $\hat{\sigma}$, increases with survey experience, using either the OLS or IV estimates. That is, more experienced survey participants tend to more actively reflect changes in inflation expectations in their consumption expectations. However, $\hat{\gamma}$ is similar for all tenure groups. Note that the range of EIS estimates in Crump et al. (2015) (around 0.5 to 0.8) are near the lower end of the range of micro estimates from prior literature. Our results indicate that the estimated EIS tends to moderately increase with survey experience. Thus the EIS of the general population who lack any prior survey experience is likely to be even lower than the original estimates of Crump et al. (2015). Survey participation that induces learning about inflation (or equivalently, greater attention to inflation) may result in larger responsiveness to reported inflation expectations by survey respondents. In this example, it results in higher estimate of the EIS.

Why does $\hat{\sigma}$ tend to be larger for experienced survey participants, but not $\hat{\gamma}$? Consistent with our findings throughout the paper, consumers' imperfect attention to aggregate shocks can account for this otherwise puzzling phenomenon. When it comes to spending decisions, consumers tend to focus on their income, but may not pay careful attention to general inflation rates. Therefore, they have sluggish responses to aggregate shocks (Carroll et al., 2020). In other words, if consumers become more attentive to inflation rates because survey experience, their consumption expectations may more quickly respond to change in future inflation (larger $\hat{\sigma}$ with survey experience). In contrast, how households' reported consumption plans respond to expected future income may not change after taking more surveys. Prior to taking a survey, they may already understand their own future income path well and have an established rule for how to adjust their consumption plan with future income changes.

This exercise shows how tenure effects can influence micro estimates. Awareness of such learning effects is useful for interpretation of analysis using survey micro data. We suggest that it would be good practice for users of survey micro data to check whether their estimates are robust using subsamples of shorter-tenured and longer-tenured respondents.

1.4 Other Surveys

One potential concern is that the tenure effects might arise from a particular feature of the SCE, such as the short time period of relatively low and stable inflation in which the SCE has been conducted. This section uses the Michigan Survey of Consumers and a survey of U.S. firms to provide evidence of tenure effects in other contexts.

1.4.1 Michigan Survey of Consumers

The Michigan Survey of Consumers (MSC), like the SCE, is a monthly survey of consumer expectations. However, whereas the SCE consecutively tracks respondents up to twelve times, the MSC only allows respondents to participate in a maximum of two interviews, with a six-month gap between interviews. A longer gap between surveys tends to reduce the size of tenure effects (Warren and Halpern-Manners, 2012). Despite its more limited panel structure and longer gap between surveys, the MSC does have the advantage of beginning in 1978, rather than 2013, allowing us to check how tenure effects have varied over time and to confirm that they are not only a feature of recent data.

First, Panel A of Figure 1.8 shows the mean inflation expectations of new and repeat respondents. As before, we apply the non-attrition restriction, so our samples consist of respondents who eventually participated in a follow-up survey.[23] The mean expectations of new respondents are typically slightly higher than those of repeat respondents—on average, the gap is 0.3 percentage points. As expected, this effect is smaller than the effect found in the SCE data. This smaller learning effect in the MSC compared to that in the SCE is consistent with recent evidence from a randomized information treatment experiment that shows providing information about inflation to households has large contemporaneous effects on their expectations but that these effects rapidly diminish over time (Coibion et al., 2020).

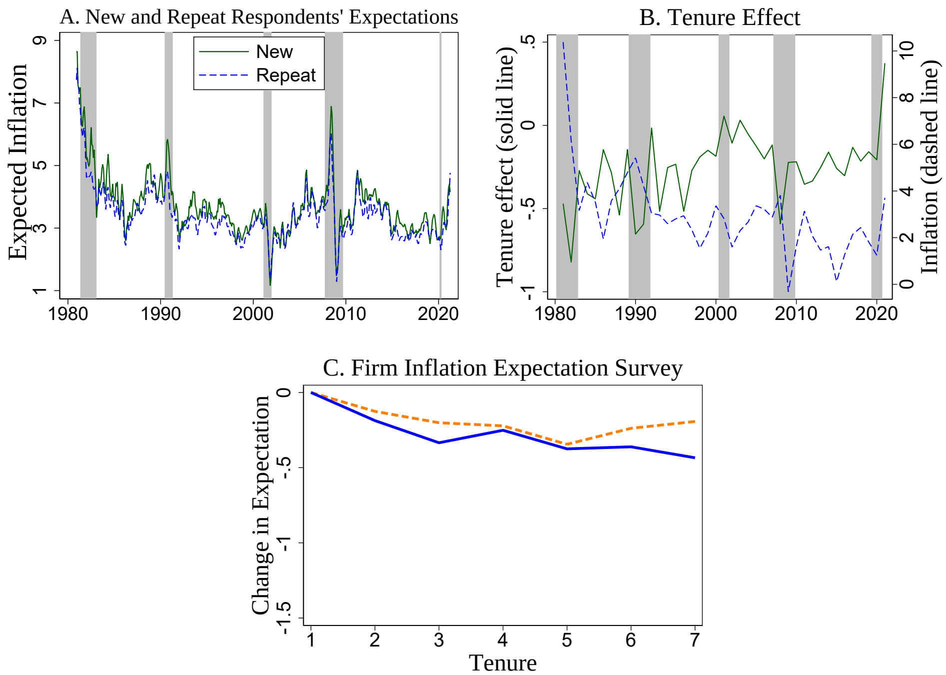
Panel A also shows that the size of the gap between repeat and new respondents' expectations can vary over time. In order to construct a time-series of tenure effects, we use a regression equation analogous to equation (1), but with fixed effects replaced by demographic control variables to account for the limitations of the MSC dataset. We estimate the following equation for each year t separately—since estimates at the monthly frequency are quite noisy—and obtain a sequence of yearly tenure effects, $\{\hat{\delta}_t\}_{t=1}^{T}$:

$$\pi^e_{its} = \alpha_t + \delta_t \tau_2 + \beta_t X_{it} + \varepsilon_{it} \tag{1.6}$$

where π^e_{its} denotes the one-year-ahead point inflation forecast of individual i in year t with tenure $s \in \{1, 2\}$, τ_2 is an indicator variable for tenure 2, X_{it} is a vector of control variables including sex, education, region, the number of

[23]Summary statistics of expectations by tenure without the non-attrition restriction are in Appendix Tables A6 and A7 (for the SCE) and Table A.8 (for the MSC).

Figure 1.8: Learning-Through-Survey Effects on the Other Surveys



Note: Panel A shows the mean inflation expectations of new and repeat respondents from the Michigan Survey of Consumers (MSC), January 1981 to September 2021. Centered three-month moving average is shown for visual clarity. Sample is restricted to non-attriters and responses are winsorized as in previous section.

Panel B shows an annual time-series of learning-through-survey effects, which are estimated from a regression equation (6). The solid line is a plot of $\{\hat{\delta}_t\}_{t=1}^{t=T}$, which are regression coefficients attached to the tenure dummies. If $\hat{\delta}_t$ is negative, then the second-time interviewees have lower inflation expectations than the first-time interviewees in the period t. The dashed line is CPI inflation. Shaded bars indicate NBER recessions.

Panel C shows tenure effects from a U.S. firm survey. The percentage points change in inflation expectations of survey participants compared to their initial responses is presented on the y-axis. The solid blue line corresponds to the results from Deaton's method, which normalizes quarterly dummy variables following Deaton and Paxson (1994). The dashed orange line corresponds to the results when macroeconomic aggregate variables are used to control for time effects in a linear panel fixed effects regression, including monthly CPI inflation rates, the S&P 500 stock price return, unemployment rate, and the log of average WTI oil prices. We restrict samples to consist of firms who eventually participate in the survey more than three times and winsorize the top and bottom 5% of the data. The sample period is 2018Q2 to 2020Q2.

kids, marital status, log of nominal household income, age, and age squared, and ε_{it} is an error term.

Panel B of Figure 1.8 shows resulting estimates of the yearly tenure effects, along with CPI inflation and shaded bars indicating recessions. These tenure effects δ_t are nearly identical to the difference between the mean expectations of repeat and new respondents in year t, and have mean 0.3 and standard deviation 0.2.

We see that tenure effects vary over time, and tend to be larger in magnitude during recessions, when economic uncertainty is high (the exception is the early 2000s recession). This suggests that households form inflation expectations in a Bayesian manner, putting more weight on new information when they are more uncertain in their beliefs. The magnitude of the tenure effects are also positively correlated with inflation uncertainty, disagreement, and volatility, with correlation coefficients of 0.2, 0.6, and 0.3, respectively.[24]

Tenure effects also vary with inflation. In particular, the largest negative values of δ_t occur when inflation is falling or about to fall, and the near-zero or positive values occur when inflation is rising or about to rise. Most notably, the most negative value of δ_t (-0.82) occurs in 1982. Inflation fell from 10% in 1981 to 6% in 1982 and would fall to 3% in 1983. New respondents, less informed about falling inflation, reported much higher inflation expectations

[24]Inflation uncertainty is the updated inflation uncertainty index from Binder (2017b). Inflation disagreement is the cross-sectional standard deviation of inflation expectations from the MSC. Inflation volatility is the rolling five-year standard deviation of annual CPI inflation.

than repeat respondents. A similar pattern occurs in 1990 and 1991: new respondents were less aware than repeat respondents that inflation was beginning to decline. Another striking example occurs in 2008. Though inflation fell during the Great Recession, many consumers expected the Recession to be inflationary, so expectations rose sharply, especially for new respondents. Repeat participants' expectations were 0.6 percentage points lower than those of new respondents. Conversely, the most positive value of δ_t (0.37) occurs in 2021. In most years, consumer inflation expectations are higher than realized inflation, so consumers revise their forecasts downward as they become more informed. Recently, as inflation has risen sharply, more informed consumers revise their expectations upward.

To summarize, the results from the MSC show that during most periods, repeat participants generally report lower inflation expectations than new participants, though the degree of the learning-through-survey effects changes over time. In addition, the learning-through-survey effect tends to be smaller for the MSC than those of the SCE, likely because there is a longer period of time between baseline and follow-up surveys.

1.4.2 Inflation Expectations of Firms

While our focus so far has been household surveys, Coibion et al. (2018) show that the inflation expectations of firm managers tend to resemble those of households, which suggests that tenure effects may exist in firm surveys as well. To study whether this is the case, we use a new firm expectation survey targeting businesses in the U.S. (Candia et al., 2021)

This firm survey, the new Survey of Firms' Inflation Expectations (SoFIE), is collected by a business intelligence company that has been collecting CEOs' and top executives' perceptions and expectations for various firm-specific economic outcomes. The panel is intended to be representative of "the underlying structure of each sector in the economy according to its contribution to the gross value added."[25] The survey covers the U.S. firms in manufacturing and services sectors. About 300 to 600 firms participate in this survey each wave and stay in the panel for about three waves on average. A question asking about one-year-ahead CPI inflation rates was added its quarterly survey in 2018. We use this data, which was collected from April 2018 to April 2020 at quarterly frequency.

We estimate the learning-through-survey effects for firms in regressions analogous to equation (1). However, since only nine survey waves are available, we relax our "non-attrition" restriction to save observations; we restrict our samples to consist of firms who participated in the survey more than three times. As before, we either use Deaton's normalization method or aggregate control variables to avert the APC problem[26], and winsorize the top and bottom 5% of data.

[25]Candia et al. (2020) and `http://firm-expectations.org` provide more detailed information about this firm survey. In particular, information about survey representativeness is at `http://firm-expectations.org/weightsbuilding.html`.

[26]We cannot use the quarterly time fixed effects method for this case since our data is quarterly.

Panel C of Figure 1.8 shows the resulting estimates of the tenure effects from the firm survey. Inflation expectations of repeat survey participants decrease with survey experience on average, consistent with what we found in the consumer surveys. This provides additional evidence that firm executives, who are likely the price-setters in the economy, typically face information constraints that may influence their expectations of aggregate inflation. For example, Mackowiak and Wiederholt (2009) show that firms pay more attention to idiosyncratic variables because firm-specific conditions are generally more important in their decision-making than aggregate conditions.

Again, tenure effects in the firm survey are smaller than those in the SCE, possibly because the time between surveys is longer. Firm respondents of tenure 3 (6 months after first participation) have about 0.20 to 0.33 percentage points lower inflation expectations than new participants. In the SCE, after 6 months of participation, repeat survey participants have about 2.2 percentage points lower inflation expectations than those of new participants. Another reason why the tenure effect in a firm survey is smaller than that of household surveys could be that the initial inflation expectations of firms are generally more accurate than those of households; after the winsorization, the average inflation expectation of firms is 2.8 percent, with standard deviation 1.86.

While the dataset is limited, the firm survey results confirm that the tenure effects are not exclusive to household surveys. It also shows that the degree of panel conditioning may depend on time between surveys and respondents' prior level of knowledge on the subjects being asked.

1.5 Discussion and Conclusion

We have shown that the reported beliefs of survey respondents significantly change over their tenure in the survey. We have documented the prevalence of these tenure effects across two surveys of consumers and a survey of firms. The size of the effects vary depending on the type of survey question, survey frequency, and respondent characteristics. The effects are particularly large and robust for inflation expectations reported on the FRBNY SCE: consumer inflation expectations and uncertainty decrease notably with survey tenure, especially after the first round of survey participation.

We believe that our results have implications both for our understanding of the expectations formation process and information rigidities, and for the design and interpretation of expectations surveys in the future.

In the literature on consumer expectations formation, two major questions that arise are (1) How do consumers allocate their limited attention and cognitive efforts?, and (2) Why do consumers disagree so much with each other and with professional forecasters? In some of this literature, households allocate their attention and cognitive efforts based on the cost and benefit of information acquisition, and as they acquire new information, they update beliefs in a Bayesian manner. For example, when households' priors are less certain, they put more weight on new information. Then, differences in beliefs can arise from different information processing constraints and different perceived benefits in acquiring certain types of information.

Households may pay little attention to inflation, either because it is costly for them to do so, or because they find it more worthwhile to devote their attention to other things. Our primary interpretation of our main results is that when a survey asks respondents about their inflation expectations, it may prompt them to devote more attention to inflation before retaking the survey (such as by looking up official inflation statistics or media reports, talking to acquaintances, or reflecting on their own experiences and the prices they have observed). Indeed, respondents significantly and predictably revise their expectations in subsequent surveys. The revisions are consistent with Bayes' rule, in the sense that forecast errors shrink and that the impact of past survey experience is larger for households who are generally less informed about inflation prior to the survey. The tenure effects are smaller in the firm survey, as respondents may be more informed about inflation prior to their first survey experience.

Moreover, the panel conditioning effects are minimal for questions about respondents' own earnings or income. Households are likely more attentive to these than to aggregate inflation, so the act of taking a survey does not much change their attention to these variables.

The implications of our findings for the design and interpretation of expectations surveys depends on whether our explanation for the panel conditioning effects is correct, and on details of how the survey data is being used. We have suggested that the panel conditioning effects arise because respondents pay more attention to inflation after being asked about inflation in one or more rounds of a survey ("cognitive stimulus" or "learning-through-survey.") One alternative explanation is a "reporting error" hypothesis.[27] Under this hypothesis, survey respondents have an underlying belief distribution about inflation that is *not* affected by survey participation. However, respondents with lower survey tenure answer questions with extra reporting error, because they must expend cognitive effort to formalize, retrieve and report their underlying beliefs accurately.

More formally, under the reporting error hypothesis, respondent i of tenure s reports an inflation expectation $r_{its} = \pi^e_{its} + \varepsilon_{its}$ where π^e_{its} is the respondent's true underlying inflation expectation in period t and ε_{its} is a reporting error. Suppose that as a respondent gains more experience in answering survey questions, then he becomes better in expressing his beliefs more accurately. That is, the distribution of ε_{its} becomes tighter as s increases. As we and others have documented, consumer inflation expectations are upward-biased, so ε_{its} is not zero in expectation, but rather positive. Suppose, for example, it has a log normal distribution, $\ln(\varepsilon_{its}) \sim N(0, \sigma_s^2)$, and σ_s^2 decreases with survey experience. This simple model can generate some of the effects documented in our paper: reported inflation expectation levels and uncertainty decrease with survey experience.

This reporting error would need to depend on demographic characteristics such as education and income in order to explain the heterogeneity in

[27]We thank an anonymous referee for providing constructive comments related to this issue.

learning effects we found. If the reporting error is simply a type of measurement error, with σ_s^2 constant over time, then this hypothesis does not explain the time-varying nature of the tenure effects, and their correlation with economic shocks. In our gas prices example in section II.A, for example, we showed that the tenure effects vary with gas prices. We also showed that the difference between new and repeat respondents' expectations was especially large at the start of the Covid-19 pandemic. So for the reporting error hypothesis to explain our results, the error or σ_s^2 would need to be allowed to depend on economic conditions. While the learning hypothesis seems more likely to us, we acknowledge that the reporting error hypothesis cannot be entirely ruled out. It is also possible that learning effects and reporting error effects are both present.

Future research to help disentangle these two hypotheses—perhaps by using focus groups and cognitive interviews of survey participants, or by adding special questionnaires that ask about behavior between survey rounds—will be quite important, because these hypotheses have different implications. The central issue is whether the change in responses throughout the survey waves reflect changing beliefs about the economy, or indicate that respondents are improving when expressing beliefs. Under the reporting error hypothesis, reported beliefs of respondents with longer tenure become *more* representative of the "true" underlying beliefs of the population. Under the learning hypothesis, in contrast, reported beliefs of respondents with longer tenure become *less* representative of population beliefs; instead they reflect the beliefs of people who are more attentive to inflation. For some applications, these beliefs may be of higher interest to researchers, if they more resemble the beliefs of price-setters or of consumers who are planning to make a major financial decision, for example.

Either way, central banks running household surveys to measure inflation expectations, and researchers working with the underlying microdata, should take note of our evidence of tenure effects. In a discussion of the SCE, Armantier et al. (2017, p. 64) argue that "the design of the panel, with a constant in- and outflow of respondents each month, ensures a stable survey tenure distribution, so the extent of learning and experience (and any associated impact on responses) is constant over time. As a result, month-to-month changes in median responses should capture real changes in population beliefs." But because of the time-varying nature of the tenure effects, month-to-month changes in median responses may not always capture changes in population beliefs in such a straightforward manner (as we showed in our gas prices example).

This is not to say that the panel component of the survey should be removed. The panel component has the clear benefit of allowing researchers to control for unobservable individual characteristics. If central banks wish to minimize the panel conditioning effects, one option is to increase the time length between surveys to minimize learning effects. Another option is to increase the size of the sample of new participants in each wave, but only invite some fraction of them to become repeat participants. This would allow researchers to conduct analysis on the full panel, on new participants only, or on repeat participants only, as appropriate to the situation. In addition, it would be good practice for users of survey microdata to check whether their estimates are robust to using subsamples of shorter-tenured and longer-tenured respondents.

Chapter 2

Common, Learning, and Long-run Components of Household Inflation Expectations

2.1 Introduction[1]

While it is conceptually easy to run a household survey asking about inflation expectations, producing a measure appropriate for monetary policy using survey data is far from straightforward. Inflation expectations of households are susceptible to survey design. Since 2013, the Federal Reserve Bank of New York (FRBNY) has been conducting the Survey of Consumer Expectations (SCE) monthly.[2] Kim and Binder (2022) show that the headline measures of the SCE suffer from learning effects of survey respondents. They show that the inflation expectations of repeat participants of the SCE tend to

[1]Some parts of Chapter 2 are published in the following citation:
Kim, Gwangmin, and Carola Binder. "Learning-through-survey in Inflation Expectations." *American Economic Journal: Macroeconomics*, Conditionally Accepted (2022).
Below briefly describes the nature of authors contribution:
Gwangmin Kim: Conceptualization, Writing - initial draft & editing, Data acquisition, Methodology, Formal statistical analysis, and Reviews. Carola Binder: Conceptualization, Writing - review & editing, Methodology, and Formal statistical analysis.

[2]See Arioli et al. (2017) and Appendix Table 1 of Coibion et al. (2019) for a list of countries running inflation expectation surveys targeting households. The European Central Bank (ECB) and the Bank of Canada are also in process of developing new household inflation expectation surveys with a survey design similar to that of the FRBNY (Bańkowska et al., 2021; Bellemare et al., 2020).

be much lower and more stable than those of new survey participants because of survey respondents' learning about inflation between survey waves. As a result, the SCE underestimates the inflation expectation level and uncertainty of the U.S. households. Similarly, Niu and Harvey (2021) show via a survey experiment that inflation expectations of households can be easily influenced by the information and context provided in a survey.

In addition to the difficulties in accurately measuring household inflation expectations, a wide variety of inflation expectations measures are increasingly available. The SCE provides household inflation expectations data in various dimensions: the type of survey respondents (new survey participants versus repeat survey participants), the type of question (point forecast versus probabilistic forecast), and the horizon of participant expectations. There are 8 different time-series indices of household inflation expectations available in the SCE alone. Extracting meaningful information about the co-movement of expectations out of many indices is challenging for monetary authority.

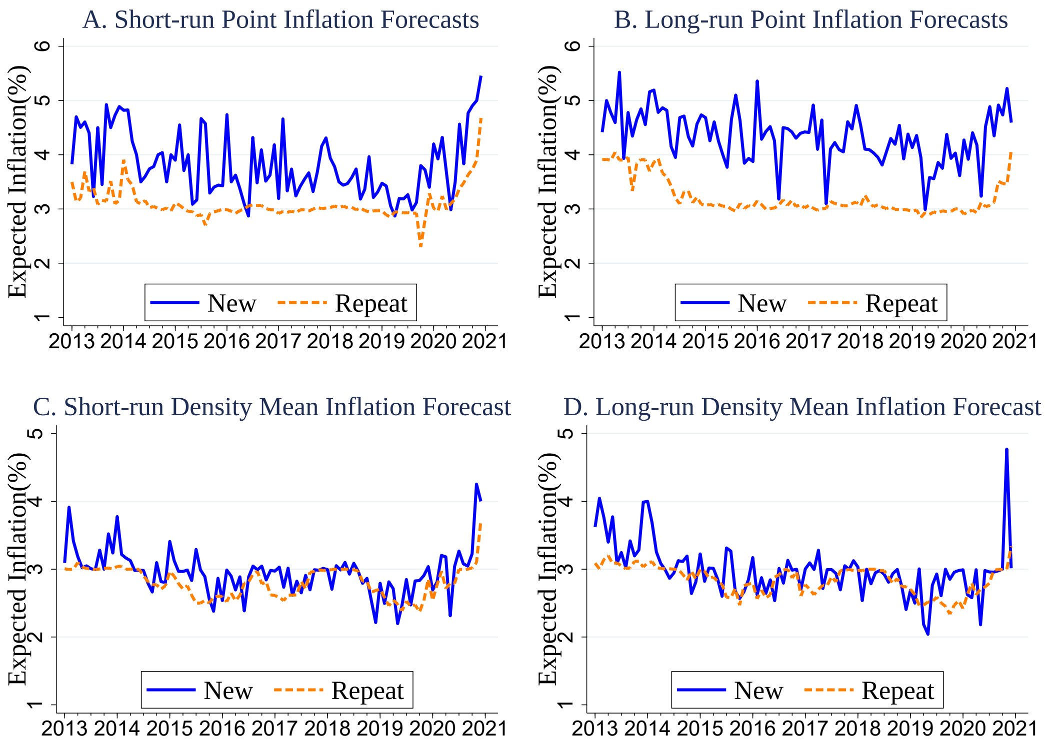
Figure 2.1 shows all 8 inflation expectation measures available in the SCE. Repeat participants display lower and more anchored inflation expectations in general when compared to new survey participants due to the learning effects described in Kim and Binder (2022). Notably, repeat survey participants comprise more than 80% of the total samples of the SCE. Only focusing on the data of new survey participants will require discarding major parts of the data. If a model only focuses on the smaller subset of new survey participants data, the aggregate statistics for inflation expectations will be noisy.

However, changing survey design or increasing the sample size is very costly and practically impossible. In short, all 8 indices are correlated with each other to some extent but can have different dynamics due to idiosyncratic components like learning effects.

In this study, I develop a statistical solution to reliably measure the household inflation expectations using the SCE. I use a dynamic factor model to extract the common factor among various household inflation expectation measures in the SCE. Dynamic factor models have been used for many applications in econometrics: forecasting (Stock and Watson, 2002a; Kotchoni et al., 2019), construction of economic indices (Ahn et al., 2020; Stock and Watson, 2002b; Forni et al., 2000), and structural modeling (Bernanke et al., 2005). Among these applications, the dynamic factor model has been most successful in summarizing a wide range of economic variables into a few indicators without missing important statistical patterns in the original data. Sargent et al. (1977) and Stock and Watson (2002b) find that a large fraction of variations in a number of economic indicators can be explained by a handful of factors.

The baseline model presented in the main result section successfully estimates the common factor that follows the dynamics of new survey participants' inflation expectations and corrects for the bias related to the learning effects (Kim and Binder, 2022). Also, the estimated common factor is significantly less noisy than the raw data of inflation expectations from new survey participants. The new measure utilizes all the data available without discard-

Figure 2.1: Average Learning-Through-Survey Effects on Inflation Expectations in the SCE



Note: Panels A and B show interpolated median point inflation forecasts of the SCE in percentage points. For Panel A, the one-year-ahead point inflation forecasts are displayed, and for Panel B, three-year-ahead point inflation forecasts are displayed. Panels C and D show the interpolated median density mean inflation forecasts of the SCE in percentage points. For Panel C, the one-year-ahead density mean inflation forecasts are displayed, and for Panel B, three-year-ahead density mean inflation forecasts are displayed. The solid blue (dashed orange) lines correspond to inflation forecasts of new (repeat) survey participants of the SCE. I winsorize the top and bottom 5% of point inflation forecasts for each period. Sampling weights are unused in the calculation of median. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to May 2021 with monthly frequency.

ing the repeat survey participants' data, which comprises the vast majority of SCE samples. As a result, the learning effects and sampling efficiency problems could be effectively resolved through the dynamic factor model.

The estimated common factor shows that inflation expectations of U.S. households could be higher than what the headline inflation expectation indicators of the FRBNY suggest. When the estimated common factor is projected onto the inflation expectations of new survey participants, the household inflation expectations of the U.S. were on average about 0.11 to 0.38 percentage points higher than the headline indicators of the FRBNY SCE. The exact numbers could be different depending on which indicator the factor is projected onto.

The dynamic factor model I develop can additionally separate out learning and long-run factors from household inflation expectations of the SCE while estimating the common expectation factor. Using the model, I further analyze two important components of household inflation expectations of the SCE other than the common factor: the learning effect of repeat survey participants and long-run factor specific to three-year-ahead long-horizon forecasts.

The estimated learning factor and long-run factor of inflation expectations suggest that inflation expectations of households are largely influenced by news coverage on inflation or oil (and gas) prices, both of which are readily available. Using data constructed from *The New York Times* (NYT), the estimated learning factor is positively correlated with news coverage of inflation. It suggests that repeat survey participants are more likely to notice and reflect inflationary news into their inflation expectations when compared to new survey participants. This is possibly due to how earlier survey participation experiences raised attention to inflation. The estimated long-run inflation expectation factor is also positively correlated to oil prices. This suggests that when forming their inflation expectations, households could be sensitive to oil and gas prices (Coibion and Gorodnichenko, 2015b).

The paper is structured as follows. Section 2.2 provides brief information about the FRBNY SCE, the main dataset. Section 2.3 provides the technical overview of the dynamic factor model I estimate. Section 2.4 presents the main results showing the estimated factors and discusses which macroeconomic variables are associated with those factors. Section 2.5 concludes.

2.2 Data

The dynamic factor model I use is constructed using 8 inflation expectation indicators available in the Federal Reserve Bank of New York's (FRBNY) Survey of Consumer Expectations (SCE) data. Table 2.1 summarizes those 8 inflation expectation indicators. The FRBNY SCE is an online survey that began in 2013. The SCE is a monthly and nationally-representative with a rotating panel structure, tracking each respondent up to 12 times consecutively. Each month, the SCE has a sample size of approximately 1,300. The number of new participants is about 150. Both short term (one-year-ahead) and long term (three-year-ahead) inflation expectation information are solicited in the survey.

In addition to inflation point forecasts, the FRBNY elicits the respondent's histogram or density forecasts for inflation by asking the respondent to assign probabilities that future inflation will fall into various bins, summing to

Table 2.1: List of Inflation Expectation Indicators Available in the SCE

	Expectation Type	Forecast Horizon	Respondents Type
$\pi_{t,SR}^{P,New}$	Point	1-year-ahead	New
$\pi_{t,SR}^{P,Repeat}$	Point	1-year-ahead	Repeat
$\pi_{t,SR}^{D,New}$	Density Mean	1-year-ahead	New
$\pi_{t,SR}^{D,Repeat}$	Density Mean	1-year-ahead	Repeat
$\pi_{t,LR}^{P,New}$	Point	3-year-ahead	New
$\pi_{t,LR}^{P,Repeat}$	Point	3-year-ahead	Repeat
$\pi_{t,LR}^{D,New}$	Density Mean	3-year-ahead	New
$\pi_{t,LR}^{D,Repeat}$	Density Mean	3-year-ahead	Repeat

Note: The data is from the FRBNY Survey of Consumer Expectations, June 2013 to May 2021. Aggregate interpolated median values are used for each month. Sampling weights are unused in the calculation of median. Repeat survey participants refers to the survey respondents with the total number of survey experiences greater than one. New survey participants means the first-time survey participant who just entered into the survey.

100%. The FRBNY provides estimates of the mean, median, and IQR of each density forecast.[3] These estimates are obtained by fitting parametric (beta) distributions to the density forecasts. See Armantier et al. (2017) for technical procedures and more details. The exact phrasing of survey questions is available in the appendix A.3.1.

[3]For inflation uncertainty, we can use the interquartile range (IQR) estimated from each individual's probabilistic forecast.

Point forecasts generally tend to be higher and more volatile than density mean forecasts partly because of the implicit guidance provided by the bin intervals used in density forecasts. When respondents complete their survey, the density forecast is solicited through probabilistic questions after the respondent provides a point forecast. Those probabilistic questions have upper and lower bounds corresponding to inflation above 12% and deflation below -12%. The bins near zero are narrower than those above 4% or below -4% and symmetrically centered around zero. From these bins, the respondent may infer that most of the probability should be placed around zero, or at least not above 12% and below 12%. Therefore, we winsorize the top and bottom 5% of all point inflation forecasts for each period to ensure the robustness of the estimates.

The FRBNY's headline inflation expectations measure is based on the weighted median of individual density mean forecasts. As an average value is susceptible to outliers, the FRBNY and most central banks generally prefer to keep the median value as their benchmark indicator. Following such conventions, I also use an aggregate (interpolated) median as the main benchmark throughout the paper.

2.3 Model

2.3.1 Dynamic Factor Model Setup

The dynamic factor model that I estimate follows a standard structure of Stock and Watson (1989) and Watson and Engle (1983). To briefly explain, there are a few underlying factors (F_t) that are not visible to econometricians. Those latent factors generate a wide variety of observable time-series data (Y_t) up to certain noises and shocks (u_t).The goal is to estimate those latent factors that are not directly visible only using the observable time-series data. Noises and shocks are assumed to be Gaussian. The dynamics of the data is modeled as a linear process. In a state representation,

$$Y_t = \Lambda F_t + u_t \tag{2.1a}$$

$$F_t = AF_{t-1} + v_t \tag{2.1b}$$

$$u_t \sim N(0, I_u) \ and \ v_t \sim N(0, I_v) \tag{2.1c}$$

$$where \ E\big[u_t v'_{t-k}\big] = 0 \ \forall k \tag{2.1d}$$

Equation 2.1a, the observation equation, models the data generating process of the observable data. $Y_t = \big[Y_{1t}, Y_{2t}, \ldots, Y_{Nt}\big]'$, is the $N \times 1$ vector of stationary variables under analysis. Equation 2.1b is the state equation that models the process of latent factors (states). $F_t = \big[F_{1t}, F_{2t}, \ldots, F_{rt}\big]'$ is the $r \times 1$ vector of unobservable factors, which are assumed to be stationary[4]. $\Lambda = \big[\lambda'_1, \lambda'_2, \ldots, \lambda'_N\big]'$ is the $N \times r$ matrix of factor loadings with $\lambda_i = \big[\lambda'_{i1}, \lambda'_{i2}, \ldots, \lambda'_{ir}\big]'$ measures the effects of factors on each dependent variable. u_t and v_t are multivariate and mutually uncorrelated Gaussian processes corresponding to the dependent variable of each equation. That is, error terms

[4]Non-stationary factors are still feasible with a slight modification of the state process. However, we focus on the case of stationary factors here. Miranda et al. (2021) shows that, when it comes to factor extraction, modeling details of a dynamic factor model have minor impacts on the general results

are uncorrelated over time with zero means and hold constant covariance matrix. I_u and I_v are assumed to be diagonal matrices. Note that even if the idiosyncratic components are correlated each other, the estimates based on the assumption of I_u being a diagonal matrix are still asymptotically valid (Stock and Watson, 2002a). That is, the assumption that I_u being a diagonal matrix is not an identification restriction.[5]

2.3.2 Factor Structure and Identification

The assumption on the structure of Λ is generally required to identify the above dynamic factor model, since a rotation of the data can result in an observationally equivalent model (Bai and Wang, 2012). The most common approach to the identification problem is assuming Λ to be a lower triangular matrix. This assumption is naturally imposed because of our factor structure, which will be explained in detail below.

Our baseline model has three factors in explaining the inflation expectations of households: Common factor, Learning factor, and Long-run factor. Any inflation expectation indicators in Table 2.1 can be expressed as a linear combination of those three factors with factor loadings as weights.

$$\begin{gathered} \pi_{t,i}^{j,k} = \lambda_{1,i}^{j,k} F_t^{Common} + \lambda_{2,i}^{j,k} F_t^{Learning} + \lambda_{3,i}^{j,k} F_t^{Long-run} + u_t \\ where\ i \in \{SR, LR\}\ \&\ j \in \{P, D\}\ \&\ k \in \{New, Repeat\} \end{gathered} \tag{2.2}$$

[5]See Doz et al. (2012) and Stock and Watson (2002a) for a more detailed discussion on this issue.

The common factor (F_t^{Common}) is the shared component for all inflation expectations. The learning factor ($F_t^{Learning}$) captures the difference between the expectations of repeat survey participants and new survey participants. This learning factor reflects the fact that repeat survey participants of the SCE generally achieve lower forecast errors and lower level of inflation expectations because of information acquisition during a repetitive panel survey (Kim and Binder, 2022). Since new survey participants do not have suchh learning effects by definition, $\lambda_{2,i}^{j,New} = 0 \ \forall \ i, j$. The long-run factor ($F_t^{Long-run}$) is attached to all three-year-ahead inflation expectations. That is, $\lambda_{3,SR}^{j,k} = 0 \ \forall \ j, k$.

This factor structure implies that Λ of equation 2.1a is a lower triangular matrix. To illustrate this point, consider only point inflation forecasts of the SCE: $\pi_{t,SR}^{P,New}$, $\pi_{t,LR}^{P,New}$, $\pi_{t,SR}^{P,Repeat}$, $\pi_{t,LR}^{P,Repeat}$. The expected values of these four inflation forecasts can be represented as below.

$$\begin{aligned}
\pi_{t,SR}^{P,New} &= \lambda_{1,SR}^{P,New} F_t^{Common} \\
\pi_{t,LR}^{P,New} &= \lambda_{1,LR}^{P,New} F_t^{Common} + \lambda_{3,LR}^{P,New} F_t^{Long-run} \\
\pi_{t,SR}^{P,Repeat} &= \lambda_{1,SR}^{P,Repeat} F_t^{Common} + \lambda_{2,Sr}^{P,Repeat} F_t^{Learning} \\
\pi_{t,LR}^{P,Repeat} &= \lambda_{1,LR}^{P,Repeat} F_t^{Common} + \lambda_{2,LR}^{R,Repeat} F_t^{Learning} + \lambda_{3,LR}^{P,Repeat} F_t^{Long-run}
\end{aligned} \tag{2.3}$$

As one can see from the structure of factor loadings in Equation 2.3, Λ follows a lower triangular shape by design. Also, notice that the common factor (F_t^{Common}) enters into every expectation with a varying degree because of different factor loadings.

In the estimation step, I use the Maximum Likelihood Estimation (MLE) and Kalman Filter (and Smoothing) to estimate latent factors. While the direct MLE calculation was computationally cumbersome in general, the EM algorithm can be used to efficiently estimate the model parameters through MLE (Shumway and Stoffer, 1982; Dempster, 1977).

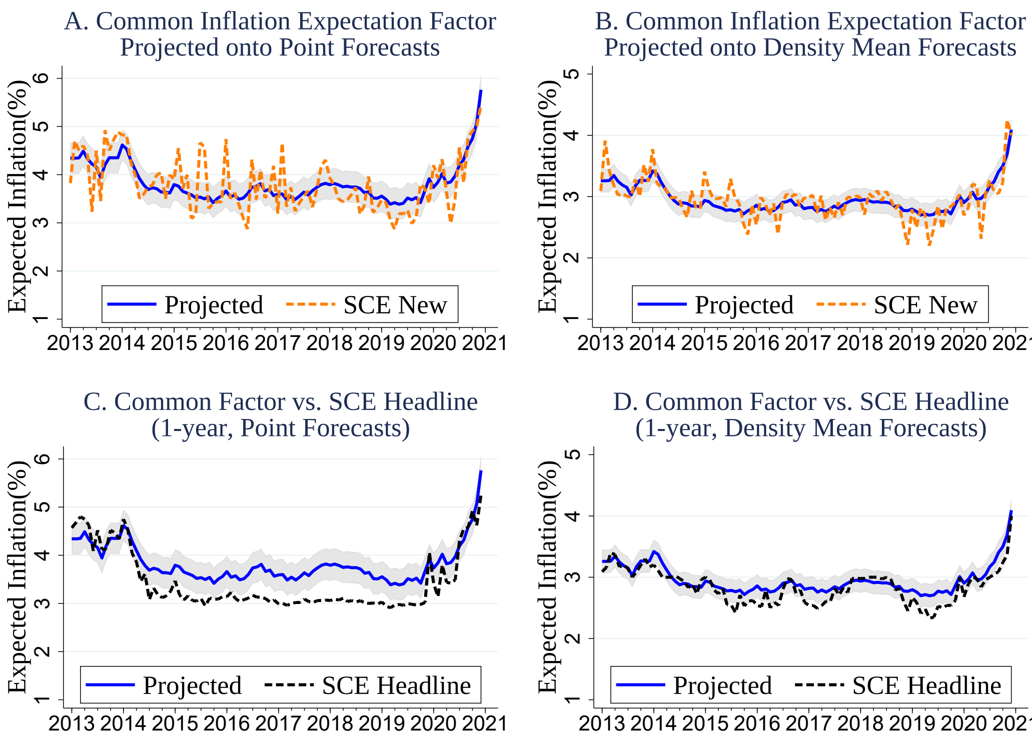
If the scales of each data series is different, a dynamic factor model is generally biased toward the data that have a larger variance. Therefore, I use the standardized data as input in the estimation stage following the standard procedure of estimating a dynamic factor model. All factors will have a mean of zero because of the standardization process. The direction of changes is still informative. To improve the interpretability of the factors, I project the estimated factors onto a data series when we discuss the results.

2.4 Main Results

2.4.1 Estimated Common Factor of Household Inflation Expectations of the SCE

Panel A and B of Figure 2.2 visually show the estimated common factor from the dynamic factor model (Equation 2.2) projected onto median inflation forecasts of new survey participants (point and density mean each). For both cases, the projected common inflation expectation factors track the inflation expectations of new survey participants of the SCE well. The estimated common factor is also significantly less noisy than the original data providing more reliability as economic indicators. In the Appendix Figure B.1, the n-projected

Figure 2.2: Estimated Common Inflation Expectation Factor of the SCE



Note: Panels A and B show the estimated common inflation expectation factors projected onto the median point and density mean inflation forecasts of the new survey participants of the SCE, in percentage points. The solid blue lines of Panel A and B correspond to the common inflation expectation factor projected onto the dashed orange line in each panel. The dashed orange lines correspond to the median point (Panel A) and density mean (Panel B) inflation forecasts of new survey participants of the SCE. Panel C and D compare the projected common inflation expectation factors with the headline inflation expectations of the SCE which are currently being published by the FRBNY. The dashed black lines correspond to the headline median point (Panel C) and density mean (Panel D) inflation forecasts of the FRBNY SCE. The gray area shows a 95% confidence interval for the solid blue line of each panel. I winsorize the top and bottom 5% of point inflation forecasts for each period. Sampling weights are unused in the calculation of median. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to May 2021 with monthly frequency.

raw factor values are plotted that have the same time-series trend.

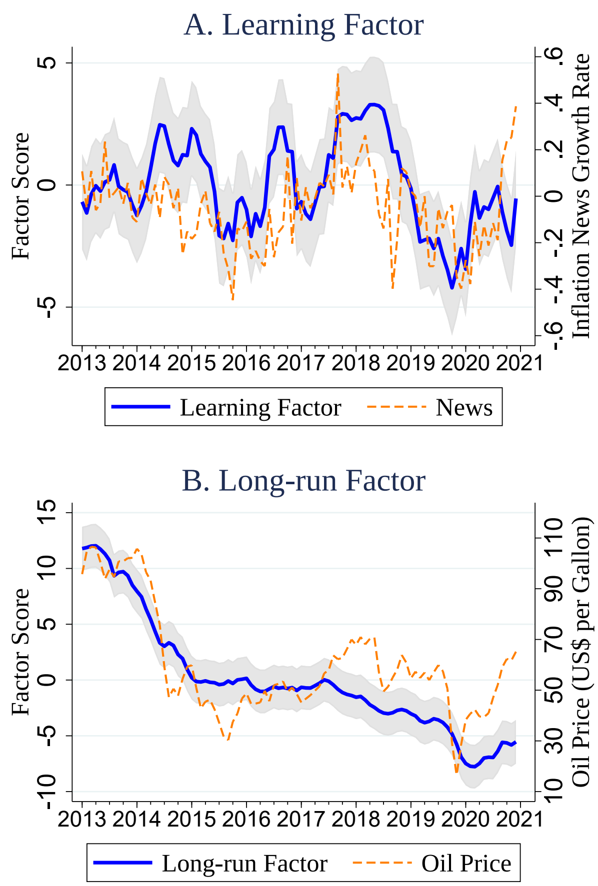
Panel C and D of Figure 2.2 compares the projected common factor with the headline short-run inflation expectations of the SCE. Those headline inflation expectations of the SCE are produced by the FRBNY. Consistent with what Kim and Binder (2022) find, when we factor out the learning effects, the household inflation expectations tend to be generally higher than what the headline estimates suggest. The effect is most salient for the point inflation expectations. On average, the common household inflation expectations of the SCE constructed through our model is 0.38 percentage points higher than those of the headline expectations. For the case of density mean inflation expectations, the estimated inflation expectations are 0.11 percentage points higher than those of the headline expectations. However, it is understandable that density mean inflation expectations display smaller learning effects when we consider the survey questionnaire design of the SCE. Survey participants of the SCE are implicitly guided by the survey questions while answering the probabilistic inflation questions of the SCE.

Note that, after providing a point forecast, the density forecasts of the survey respondent of the SCE are solicited. As it can be directly seen in the actual survey questions presented in Appendix A.3.1, the bins are upper- and lower-bounded at 12% and -12%. In addition, the bins are displayed symmetrically around zero and are narrower than those above 4% or below -4% near zero. While this implicit guidance looks innocent to economists who are familiar with the historical level of inflation rates in general, over half of new survey respondents actually place point forecasts outside of $[-4\%, 4\%]$, before taking the density forecast questions. The learning factor from the implicit information given by the survey design itself can result in smaller observed tenure effects in subsequent rounds (Niu and Harvey, 2021; Kim and Binder, 2022; Halpern-Manners et al., 2017).

While the factor structure of Equation 2.2 is imposed to the data by design in the model, a non-parametric approach such as PCA further confirms the robustness of the common factor estimated. I apply PCA to all inflation expectations of the SCE available (Table 2.1) and estimate the first principal component. The first principal component of inflation expectations of the SCE alone can account for about 64.2% of the total variation in the data. Appendix Figure B.2 shows that the first principal component and estimated common inflation expectation factor are highly correlated to each other. Indeed, the correlation coefficient between those two time series is 0.90. That is, the common factor I estimated is not heavily influenced by the factor structure the model imposed a priori. Rather, it captures the central movements of the data well.

Overall, the estimated common inflation expectation factor successfully captures the inflation expectations of the SCE without suffering from learning effects factor. It consistently follows the dynamics of inflation expectations of new survey participants of the SCE as expected and is significantly less noisy than the raw inflation expectation data of new survey participants.

Figure 2.3: Estimated Learning Factor and Long-run Factor of Inflation Expectations of the SCE



Note: The solid blue lines of Panel A and B show the estimated learning factor and long-run factor of Equation 2.2. The left y-axis of both panel display the raw factor value without a projection (solid blue line; "Learning Factor" and "Long-run Factor"). For Panel A, the annual growth rate of the number of inflation news articles on NYT, measured as $log(news_t) - log(news_{t-12})$, is on the right y-axis (dashed orange line; "News"). For Panel B, the monthly average nominal WTI crude oil price per barrel in US$ is on the right y-axis (dashed orange line; "Oil Price"). The gray area shows a 95% confidence interval for the solid blue line of each panel. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to May 2021 with monthly frequency. For the NYT news coverage data, monthly average is used.

2.4.2 Estimated Learning and Long-run Factors of Household Inflation Expectations of the SCE

The dynamic factor model developed in Equation 2.2 can separate out the learning and long-run factors from household inflation expectations of the SCE while estimating the common expectation factor. Using the model, I further analyze two important components of household inflation expectations of the SCE other than the common factor: the learning effect of repeat survey participants and long-run factor specific to three-year-ahead long-horizon forecasts. I discuss what macroeconomic variables drive the learning component and the long-run component of household inflation expectations of the SCE.

The solid blue line in Panel A of Figure 2.3 shows a time series pattern of the estimated learning factor. The learning factor measures the difference between inflation expectations of new survey participants and repeat participants. When we assume the inflation expectations of new survey participants stay the same, higher learning factor means inflation expectations of repeat survey participants get higher.

I constructed the measure of news coverage of inflation by counting the number of news articles containing the word "inflation" in the *New York Times* each month. The dashed orange line in Panel A of Figure 2.3 shows the annual growth rate of the number of inflation news articles on NYT.[6] While the level of a factor is not informative because of standardization of input data which involves demeaning, Panel A of Figure 2.3 shows that the learning

[6] The annual growth rate is measured in a log difference: $log(news_t) - log(news_{t-12})$

factor (solid blue line) is positively correlated with the news coverage measure of inflation (dashed orange line). This alludes that repeat survey participants are more likely to notice news coverage of inflation and incorporate the news into their inflation expectations than new survey participants do, nudged by their earlier survey participation.

Long-run inflation expectations are particularly of interest for monetary policy authority as one of the main policy objectives is to keep long-run inflation expectations staying around the target inflation rates. Therefore, it would be important to understand what drives inflation expectations of households over a long horizon in particular.

Panel B of Figure 2.3 shows the positive correlation between the long-run inflation expectation factor of households and nominal oil prices. The long-run factor is attached to all three-year-ahead inflation forecasts of the SCE and measures an expectation component specific to three-year-ahead (long-run) inflation expectations compared to one-year-ahead (short-run) inflation expectations. The solid blue line in Panel B of 2.3 shows the estimated long-run factor. The dashed orange line in the same panel shows the monthly average nominal WTI crude oil price per barrel in US dollars.[7]. The positive correlation between the long-run inflation expectation factor and oil prices indicates that long-run inflation expectations of households could be largely

[7]The positive correlation between the long-run factor and oil prices get stronger when we use the real oil prices based on CPI, which may better measure the importance of energy prices in consumers' point of view

influenced by salient prices. This finding is consistent with what previous literature documented about household inflation expectations. Household inflation expectations are generally sensitive to oil and gas prices. Coibion and Gorodnichenko (2015b) find that the increase in inflation expectations of households during the Great Recession can be attributed to the increase in the oil price, since consumers regularly purchase gasoline for their daily life.

2.5 Conclusion

Using a dynamic factor model, I estimated three components of the household inflation expectations of the SCE: Common factor, Learning factor, and Long-run factor. The estimated model successfully captures a common inflation expectation factor shared by all survey participants of the SCE, correcting for the bias due to the learning-through-survey effects of repeat survey participants (Kim and Binder, 2022). The model utilizes all the data available without discarding the repeat survey participants' data that consist of more than 80% of the total SCE samples, thereby improving the efficiency of the resulting indicator.

The estimated common factor is significantly less noisy than the raw data of new survey participants' inflation expectations is. When the estimated common factor is projected onto the inflation expectations of new survey participants, it suggests that the household inflation expectations of the U.S. could be 0.11 to 0.38 percentage points higher on average than what the headline indicators of the FRBNY SCE suggests, depending on the indicator on which the factor is being projected onto.

In addition, the estimated learning factor and long-run factor of inflation expectations suggest that inflation expectations of households are largely influenced by news coverage on inflation or oil (and gas) prices which are readily available information in general. For example, the estimated learning factor is positively correlated with the news coverage of inflation that I constructed using the *New York Times* data. It suggests that repeat survey participants are more likely to notice inflationary news and reflect the news into their inflation expectations than new survey participants do, partly due to raised attention to inflation because of their earlier survey participation experience. Also, the estimated long-run inflation expectation factor is positively correlated to oil prices. This suggests that when forming their inflation expectations, households could be sensitive to oil and gas prices since those are generally salient to consumers (Coibion and Gorodnichenko, 2015b).

Chapter 3

Product Life Cycle Effects on Inflation Inequality

3.1 Introduction[1]

Inflation rates generally refer to the rate of increase in an aggregate price index. Central banks often focus on Consumer Price Index (CPI), which is based on the aggregate consumption bundle of a country. However,the rate of price level increases that individual consumers face can be vastly heterogeneous from the national aggregate average, because taste and other factors shape the consumption bundle of each household.

Take, for instance, the price of a new luxury cosmetic by Dior. The price will be highest right after its launch and initial marketing campaign. As the novelty dissipates, the prices of a new luxury product slowly come down similar to the level of existing products. As a result, an early adopter who purchases a product early on the product life cycle may experience a more

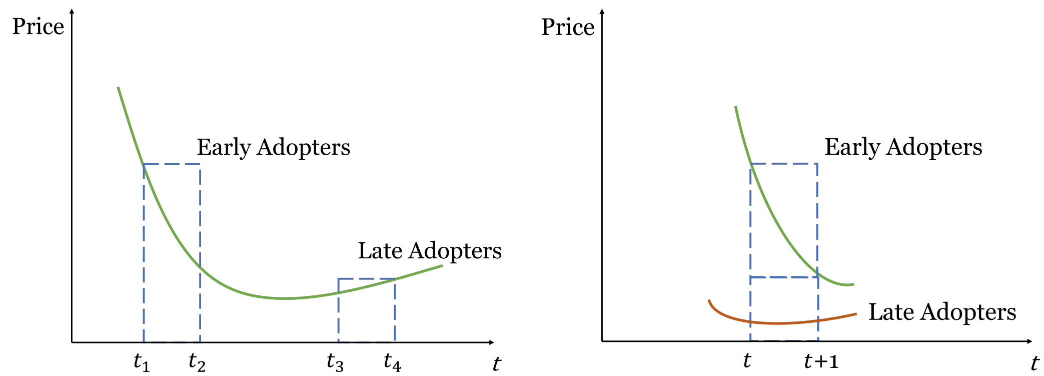
[1]Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

sharp price decrease compared to the initial prices. However, most regular consumers who purchase a new product later will experience price increases in general. The price for the product has stabilized from its initial state, and will begin steadily increasing following general inflation rate trends. While the physical properties of the product stay the same, the price changes that each consumer experiences could be very different depending on how much they value novelty in their consumption bundle.

In this chapter, I quantitatively evaluate how much the product cycle channel can account for the inflation inequality among the U.S. households. Previous research on household-level inflation rates in the U.S. has frequently shown that low-income households have experienced higher inflation rates than high-income households, a finding termed in extensive literature as "inflation inequality" (Hobijn and Lagakos, 2005). While inflation inequality depends on how inflation rates are measured, households belonging to the bottom 20% income quintile have experienced an average about 0.5% points to 0.6% points higher annual inflation rates compared to top 20% income quintile from 2004 to 2018 (Kaplan and Schulhofer-Wohl, 2017; Argente and Lee, 2021; Jaravel, 2019).

Notably, a large portion of inflation inequality is attributable to the difference in specific products *within* the same category of goods instead of the difference in shares of expenditures in broader categories (Kaplan and Schulhofer-Wohl, 2017; Jaravel, 2019). That is, even if two households spend exactly the same amount of money on the same grocery category, those house-

Figure 3.1: Price Changes over Product Life Cycle



hold that prefers premium products are more likely to experience lower inflation rates than a household preferring nonpremium goods.

Several channels have been suggested to account for the inflation inequality. Jaravel (2019) emphasizes the endogenous responses of the supply side, in particular the expansion of the U.S. premium market. This expansion is a result of rising income inequality, which encouraged firms to produce more high-end products curated for wealthy households. Argente and Lee (2021) document that the inflation inequality has been largest during the Great Recession. They show that the substitution toward cheaper products, which is the margin generally more available to high-income households, can explain approximately 40% of the inflation inequality between the top and bottom income quartile during the Great Recession. Most recently, Jaravel (2021) provides a comprehensive literature review about inflation inequality in the U.S.

I suggest another potential channel for inflation inequality: product life cycle. Figure 3.1 conceptually illustrates how the product life cycle effect can generate an inflation rates gap between early adopters and late adopters.

Suppose there is a new high-end cosmetic brand or organic dairy product that consumers repeatedly purchase. Generally, prices for new products are often marked high initially, but slowly begin decreasing after its first launch (Ueda et al., 2019; Bils, 2009).[2] In the presence of non-linear pricing, early adopters on average can experience lower price increase rates (if not decreases) when compared to late adopters. This is because late adopters join the product life cycle when prices bottom out and begin increasing following general inflation rates. That is, rich households who tend to be more early adopters on average can display lower inflation rates when compared to other households.

Using the Nielsen Retail Scanner and Consumer Panel data, I show that the product life cycle channel can account for approximately 21% of the inflation inequality between the top and bottom quintiles of household income in the U.S. Furthermore, I empirically show the non-linear pricing model of a product life cycle that contrasts to what previous studies have found. Argente et al. (2019) documented that prices of retail products fall over product life log-linearly in a relatively short time horizon. Ueda et al. (2019) also find that the product life cycle effect is generally short-lived when using Japanese retail

[2]Potential factors behind such product life cycle effects in pricing could include: more investment in low-cost mass production technology as well as a broader trend that consumers prefer new goods over old products. Regardless of the reason, such non-linear pricing robustly exists even for small consumable goods sectors. See Ueda et al. (2019) for the evidence drawn from the Japanese retail market.

data. However, I find that the product life cycle effect is highly *non-linear* when we consider a long-run time horizon. Such non-linear pricing over a product life cycle, combined together with consumers' preferences for novelty, provides a new channel that can generate inflation inequality in addition to other confirmed sources.

The paper is structured as follows. Section 3.2 provides information about the main data sets. Section 3.3.1 and 3.3.2 empirically estimate the non-linear product life cycle effect in retail prices. Section 3.3.3 conducts a back-of-the-envelope calculation on how much a product age effect can account for inflation inequality between the top and bottom income quintile for U.S. households. Section 3.4 concludes.

3.2 Data

3.2.1 The Nielsen Retail Scanner Data

The Nielsen Retail Scanner Data and Kilts-Nielsen Consumer Panel Data are our main datasets. The Nielsen Retail Measurement Services Data (RMS) is commonly called the Nielsen Retail Scanner Data, as it is directly collected from participating retail stores using barcode scanners. The RMS tracks sales records for more than 100 retail chains and 45,000 stores across U.S. retail markets. I use RMS data from 2006 to 2017.

The sampling error of the RMS tends to be significantly smaller than what can be found in a household panel survey data, because it is directly collected from the stores. The RMS data contains twelve-digit universal product codes (UPC) for each product; weekly prices; sales associated with the product; package volumes; unit quantities sold; and store information. In total, the RMS covers more than 2 trillion dollars in sales and 100 billion data points. Not surprisingly, as documented in Argente and Lee (2021) and Jaravel (2019), the RMS covers a large portion of the U.S. consumption space – about 40% of all expenditures on products in the CPI. Such large coverage of the RMS enables us to track the product age and life cycle over extended periods. While the RMS generally focuses on the consumer packaged goods sector, it does not limit its data set to only food and grocery categories. For example, the RMS also includes sales data on small durable goods like personal health supplies, cosmetics, and cookware. For these reasons, I primarily rely on the RMS for estimating product life cycle and product age.

3.2.2 The Kilts-Nielsen Consumer Panel Data

When focusing on household-level inflation rate, I use the Kilts-Nielsen Consumer Panel (KNCP) data. The KNCP tracks the yearly purchasing habits of about 50,000 nationally-representative households in the U.S. Each panelist records their purchases with provided barcode scanners that are located in their home. The items purchased by the panelists are identified by a UPC barcode. As a result, the KNCP contains the demographic information of participating households matched together with their consumption expenditure information by barcode-level. The data set covers approximately 3 million unique UPC barcodes organized into 1,200 narrowly-defined product modules. As reported in Argente and Lee (2021) and Jaravel (2019), a price index constructed using the Nielsen Consumer Panel data set and the BLS's food-at-home CPI for all urban consumers in general.

3.2.3 Product and Age Definition

Following Argente and Lee (2021) and Kaplan and Schulhofer-Wohl (2017), I use barcodes as my baseline definition of products throughout the paper. A barcode is a Universal Product Code (UPC) that consists of 12 digits and is uniquely assigned to each specific good available in stores. UPCs were created so retail outlets could determine prices and inventory accurately and to improve transactions along the supply chain.[3]

I measure product age as their in-sample age using the RMS and KNCP data. Any product barcode that appears for the first time on either data set will be treated as a ”newborn” product. However, as one can expect, every product would look like a ”newborn” product based on the the first date of data. To properly measure the birthdate of products, I left-censor the product space with a one year margin. That is, I focus on the products that appeared on the RMS and KNCP data for the first time after the first year of data collection.

[3] Argente et al. (2019) and Jaravel (2019) also use the same definition of products for similar reasons.

3.3 Main Results

3.3.1 Model and Identification Issues in Estimating the Product Life Cycle Effects in Prices

I begin by documenting the presence of the (non-linear) product life cycle effect in pricing, using a panel linear regression of the form:

$$\Delta_0 ln(p_{it}) \equiv ln(p_{it}) - ln(p_{io}) = \sum_{s=1}^{L} \beta_s I_s(Age_{it}) + \Gamma X_t + \varepsilon_{it} \tag{3.1}$$

where the dependent variable $\Delta_0 ln(p_{it})$ is the normalized log prices of product i at time t in nominal or real terms by long differencing with its prices of the initial month ($ln(p_{io})$), L is the lifespan of products in months, $I_s(Age_{it})$ is an age indicator variable that if product i is the age of s at time t then it is 1 and otherwise zero, and X_t is a vector of control variables that consists of aggregate time series data: first-order log-difference of industrial production index and oil price, first difference of unemployment rate and effective federal fund rates, and also monthly seasonal dummy variables.

In Equation 3.1, our main interest is the regression coefficients on age dummy variables, $\{\beta_s\}_{s=1}^{L}$, which estimate how much the price has changed on average from the initial price by percentage point.

One potential identification concern would be a survival bias of the regression coefficients since not every product survive for long periods. For example, a product that cannot meet the demand of consumers can exit the market, which will decrease the overall supply of goods. The prices of long-surviving products could be naturally higher than the others in such a case.

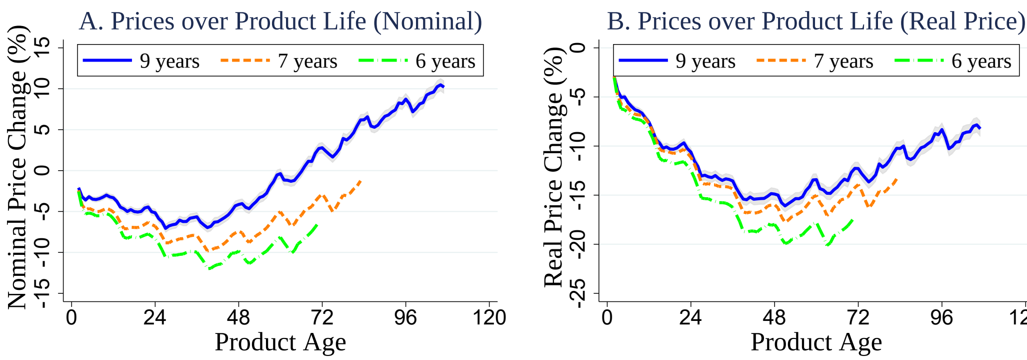
Considering the potential survival bias that can prevent correct measurement of product life cycle effects, I focus on the products that at least survive for the selected lifespan (6, 7 and 9 years). I run separate regressions for each lifespan. The regression coefficients outside of the lifespan will not be used as those coefficients may contain the survival bias. Our sampling and regression coefficients choice prevents the confounding survival bias – at least for the selected lifespan – since panel attrition does not occur during the selected lifespan.[4]. Note also that long differencing of the dependent variable in Equation 3.1 removes possible individual fixed effects for each product, estimating the average within-product effects only and controlling for the unobserved time-constant characteristics.

3.3.2 Estimated Product Life Cycle Effects in Prices

Panel A and B of Figure 3.2 shows the estimated product life cycle effects in prices using Equation 3.1. Panel A shows the estimates based on the nominal prices; Panel B shows estimates based on real prices using CPI. While price changes over the product life cycle is seemingly linear for the first 4 years as documented in Argente et al. (2019), it is clearly non-linear overall over the long run. In Panel A, prices of new products steadily decrease on average in the first 4 years and get approximately 5% to 10% cheaper compared to their initial prices. Then, prices bottom up and begin increasing.

[4]This sampling choice is similar to Halpern-Manners et al. (2017) and Kim and Binder (2022) that focus on "non-attriters" to avoid sample selection effects that can occur due to panel attrition.

Figure 3.2: Estimated Product Life Cycle Effects in Prices



Note: Panels A and B show the change in prices of of new products over product life compared to their initial price levels in percentage points, estimated from Equation 3.1. The solid blue (dashed orange and long-dash dot lime) lines correspond to prices of products that survive at least 9 years (7 years and 6 years) after its first introduction in the U.S. retail market. The gray area shows a 95% confidence interval for the solid blue line with Huber-White standard errors. Product age measured in months is shown on the x-axis. Data is drawn from the RMS and KNPC from January 2006 to December 2017.

A similar but stronger non-linear pattern can be observed in Panel B as well. Note that the overall inflation rates have been mostly positive during the sample periods except for the Great Recession. The real prices of new products fall more sharply than nominal prices do. Nevertheless, real prices eventually increases if the new products survive more than 4-5 years in the end. Additionally, both in Panel A and B, the magnitude of drops in prices tend to be larger for short-lived products.

The degree of the initial price drop is quite large particularly when we consider that price levels generally tend to increase over time following aggregate inflation rates. To illustrate, if 5% of expenditure of a household gets exposed to 10% of deflation rates, it can lower inflation rate of the household by 0.5%. Note that the aggregate CPI inflation rate was about 2.6% in 2004.

As a robustness check, in Appendix Figure C.1 conducts a similar analysis but with price increase rates, not levels, with a dependent variable after aggregating prices for each quarterly cohort. Also, instead of using aggregate time-series variables as control variables, time and module fixed effects are employed. While it has a quite different empirical setup than Equation 3.1, the product life cycle patterns of prices I find in Figure 3.2 remain similar. In general, new products display a large degree of deflation in the first four years since its initial introduction to the market.

3.3.3 Back-of-the-Envelope Calculation of the Effects of Product Life Cycle on Household Inflation Rates

I use three superlative indices to measure the inflation rates of the individual households: Tornqvist, Fisher, and Walsh. Formal definitions for each of them are presented in Equation 3.2 below. While conventional measures like Laspeyres and Paasche indices are more popular, I use the Tornqvist price index as a main benchmark. Theoretically speaking, the Tornqvist index provides a non-parametric 2nd-order approximation to inflation rates under smooth utility functions (Jaravel, 2019).

$$\text{(3.2a)} \qquad ln\left(\pi_t^{Tornqvist}\right) \equiv \frac{1}{2}\sum_{i=1}^{n}\left(\frac{p_{i,t-1}q_{i,t-1}}{\sum_j p_{j,t-1}q_{j,t-1}} + \frac{p_{i,t}q_{i,t}}{\sum_j p_{j,t}q_{j,t}}\right) ln\left(\frac{p_{i,t}}{p_{i,t-1}}\right)$$

$$\text{(3.2b)} \qquad \pi_t^{Fisher} \equiv \sqrt{\pi_t^{Laspeyres} \times \pi_t^{Paasche}}$$

$$\text{(3.2c)} \qquad \pi_t^{Laspeyres} \equiv \frac{\sum_{i=1}^{n}\left(p_{i,t} \times q_{i,t-1}\right)}{\sum_{i=1}^{n}\left(p_{i,t-1} \times q_{i,t-1}\right)}$$

$$\text{(3.2d)} \qquad \pi_t^{Paasche} \equiv \frac{\sum_{i=1}^{n}\left(p_{i,t} \times q_{i,t}\right)}{\sum_{i=1}^{n}\left(p_{i,t-1} \times q_{i,t}\right)}$$

$$\text{(3.2e)} \qquad \pi_t^{Walsh} \equiv \frac{\sum_{i=1}^{n}\left(p_{i,t} \times \sqrt{q_{i,t} \times q_{i,t-1}}\right)}{\sum_{i=1}^{n}\left(p_{i,t-1} \times \sqrt{q_{i,t} \times q_{i,t-1}}\right)}$$

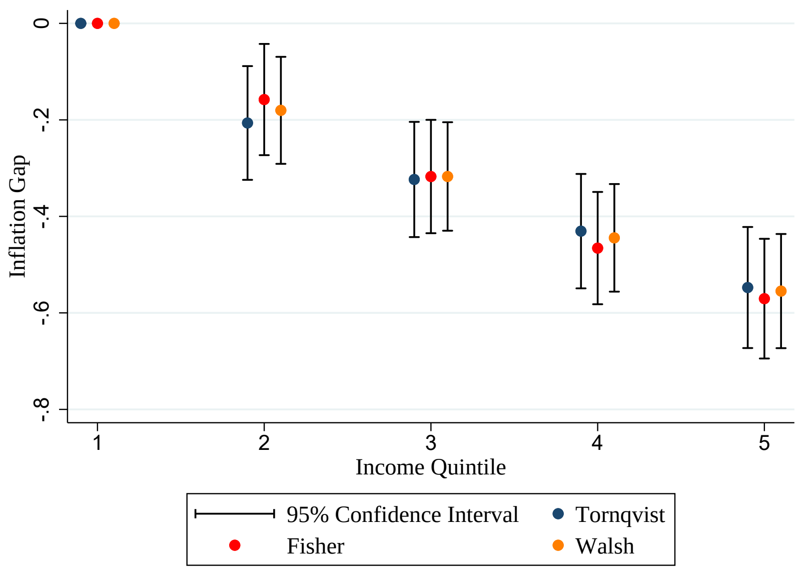
where i denotes each product in the consumption bundle of a household, and n denotes the total number of products in the consumption bundle.[5]

Figure 3.3 shows the estimated average household inflation rate across income quintiles in the U.S. from 2004 to 2017 compared to the average household inflation rates of the lowest income quintile (1st quintile) in percentage points. Consistent with Jaravel (2019), Argente and Lee (2021), and Kaplan and Schulhofer-Wohl (2017), approximately 0.5 to 0.6 percentage points of inflation gap between the highest income quintile and lowest income quintile are found. The highest income quintile (5th quintile) displays lower inflation rates than the others on average.

To evaluate the effects of the product life cycle on household inflation rates on average, I use a simple linear regression with the product age of the

[5]For simplicity of notations, the index for each household is dropped in the formula without loss of generality.

Figure 3.3: Average U.S. Household Inflation Rates across Income Quintiles



Note: The figure shows average household inflation rate across income quintiles in the U.S. compared to the average household inflation rates of the lowest income quintile (1st quintile), in percentage points. Sampling weights provided by Nielsen are used. Data is from the KNCP, from January 2004 to December 2017.

consumption basket of each household. The product age of households' consumption basket is measured as a weighted average using the product barcodes. As a result, the product age could be identified and their expenditure shared as weights.

$$\pi_{it} = f(S_{it}) + \delta_t + \alpha_c + X_{it} + \varepsilon_{it} \tag{3.3}$$

where π_{itc} is annual household-level inflation rates using the Tornqvist price index, δ_t is quarterly time fixed effect, α_c is a county fixed effect, and

X_{it} is a vector of demographic controls including the gender of household head, age of the household head, occupational information, and the number of household members. $f(S_{it})$ is a potentially non-linear function of S_{it} which measures the product age of the household i's consumption basket at time t. I set $f(S_{it})$ which is a potentially non-linear function of the product age of the households as a standard quadratic function. Following Kaplan and Schulhofer-Wohl (2017), the data is transformed in quarterly basis.

By plugging the average product age of the consumption baskets of the top and bottom income quintiles into the estimated function, $f(S_{it})$, the effects of product age on the household-level inflation rates could be obtained. Approximately 21.3% of the inflation inequality (0.126 percentage points out of 0.59 percentage points of the inflation gap) in 2004 to 2017 could be accounted by the difference in product age in the consumption baskets between the top 20% income group and bottom 20% income group. The back-of-the-envelope calculation suggests that consisting the consumption basket older in terms of the product age can increase the household inflation rates to some extent. For more robustness check, I used a cubic spline as well, but the results remain similar both quantitatively and qualitatively.

3.4 Conclusion

Using the RMS and KNCP data sets, I find the presence of substantial non-linearity in prices over product life cycle in the U.S. retail sector. Such a non-linear pattern of prices is in contrast to Argente et al. (2019) that find the log prices of retail products fall over product life linearly.

The presence of non-linear pricing over a product life cycle that is combined together with consumers' preferences over newness provides a novel channel that can generate inflation inequality on top of the other sources. I estimated the effects of the product life cycle on household-level inflation rates using a linear regression. A simple back-of-the-envelope calculation based on the estimated regression coefficients suggest that approximately 21% of inflation inequality between the top and bottom income quintile can be explained by the product life cycle effect.

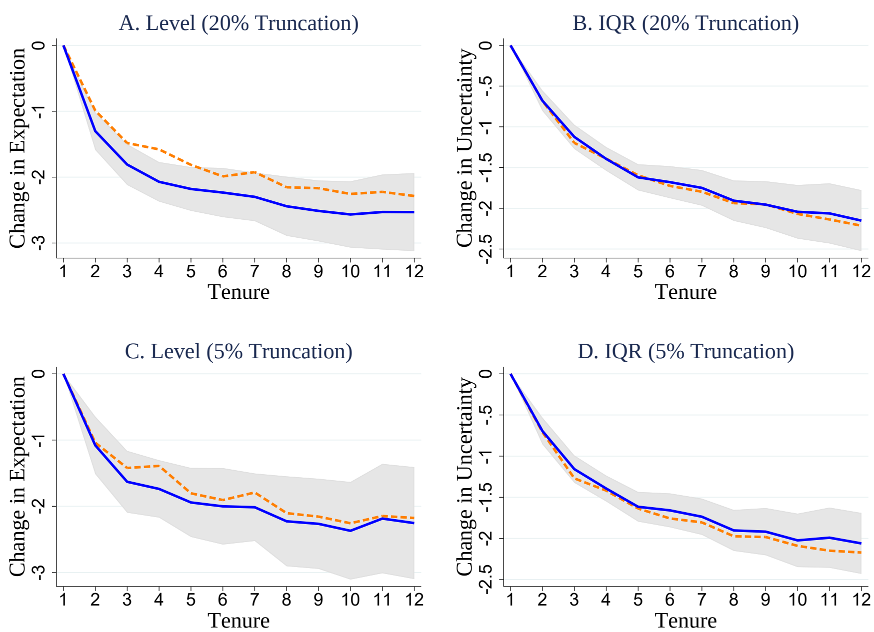
Appendices

Appendix A

Appendix of Chapter1

A.1 Figures

Figure A.1: Average Survey Effects on Inflation Expectations in the SCE by Different Thresholds



Note: Panels A and B reproduce the results of Figure 1 for higher thresholds (trimming top and bottom 10%). Panels C and D reproduce the results of Figure 1 for lower thresholds (trimming top and bottom 2.5%). The y-axis shows the change in responses of survey participants compared to their initial responses, which is estimated from the regression (1). The y-axis is measured in percentage points. For Panel A and C, the dependent variable of the regression is the point inflation rate forecast, and for Panel B and D, the dependent variable is the IQR of consumers' inflation expectations. The solid blue (dashed orange) lines correspond to one-year (three-year) ahead inflation forecasts. The gray area shows a 95% confidence interval for the solid blue line with Driscoll-Kraay standard errors of lag one. Tenure is shown on the x-axis and corresponds to the total number of survey experiences of each respondent (including the current survey wave). Data is from the FRBNY Survey of Consumer Expectation, June 2013 to October 2020.

A.2 Tables

Table A.1: Average Survey Effects on Various Expectations in the SCE

Dependents:	(1) $\pi^{e,point}_{t,t+12}$	(2) $\pi^{e,point}_{t+24,t+36}$	(3) $\pi^{e,density}_{t,t+12}$	(4) $\pi^{e,density}_{t+24,t+36}$	(5) $\pi^{e,IQR}_{t,t+12}$	(6) $\pi^{e,IQR}_{t+24,t+36}$	(7) $\pi^{e,point}_{t,t+12;2015}$
Tenure 2	-1.24	-1.01	-0.50	-0.27	-0.70	-0.71	-1.76
	(0.18)	(0.17)	(0.09)	(0.12)	(0.07)	(0.06)	(0.39)
Tenure 3	-1.78	-1.50	-0.60	-0.46	-1.15	-1.23	-2.37
	(0.20)	(0.21)	(0.09)	(0.11)	(0.08)	(0.06)	(0.21)
Tenure 4	-2.00	-1.56	-0.54	-0.38	-1.38	-1.39	-2.67
	(0.19)	(0.18)	(0.10)	(0.10)	(0.07)	(0.06)	(0.23)
Tenure 5	-2.16	-1.86	-0.61	-0.40	-1.58	-1.58	-2.81
	(0.21)	(0.22)	(0.10)	(0.11)	(0.09)	(0.07)	(0.41)
Tenure 6	-2.24	-2.05	-0.51	-0.34	-1.64	-1.69	-3.37
	(0.24)	(0.23)	(0.13)	(0.14)	(0.10)	(0.08)	(0.33)
Tenure 7	-2.29	-1.97	-0.48	-0.31	-1.71	-1.75	-3.46
	(0.21)	(0.25)	(0.12)	(0.13)	(0.11)	(0.09)	(0.32)
Tenure 8	-2.46	-2.26	-0.57	-0.37	-1.86	-1.91	-3.40
	(0.27)	(0.28)	(0.14)	(0.15)	(0.13)	(0.10)	(0.56)
Tenure 9	-2.49	-2.30	-0.56	-0.37	-1.89	-1.92	-3.64
	(0.28)	(0.31)	(0.16)	(0.16)	(0.14)	(0.12)	(0.55)
Tenure 10	-2.60	-2.35	-0.71	-0.48	-2.00	-2.04	-3.80
	(0.30)	(0.34)	(0.17)	(0.17)	(0.16)	(0.13)	(0.66)
Tenure 11	-2.48	-2.32	-0.58	-0.39	-1.98	-2.10	-3.72
	(0.34)	(0.37)	(0.19)	(0.18)	(0.18)	(0.14)	(0.59)
Tenure 12	-2.55	-2.38	-0.58	-0.40	-2.04	-2.14	-3.57
	(0.34)	(0.42)	(0.20)	(0.21)	(0.19)	(0.15)	(0.69)
Observations	55879	55924	55070	55125	55070	55125	8120
R^2	0.01	0.01	0.00	0.00	0.02	0.02	0.02

Note: Driscoll-Kraay standard errors of lag one are in parentheses. Dependent variables of regressions are represented under the corresponding column numbers. For example, for column (1), the dependent variable, $\pi^{e,point}_{t,t+12}$, is one-year-ahead point inflation forecast. $\pi^{e,point}_{t+24,t+36}$ is three-year-ahead point inflation forecast. $\pi^{e,density}_{t,t+12}$, is one-year-ahead density mean inflation forecast. $\pi^{e,density}_{t+24,t+36}$ is three-year-ahead density mean inflation forecast. $\pi^{e,IQR}_{t,t+12}$ is IQR of one-year-ahead point inflation forecast which is estimated at individual-level using probabilistic forecasts of each respondent. $\pi^{e,IQR}_{t+24,t+36}$ is IQR of three-year-ahead point inflation forecast which is estimated at individual-level using probabilistic forecasts of each respondent. All units are in percentage points. We run a linear panel regression with individual and quarterly fixed effects, $y_{its} = \sum_{s=2}^{12} \beta_s \tau_s + \alpha_i + \gamma_t + \varepsilon_{it}$, where τ_s is a tenure dummy variable for s number of total survey experience. We restrict samples to consist of respondents who eventually participate in the survey for twelve waves (non-attriters). We winsorize the top and bottom 5% of each dependent variable for each tenure group and period. Data is from the FRBNY Survey of Consumer Expectations, June 2013 to October 2020.In column (7), sample is restricted to 2015.

Table A.2: Average Survey Effects on Updating of Expectations and Absolute Forecast Errors in the SCE

	(1)	(2)	(3)	(4)
Dependents:	$Update(\pi^e_{t,t+12})$	$Update(\pi^e_{t+24,t+36})$	$\lvert\pi^e_{t,t+12} - \pi_{t,t+12}\rvert$	$\lvert\pi^e_{t+24,t+36} - \pi_{t+24,t+36}\rvert$
Tenure 2			-2.02	-1.46
			(0.16)	(0.14)
Tenure 3	-5.06	-3.11	-2.92	-2.20
	(0.98)	(0.82)	(0.16)	(0.16)
Tenure 4	-7.31	-6.79	-3.34	-2.51
	(1.15)	(0.90)	(0.18)	(0.18)
Tenure 5	-9.20	-8.75	-3.56	-2.86
	(1.31)	(1.12)	(0.21)	(0.19)
Tenure 6	-12.50	-9.97	-3.69	-3.12
	(1.48)	(1.27)	(0.23)	(0.22)
Tenure 7	-13.93	-11.82	-3.81	-3.16
	(1.80)	(1.29)	(0.25)	(0.26)
Tenure 8	-14.65	-11.85	-4.10	-3.43
	(1.93)	(1.47)	(0.31)	(0.28)
Tenure 9	-16.68	-12.58	-4.17	-3.41
	(2.43)	(1.74)	(0.33)	(0.31)
Tenure 10	-16.59	-11.50	-4.24	-3.59
	(2.52)	(1.98)	(0.35)	(0.34)
Tenure 11	-17.65	-11.72	-4.18	-3.56
	(2.85)	(2.34)	(0.41)	(0.39)
Tenure 12	-18.08	-12.74	-4.29	-3.60
	(3.31)	(2.42)	(0.44)	(0.43)
Observations	51162	51210	55812	41660
R^2	0.00	0.00	0.03	0.02

Note: Driscoll-Kraay standard errors of lag one are in parentheses. Dependent variables of regressions are represented under the corresponding column numbers. For example, for column (1), the dependent variable, $Update(\pi^e_{t,t+12})$, is an indicator variable for an update of one-year-ahead point inflation forecast. For example, if $\pi^e_{t,t+12} \neq \pi^e_{t-1,t+11}$ then $Update(\pi^e_{t,t+12}) = 100$ (percentage points unit) but otherwise $Update(\pi^e_{t,t+12})$ is zero. A similar rule applies to $Update(\pi^e_{t+24,t+36})$. $Update(\pi^e_{t+24,t+36})$ is an indicator variable for an update of three-year-ahead point inflation forecast. $|\pi^e_{t,t+12} - \pi_{t,t+12}|$ measures absolute forecast error of one-year-ahead point inflation forecast. $\pi_{t,t+12}$ corresponds to realized seasonally-adjusted CPI inflation rates from period t to period $t+12$ (all urban consumer items). A similar rule applies to three-year-ahead point inflation forecast. $|\pi^e_{t+24,t+36} - \pi_{t+24,t+36}|$ measures absolute forecast error of three-year-ahead point inflation forecast. All units are in percentage points. We run a linear panel regression with individual and quarterly fixed effects, $y_{its} = \sum_{s=2}^{12} \beta_s \tau_s + \alpha_i + \gamma_t + \varepsilon_{it}$, where τ_s is a tenure dummy variable. We restrict samples to consist of respondents who eventually participate in the survey for twelve waves (non-attriters). For column (3) and (4), we winsorize the top and bottom 5% of point inflation forecasts for each tenure group and period. Data is from the FRBNY Survey of Consumer Expectations and Federal Reserve Economic Data, June 2013 to October 2020.

Table A.3: Tenure Effects on Other Expectations Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dependents:	Unemp	Gas	Food	Med	College	Rent	Gold	Home^{iqr}	Inc	Earn
Tenure 2	-3.28								-0.55	-0.35
	(0.39)								(0.20)	(0.13)
Tenure 3	-5.72	-0.34	-0.25	-0.56	-0.51	-0.50	-0.30	-0.41	-0.48	-0.36
	(0.54)	(0.19)	(0.07)	(0.20)	(0.14)	(0.10)	(0.18)	(0.07)	(0.21)	(0.15)
Tenure 4	-7.66	-0.68	-0.54	-1.23	-1.09	-0.92	-0.38	-0.63	-0.77	-0.59
	(0.72)	(0.29)	(0.09)	(0.21)	(0.16)	(0.09)	(0.19)	(0.06)	(0.22)	(0.16)
Tenure 5	-9.12	-1.10	-0.76	-1.74	-1.77	-1.26	-0.49	-0.84	-1.05	-0.79
	(0.89)	(0.34)	(0.11)	(0.24)	(0.18)	(0.12)	(0.18)	(0.06)	(0.28)	(0.19)
Tenure 6	-9.65	-1.21	-0.78	-1.90	-1.80	-1.54	-0.49	-0.91	-1.22	-0.81
	(1.03)	(0.44)	(0.12)	(0.28)	(0.20)	(0.13)	(0.23)	(0.07)	(0.30)	(0.20)
Tenure 7	-9.93	-1.40	-0.82	-2.06	-2.17	-1.69	-0.60	-1.01	-1.29	-0.94
	(1.11)	(0.48)	(0.14)	(0.31)	(0.21)	(0.15)	(0.24)	(0.08)	(0.32)	(0.22)
Tenure 8	-11.18	-1.42	-0.91	-2.25	-2.45	-1.84	-0.70	-1.10	-1.31	-0.92
	(1.44)	(0.63)	(0.16)	(0.31)	(0.24)	(0.18)	(0.26)	(0.09)	(0.37)	(0.23)
Tenure 9	-11.46	-1.66	-1.00	-2.55	-2.56	-2.03	-0.85	-1.13	-1.29	-0.91
	(1.64)	(0.71)	(0.19)	(0.37)	(0.25)	(0.19)	(0.30)	(0.10)	(0.41)	(0.26)
Tenure 10	-12.10	-1.99	-1.11	-2.79	-2.78	-2.21	-0.88	-1.19	-1.24	-0.87
	(1.84)	(0.80)	(0.21)	(0.43)	(0.31)	(0.22)	(0.33)	(0.10)	(0.47)	(0.27)
Tenure 11	-12.32	-1.96	-1.16	-2.68	-2.64	-2.23	-0.91	-1.27	-1.23	-0.89
	(2.03)	(0.89)	(0.24)	(0.46)	(0.32)	(0.26)	(0.37)	(0.11)	(0.51)	(0.32)
Tenure 12	-12.81	-2.10	-1.12	-2.73	-2.79	-2.37	-0.86	-1.32	-1.39	-0.78
	(2.37)	(1.01)	(0.25)	(0.49)	(0.35)	(0.27)	(0.43)	(0.12)	(0.57)	(0.35)
Observations	45522	51353	51403	51398	51366	51395	51330	50926	55968	34883
R^2	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00

Note: Driscoll-Kraay standard errors of lag one are in parentheses. For column (1), the dependent variable is the percent chance that unemployment will be higher in 12 months, and sample is limited to months when realized unemployment is lower in 12 months (so smaller responses are more accurate). For columns (2) to (7), the dependent variable is the expected percent change in prices in the next 12 months for the indicated category. In column (8), the dependent variable is the interquartile range of the respondent's density forecast for national home prices. In columns (9) and (10), the dependent variable is the point forecast for household income or personal earnings growth in the next 12 months. The dependent variables in (2) through (8) are only asked of respondents with tenure 2 or greater. We run a linear panel regression with individual and quarterly fixed effects, $y_{its} = \sum_{s=2}^{12} \beta_s \tau_s + \alpha_i + \gamma_t + \varepsilon_{it}$, where τ_s is a tenure dummy variable. We restrict samples to consist of respondents who eventually participate in the survey for twelve waves (non-attriters). We winsorize the top and bottom 5% of each dependent variable for each tenure group and period. Data is from the FRBNY Survey of Consumer Expectations, June 2013 to October 2020.

Table A.4: Panel Regression Estimation of Responses to Gas Prices by Survey Tenure

	(1)	(2)	(3)	(4)	(5)	(6)
Tenure 1×$log(Gas)$	2.75	1.68	4.85	1.33	1.25	1.66
	(0.57)	(0.49)	(1.13)	(0.42)	(0.18)	(0.17)
Tenure 2×$log(Gas)$	2.33	1.26	3.20	0.84	0.78	1.27
	(0.43)	(0.38)	(0.65)	(0.43)	(0.31)	(0.37)
Tenure 3×$log(Gas)$	2.34	1.03	2.95	0.75	0.46	1.30
	(0.53)	(0.44)	(1.00)	(0.42)	(0.24)	(0.18)
Tenure 4×$log(Gas)$	2.21	1.07	2.50	0.82	0.71	1.18
	(0.58)	(0.50)	(1.07)	(0.41)	(0.30)	(0.22)
Tenure 5×$log(Gas)$	2.01	0.89	2.29	0.77	0.47	1.00
	(0.52)	(0.48)	(0.94)	(0.41)	(0.31)	(0.38)
Tenure 6×$log(Gas)$	2.38	1.11	2.18	0.90	0.55	1.40
	(0.52)	(0.53)	(1.14)	(0.42)	(0.27)	(0.30)
Tenure 7×$log(Gas)$	2.34	0.99	2.41	0.94	0.40	1.39
	(0.55)	(0.56)	(1.22)	(0.42)	(0.27)	(0.31)
Tenure 8×$log(Gas)$	2.07	0.63	1.73	0.86	0.15	1.14
	(0.60)	(0.58)	(1.12)	(0.41)	(0.38)	(0.35)
Tenure 9×$log(Gas)$	1.93	0.42	1.52	0.87	0.10	1.03
	(0.70)	(0.69)	(1.36)	(0.41)	(0.52)	(0.50)
Tenure 10×$log(Gas)$	1.88	0.30	1.70	0.75	0.02	1.02
	(0.68)	(0.66)	(1.40)	(0.41)	(0.43)	(0.43)
Tenure 11×$log(Gas)$	1.95	0.29	1.44	0.90	-0.07	1.12
	(0.77)	(0.66)	(1.46)	(0.42)	(0.45)	(0.43)
Tenure 12×$log(Gas)$	1.84	0.32	1.40	0.91	-0.17	1.04
	(0.80)	(0.72)	(1.55)	(0.42)	(0.58)	(0.58)
Expectation Type	Mean	Mean	Point	Mean	Mean	Mean
10% Winsorization	N	Y	Y	N	N	N
Full Survey Participation	Y	Y	Y	Y	N	Y
Individual FE	Y	Y	Y	Y	Y	Y
Quarterly Time FE	Y	Y	Y	Y	Y	N
Sample Period	14m7-15m2	14m7-15m2	14m7-15m2	13m6-19m12	14m7-15m2	14m7-15m2
Observations	5360	5360	5454	54694	9859	5360
F Statistic	33.8	57.0	3.5	5.9	60.9	36.5

Note: Driscoll-Kraay standard errors of lag one are in parentheses. Tenure corresponds to the the total number of of survey experiences of each respondent (including the current survey wave). The independent variables are interaction terms between monthly average U.S. Regular Conventional Gas price per Gallon in $US and tenure dummy variables (τ_s). The dependent variable is the one-year-ahead density mean inflation expectation (in percentage points) estimated by the NY Fed, except for model (3) which uses point inflation expectations. For model (2) and (3), we winsorize the top and bottom 5% of dependent variable for each tenure group and period. Except for model (5), we restrict samples to consist of respondents who eventually participate in the survey for twelve waves (non-attriters). Data is from the FRBNY Survey of Consumer Expectations and Federal Reserve Economic Data.

Table A.5: Panel Regression Estimation of the EIS and Excess Sensitivity by Survey Tenure

	(1)		(2)		(3)	
	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\gamma}$
Pooled	0.70	0.24				
	(0.02)	(0.01)				
Tenure1			0.62	0.22	0.36	0.27
			(0.03)	(0.02)	(0.07)	(0.02)
Tenure2			0.66	0.24	0.38	0.29
			(0.03)	(0.02)	(0.06)	(0.02)
Tenure3			0.68	0.24	0.47	0.28
			(0.03)	(0.02)	(0.05)	(0.02)
Tenure4			0.70	0.25	0.48	0.30
			(0.03)	(0.02)	(0.05)	(0.02)
Tenure5			0.70	0.24	0.49	0.28
			(0.03)	(0.0175)	(0.05)	(0.02)
Tenure6			0.75	0.21	0.53	0.26
			(0.03)	(0.02)	(0.06)	(0.02)
Tenure7			0.74	0.24	0.55	0.29
			(0.03)	(0.02)	(0.05)	(0.02)
Tenure8			0.74	0.24	0.54	0.28
			(0.03)	(0.02)	(0.05)	(0.02)
Tenure9			0.76	0.24	0.62	0.28
			(0.03)	(0.02)	(0.05)	(0.03)
Tenure10			0.78	0.22	0.61	0.26
			(0.0290)	(0.02)	(0.06)	(0.02)
Tenure11			0.70	0.26	0.53	0.30
			(0.03)	(0.02)	(0.06)	(0.02)
Tenure12			0.71	0.24	0.55	0.28
			(0.03)	(0.02)	(0.05)	(0.03)
Regression Type	OLS		OLS		IV	
Observations	54970		54970		54850	

Note: Driscoll-Kraay standard errors of lag one are in parentheses. We run a linear panel regression of Crump et al. (2015), allowing regression coefficients to vary by survey experience of respondents: $ExpCG^i_{t,t+12} = -\sum_{s=1}^{12} \tau_s \sigma_s ExpInf^i_{t,t+12} + \sum_{s=1}^{12} \tau_s \gamma_s ExpIG^i_{t,t+12} + \alpha_i + \beta_t + \varepsilon_{i,t}$. The dependent variable is expected real consumption growth over the next twelve months of households, $ExpCG^i_{t,t+12}$. Independent variables are density-implied mean inflation rates, $ExpInf^i_{t,t+12}$, and expected real household income growth, $ExpIG^i_{t,t+12}$. α_i and β_t are individual and quarterly time fixed effects. τ_s is a dummy variable for respondents whose tenure of s. Tenure corresponds to the the total number of of survey experiences of each respondent (including the current survey wave). For the case of IV, the point inflation expectation is used as an instrument of density-implied mean inflation expectation. All units of variables are in percentage points. We winsorize the top and bottom 5% of each dependent variable for each tenure group and period. We restrict samples to consist of respondents who eventually participate in the survey for twelve waves (non-attriters). Data is from the FRBNY Survey of Consumer Expectations, June 2013 to October 2020.

Table A.6: Summary Statistics of SCE One-year-ahead Inflation Point Forecasts

Tenure	Mean	Std.	N
1	7.0	12.5	15,050
2	5.5	8.7	12,585
3	5.1	7.4	11,702
4	4.8	6.4	11,092
5	4.7	6.0	10,509
6	4.5	5.6	10,035
7	4.5	5.5	9,588
8	4.3	5.0	8,858
9	4.3	5.0	8,177
10	4.2	5.0	7,411
11	4.2	5.2	6,385
12	4.1	5.0	4,717
Total	5.0	7.4	116,109

Note: Tenure refers to the total number of survey experiences including the current survey experience. The one-year-ahead point inflation expectations are used. Std. denotes the standard deviation. By each tenure group and period, the top and bottom 5% of observations are winsorized. Sampling weights are unused, and the maximum tenure is not restricted in the calculation of the summary statistics. The data is from the FRBNY Survey of Consumer Expectation, June 2013 to October 2020.

Table A.7: Summary Statistics of One-year-ahead Density-implied Mean Inflation Expectations of the SCE

Tenure	Mean	Std.	N
1	5.5	4.8	14,258
2	4.5	4.3	12,231
3	4.0	3.9	11,512
4	3.7	3.6	10,927
5	3.6	3.4	10,379
6	3.5	3.4	9,943
7	3.4	3.3	9,512
8	3.3	3.2	8,789
9	3.3	3.3	8,113
10	3.2	3.2	7,342
11	3.2	3.2	6,338
12	3.1	3.1	4,681
Total	3.8	3.8	114,025

Note: Tenure refers to the total number of survey experiences including the current survey experience. The one-year-ahead density-implied mean inflation expectations are used. Std. denotes the standard deviation. By each tenure group and period, the top and bottom 5% of observations are winsorized. Sampling weights are unused, and the maximum tenure is not restricted in the calculation of the summary statistics. The data is from the FRBNY Survey of Consumer Expectations, June 2013 to October 2020.

Table A.8: Summary Statistics of One-year-ahead Inflation Expectations of the MSC

Tenure	Mean	Std.	N
1	4.7	4.9	170,066
2	3.5	3.5	101,705
Total	4.3	4.5	271,771

Note: Tenure refers to the total number of survey experiences, including the current experience. The one-year-ahead point inflation expectations are used. Std. denotes the standard deviation. By each tenure group and period, the top and bottom 5% of observations are winsorized. Sampling weights are unused, and the maximum tenure is not restricted in the calculation of the summary statistics. The data is from the Michigan Survey of Consumers, July 1980 to October 2021.

Table A.9: Tests for Significant Differences between the Coefficients across Survey Tenure in Table A.5

	$\hat{\sigma}$	$\hat{\gamma}$	$\hat{\sigma}$	$\hat{\gamma}$
F-statistic	2.62	0.93	3.49	0.91
P-value	0.0061	0.5204	0.0004	0.5303
Regression Type	OLS	OLS	IV	IV
Observations	54970	54970	54850	54850

Note: We have conducted F-tests for H0: $\hat{\sigma} = \hat{\sigma}_s \ \forall s$ and H0: $\hat{\gamma} = \hat{\gamma}_s \ \forall s$ for both OLS and IV cases in Column (2) and (3) of Table A5. Driscoll-Kraay standard errors of lag one are used. More detailed regression setups are described in Table A5. For $\hat{\sigma}$, we could reject the null hypothesis at 1% significance level that the coefficients across the tenure groups are equal. On the contrary, we could not reject the null for the case of $\hat{\gamma}$.

A.3 Questionnaire

A.3.1 SCE questions related to inflation, unemployment, and other price changes

- **Q2**

  – And looking ahead, do you think you (and any family living with you) will be financially better or worse off **12 months from now** than you are these days?

    Instruction H1

    ◯ Much worse off

    ◯ Somewhat worse off

    ◯ About the same

    ◯ Somewhat better off

    ◯ Much better off

    If not response: error E1

- **Q8v2**

  – The next few questions are about inflation. **Over the next 12 months,** do you think that there will be inflation or deflation? (Note: deflation is the opposite of inflation)

    Instruction H8

    ◯ Inflation

◯ Deflation

- **Q8v2part2**
  - What do you expect the rate of [inflation/deflation as in Q8v2] to be **over the next 12 months**? Please give your best guess.

    Instruction H9

    Over the next 12 months, I expect the rate of [inflation/deflation] to be _____ %

- **Q9**
  - Now we would like you to think about the different things that 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...**

    Instruction H4

    the rate of inflation will be 12% or higher: _____ percent chance

    the rate of inflation will be between 8% and 12%: _____ percent chance

    the rate of inflation will be between 4% and 8%: _____ percent chance the rate of inflation will be between 2% and 4%: _____ percent chance

the rate of inflation will be between 0% and 2%: _____ percent chance

the rate of deflation (opposite of inflation) will be between 0% and 2%: _____ percent chance

the rate of deflation (opposite of inflation) will be between 2% and 4%: _____ percent chance

the rate of deflation (opposite of inflation) will be between 4% and 8%: _____ percent chance

the rate of deflation (opposite of inflation) will be between 8% and 12%: _____ percent chance

the rate of deflation (opposite of inflation) will be 12% or higher: _____ percent chance

TOTAL 100

If sum not equal to 100: "Your total adds up to XX" followed by an error message

- **C2**

  And in your view, what would you say is the percent chance that, over the next 12 months, the average home price nationwide will...

  Instruction H4

increase by 12% or more: ______ percent chance

increase by 8% to 12%: ______ percent chance

increase by 4% to 8%: ______ percent chance

increase by 2% to 4%: ______ percent chance

increase by 0% to 2%: ______ percent chance

decrease by 0% to 2%: ______ percent chance

decrease by 2% to 4%: ______ percent chance

decrease by 4% to 8%: ______ percent chance

decrease by 8% to 12%: ______ percent chance

decrease by 12% or more: ______ percent chance

TOTAL 100

If sum not equal to 100: "Your total adds up to XX" followed by an error message

- **Q4new**
  - What do you think is the percent chance that 12 months from now the unemployment rate in the U.S. will be higher than it is now?

    Instruction H2

- **C4info**

- Twelve months from now, what do you think will have happened to the price of the following items? *Instruction H11* I expect...

  The price of a gallon of gas to have increased by _____ percent or decreased by _____ percent

  The price of food to have increased by _____ percent or decreased by _____ percent

  The price of medical care to have increased by _____ percent or decreased by _____ percent

  The price of a college education to have increased by _____ percent or decreased by _____ percent

  The cost of renting a typical house/apartment to have increased by _____ percent or decreased by _____ percent

  The price of gold to have increased by _____ percent or decreased by _____ percent

A.3.2 Questions related to Future Income/Earning in SCE

- **Q23v2**

  - Please think ahead to **12 months from now.** Suppose that you are working in the exact same job at the same place you currently work, and working the exact same number of hours. What do you expect to have happened to your earnings on this job, before taxes and deductions?

Instruction H8

Twelve months from now, I expect my earnings to have

◯ Increase by 0% or more

◯ Decrease by 0% or more

- **Q23v2part2**
  - By about what percent do you expect your earnings to have [increased/decreased as in Q23v2]? Please give your best guess.

    Instruction H9

    Twelve months from now, I expect my earnings to have [increased/decreased] by _____ %

- **Q24**
  - Suppose again that, 12 months from now, you are working in the exact same job at the same place you currently work, and working the exact same number of hours. In your view, what would you say is the percent that 12 months from now...

    Instruction H4

    Your earnings on this job, before taxes and deductions, will have...

    increase by 12% or more: _____ percent chance

    increase by 8% to 12%: _____ percent chance increase by 4% to 8%: _____ percent chance

increase by 2% to 4%: _____ percent chance

increase by 0% to 2%: _____ percent chance

decreased by 0% to 2%: _____ percent chance

decreased by 2% to 4%: _____ percent chance

decreased by 4% to 8%: _____ percent chance

decreased by 8% to 12%: _____ percent chance

decreased by 12% or more: _____ percent chance

TOTAL 100

If sum not equal to 100: "Your total adds up to XX" followed by an error message

- **Q25v2**
  - Next we would like to ask you about your overall household income going forward. By household we mean everyone who usually lives in your primary residence (including yourself), excluding roommates and renters.

    Over the next 12 months, what do you expect will happened to the total income of all members of your household (including you), from all sources before taxes and deductions?

    Instruction H8

Over the next 12months, I expect my total household income to...

◯ increase by 0% or more

◯ decrease by 0% or more

- **Q25v2part2**
  - By about what percent do you expect your total household income to

    [increased/decreased as in Q25v2]? Please give your best guess.

    Instruction H9

    Over the next 12 months, I expect my total household income to

    [increased/decreased] by _____ %

A.3.3 Questions related to inflation in MSC

- **A12**
  - During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?

    1.GO UP 2.STAY THE SAME 5.GO DOWN 8.DON'T KNOW

    (If answer 2 is chosen then go to A12a. For 1, go to A12b. For 5, go to A12c.)

- **A12a**
  - Do you mean that prices will go up at the same rate as now, or that prices in general will not go up during the next 12 months?

    2.GO UP 3.WILL NOT GO UP

- **A12b**
  - By about what percent do you expect future prices to go (up/down) on the average, during the next 12 months?

    _____ PERCENT

  - DON'T KNOW (Go to A12c if this is chosen)

- **A12c**

    (AFTER A DON'T KNOW RESPONSE IS PROVED, IF R SAYS, "I DON'T KNOW" USE THE FOLLOWING PROBE:)

    (USE PROBE BELOW IF ANSWER IS GREATER THAN 5%)

  - How many cents on the dollar do you expect prices to go (up/down) on the average, during the next 12 months?

    _____ CENTS ON DOLLAR

  - DON'T KNOW

- IF R GIVES AN ANSWER THAT IS GREATER THAN 5%, PLEASE PROBE WITH:

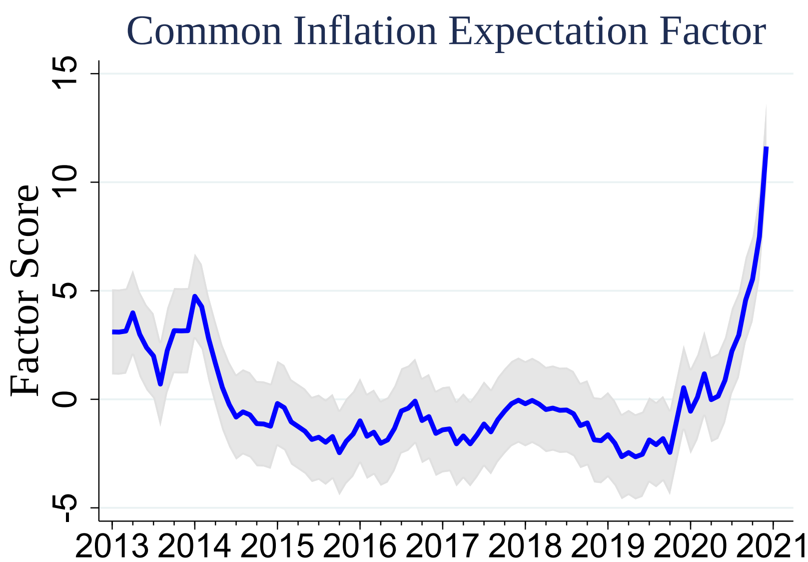
"Let me make sure I have that correct. You said that you expect prices to go (up/down) during the next 12 months by (X) percent. Is that correct?"

Appendix B

Appendix of Chapter2

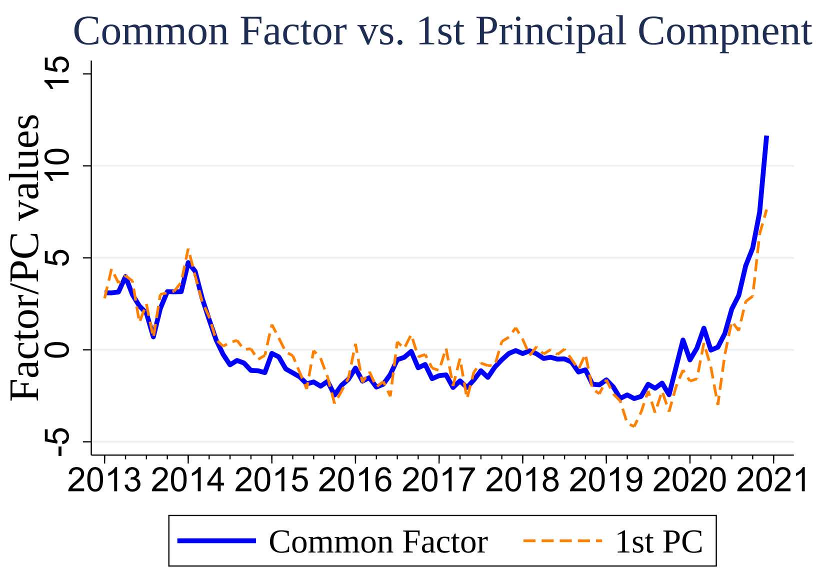
B.1 Figures

Figure B.1: Estimated Common Inflation Expectation Factor of the SCE



Note: The raw values of common inflation expectation factor of the SCE are displayed without a projection (F_t^{Common} of Equation 2.2). y-axis shows raw factor scores. The gray area shows a 95% confidence interval for the solid blue line of each panel. I winsorize the top and bottom 5% of point inflation forecasts for each period. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to May 2021 with monthly frequency.

Figure B.2: Estimated Common Factor vs. 1st Principal Component of Inflation Expectation of the SCE



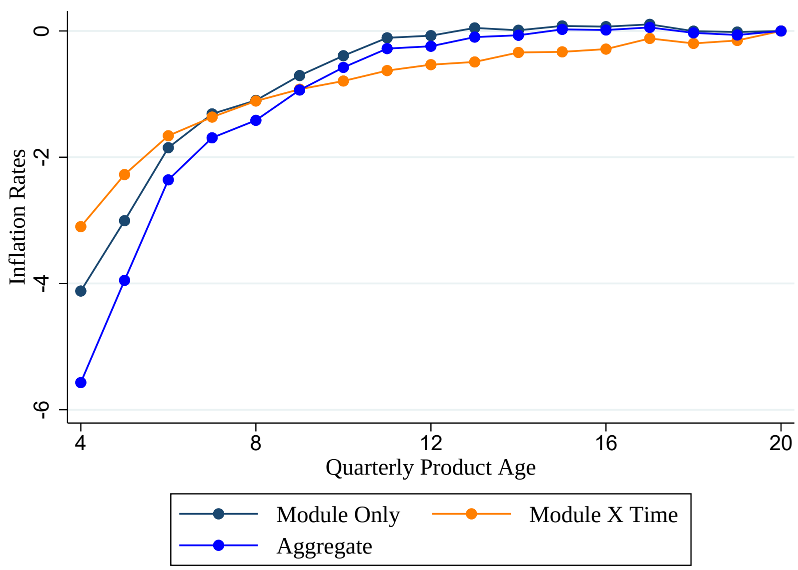
Note: The solid blue line ("Common Factor") shows estimated common inflation expectation factor of the SCE without a projection (F_t^{Common} of Equation 2.2). The dashed orange line ("1st PC") shows the first principal component of all inflation expectations of the SCE (Table 2.1). y-axis shows raw factor scores or principal component values. The correlation coefficient between those two time-series is 0.90. I winsorize the top and bottom 5% of point inflation forecasts for each period. Sampling weights are unused in the calculation of median. Data is from the FRBNY Survey of Consumer Expectations, from June 2013 to May 2021 with monthly frequency.

Appendix C

Appendix of Chapter3

C.1 Figures

Figure C.1: Price Increase Rate Changes over Product Life Cycle



Note: The figure shows the change in price increase rates of new products over product life compared to their initial price levels, in percentage points. Product age measured in quarters is shown on the x-axis. Data is from the RMS, from January 2006 to December 2017.

Bibliography

Aguiar, M. and E. Hurst (2007). Life-cycle prices and production. *American Economic Review 97*(5), 1533–1559.

Ahn, H. J., C. Fulton, et al. (2020). Index of common inflation expectations. Technical report, Board of Governors of the Federal Reserve System (US).

Argente, D. and M. Lee (2021). Cost of living inequality during the great recession. *Journal of the European Economic Association 19*(2), 913–952.

Argente, D., M. Lee, and S. Moreira (2019). The life cycle of products: Evidence and implications. *Available at SSRN 3163195*.

Arioli, R., C. Bates, H. Dieden, I. Duca, R. Friz, C. Gayer, G. Kenny, A. Meyler, and I. Pavlova (2017). Eu consumers' quantitative inflation perceptions and expectations: An evaluation. Technical report, ECB Occasional Paper.

Armantier, O., G. Topa, W. van der Klaauw, and B. Zafar (2017). An overview of the survey of consumer expectations. *Federal Reserve Bank of New York Economic Policy Review 23*(2), 51.

Baffes, J., M. A. Kose, F. Ohnsorge, and M. Stocker (2015). The great plunge in oil prices: Causes, consequences, and policy responses. *Consequences, and Policy Responses (June 2015)*.

Bai, J. and P. Wang (2012). Identification and estimation of dynamic factor models.

Bańkowska, K., A. Borlescu, E. Charalambakis, A. Dias da Silva, D. Di Laurea, M. Dossche, D. Georgarakos, J. Honkkila, N. Kennedy, G. Kenny, et al. (2021). Ecb consumer expectations survey: an overview and first evaluation. *ECB Occasional Paper* (2021/287).

Bellemare, C., R. KpekouTossou, and K. Moran (2020). The determinants of consumers' inflation expectations: Evidence from the us and canada. Technical report, Bank of Canada Staff Working Paper.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (favar) approach. *The Quarterly journal of economics 120*(1), 387–422.

Bils, M. (2009). Do higher prices for new goods reflect quality growth or inflation? *The Quarterly Journal of Economics 124*(2), 637–675.

Binder, C. (2017a). Consumer forecast revisions: Is information really so sticky? *Economics Letters 161*, 112–115.

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

Binder, C. C. (2020). Coronavirus fears and macroeconomic expectations. *Review of Economics and Statistics 102*(4), 721–730.

Candia, B., O. Coibion, and Y. Gorodnichenko (2020). Communication and the beliefs of economic agents. Technical report, National Bureau of Economic Research.

Candia, B., O. Coibion, and Y. Gorodnichenko (2021). The inflation expectations of us firms: Evidence from a new survey. Technical report, National Bureau of Economic Research.

Carroll, C. D., E. Crawley, J. Slacalek, K. Tokuoka, and M. N. White (2020). Sticky expectations and consumption dynamics. *American economic journal: macroeconomics 12*(3), 40–76.

Coibion, O. and Y. Gorodnichenko (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Coibion, O. and Y. Gorodnichenko (2015b). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics 7*(1), 197–232.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018). How do firms form their expectations? new survey evidence. *American Economic Review 108*(9), 2671–2713.

Coibion, O., Y. Gorodnichenko, and M. Ulate (2019, May). Is inflation just around the corner? the phillips curve and global inflationary pressures. *AEA Papers and Proceedings 109*, 465–69.

Coibion, O., Y. Gorodnichenko, and M. Weber (2020). Monetary policy communications and their effects on household inflation expectations. Technical report, National Bureau of Economic Research.

Crump, R. K., S. Eusepi, A. Tambalotti, and G. Topa (2015). Subjective intertemporal substitution. *FRB of New York Staff Report* (734).

Deaton, A. S. and C. Paxson (1994). Saving, growth, and aging in taiwan. In *Studies in the Economics of Aging*, pp. 331–362. University of Chicago Press.

Dempster, A. P. (1977). Maximum likelihood estimation from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 39*, 1–38.

Doz, C., D. Giannone, and L. Reichlin (2012). A quasi–maximum likelihood approach for large, approximate dynamic factor models. *Review of economics and statistics 94*(4), 1014–1024.

Dräger, L. and M. J. Lamla (2017). Imperfect information and consumer inflation expectations: Evidence from microdata. *Oxford Bulletin of Economics and Statistics 79*(6), 933–968.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics 82*(4), 540–554.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics 129*(4), 1661–1710.

Halpern-Manners, A., J. R. Warren, and F. Torche (2017). Panel conditioning in the general social survey. *Sociological methods & research 46*(1), 103–124.

Heckman, J. and R. Robb (1985). Using longitudinal data to estimate age, period and cohort effects in earnings equations. In *Cohort analysis in social research*, pp. 137–150. Springer.

Hobcraft, J., J. Menken, and S. Preston (1985). Age, period, and cohort effects in demography: a review. In *Cohort analysis in social research*, pp. 89–135. Springer.

Hobijn, B. and D. Lagakos (2005). Inflation inequality in the united states. *review of income and Wealth 51*(4), 581–606.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the us retail sector. *The Quarterly Journal of Economics 134*(2), 715–783.

Jaravel, X. (2021). Inflation inequality: Measurement, causes, and policy implications. *Annual Review of Economics 13*, 599–629.

Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics 91*, 19–38.

Karahan, F., S. Mihaljevich, and L. Pilossoph (2017). Understanding permanent and temporary income shocks. *URL link retrieved on 3*(02), 2018.

Kim, G. and C. Binder (Conditionally Accepted, 2022). Learning-through-survey in inflation expectations. *American Economic Journal: Macroeconomics*.

Kotchoni, R., M. Leroux, and D. Stevanovic (2019). Macroeconomic forecast accuracy in a data-rich environment. *Journal of Applied Econometrics 34*(7), 1050–1072.

Machado, J. A. and J. S. Silva (2019). Quantiles via moments. *Journal of Econometrics 213*(1), 145–173.

Mackowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review 99*(3), 769–803.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Miranda, K., P. Poncela, and E. Ruiz (2021). Dynamic factor models: Does the specification matter? *SERIEs*, 1–32.

Niu, X. and N. Harvey (2021). Context effects in inflation surveys: The influence of additional information and prior questions. *International Journal of Forecasting*.

Powell, J. (2019). Hearing before the committee on banking, housing, and urban affairs. Technical report, The Semiannual Monetary Policy Report to the Congress.

Reis, R. (2006). Inattentive consumers. *Journal of Monetary Economics 53*(8), 1761–1800.

Sargent, T. J., C. A. Sims, et al. (1977). Business cycle modeling without pretending to have too much a priori economic theory. *New methods in business cycle research 1*, 145–168.

Shumway, R. H. and D. S. Stoffer (1982). An approach to time series smoothing and forecasting using the em algorithm. *Journal of time series analysis 3*(4), 253–264.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. *NBER macroeconomics annual 4*, 351–394.

Stock, J. H. and M. W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Ueda, K., K. Watanabe, and T. Watanabe (2019). Product turnover and the cost-of-living index: Quality versus fashion effects. *American Economic Journal: Macroeconomics 11*(2), 310–47.

Verbrugge, R. J. and C. C. Binder (2016). Digging into the downward trend in consumer inflation expectations. *Economic Commentary* (2016-11).

Warren, J. R. and A. Halpern-Manners (2012). Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research 41*(4), 491–534.

Watson, M. W. and R. F. Engle (1983). Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models. *Journal of Econometrics 23*(3), 385–400.