How Does Learning Affect Overconfidence? Evidence from Survey Data

Chenguang Shang

cshang@suffolk.edu

Sawyer Business School, Suffolk University, Boston, MA 02108, USA

Abstract

Using the Survey of Consumer Expectations (SCE) data on individuals' repeat forecasts on future inflation, this study empirically examines how learning affects overconfidence. Analyses show that as respondents learn more about inflation between survey waves, their overconfidence increases. The effect is mainly seen among those who are initially uncertain but accurate. The influence of learning on overconfidence is then confirmed in survey participants' predictions of home price changes. Overall, the evidence provided in this study sheds new light on the dynamic nature of overconfidence and underscores the challenge of overcoming behavioral biases in a novel financial economic setting.

1. Introduction

Overconfidence is one of the most significant cognitive biases and has been associated with various undesirable consequences, such as investors' excessive trading (Odean, 1999), corporate investment distortions (Malmendier and Tate, 2005), business failures (Koellinger, Minniti, and Schade, 2007), medical misdiagnosis (Berner and Graber, 2008), and even wars (Johnson, 2004).[1] While it is often considered an innate personal trait, overconfidence is not necessarily static -- people continuously learn from their observations and experiences. It is therefore plausible that human behavioral biases, such as overconfidence, are dynamic and can evolve over time. The main objective of this study is to empirically examine whether and how learning affects overconfidence.

Several important works in the psychology and financial economics literature offer guidance on the relation between learning and overconfidence. For instance, Gervais and Odean (2001) provide a theoretical framework in which, because investors tend to attribute too much of their success to their own ability rather than luck, the process of learning can potentially intensify overconfidence. Sanchez and Dunning (2018) propose a "beginner's bubble" hypothesis and employ an experimental approach to demonstrate that, although people are typically cautious when approaching unfamiliar tasks, they quickly become overconfident after a few learning experiences. Using forecasts of one-year S&P 500 returns made by Chief Financial Officers (CFOs), Boutros, Ben-David, Graham, Harvey, and Payne (2020) show that CFOs are overconfident in their forecast precision, and despite their updated beliefs upon observing the outcomes of their forecasts, CFOs' adjustments are not sufficient to eliminate miscalibration. The implication derived from this line of research, therefore, is that even though people may be able to correctly update their beliefs

[1] Meanwhile, research has shown some positive effects of overconfidence, such as the achievement of higher social status (Anderson, Brion, Moore, and Kennedy, 2012) and greater innovative successes (Hirshleifer, Low, and Teoh, 2012).

through learning, the improvement in their performance can be countered or even outpaced by the growth in their confidence along the way. A classic example of this learning-overconfidence relation is that fatal crash rates increase for new pilots as they accumulate more flying hours (Craig, 2013).

This paper advances this important yet still scarce line of research by employing a novel setting that leverages the Survey of Consumer Expectations (SCE) data on individuals' repeat forecasts on future inflation. Several features of the SCE data make it ideal to test the relation between learning and overconfidence. First of all, the SCE collects a nationally representative sample of consumers. To the extent that most of the empirical analyses in this field are limited to laboratory experiments, subject to relatively small samples, or focused on specific groups of the population (e.g., corporate executives or pilots), the sample employed in this study facilitates an investigation of the issue at hand among the general public.

Second, by eliciting respondents' subjective probability distributions, the SCE is able to compute the density inter-quartile difference between the 75th percentile and the 25th percentile, which is essentially the 50% confidence intervals of respondents' inflation forecasts (i.e., individuals are 50% certain that the true inflation will fall within this range). Thus, the analysis of whether realized future inflation is captured by the 50% confidence intervals provides valuable insights into whether individuals are overconfident, a strategy commonly adopted to evaluate overconfidence (see, for example, Soll and Klayman, 2004).

Third, critically pertinent to this study, the respondents in the SCE make inflation forecasts every month for up to 12 months. The repeat nature of the survey enables enable the construction of overconfidence indicators for each individual over a 12-month period and thereby track the evolution of their level of confidence over time.

Finally, Kim and Binder (2023), also using the SCE data, demonstrate "learning-through-survey" effects by documenting that respondents significantly revise their inflation expectations and reduce their forecast errors in subsequent surveys. The evidence of learning that is observed in this data serves as a crucial foundation upon which the analyses in this paper are performed.

This study shows that overconfidence is a common bias embedded in survey participants' inflation forecasts: realized one-year-ahead inflation falls within respondents' 50% confidence intervals only 32% of the time. This phenomenon is observed across different demographic groups such as gender, age, race, wealth, and education. As individuals repeat their inflation forecasts over the course of 12 surveys, they become more overconfident about the precision of their forecasts: the likelihood of the 50% confidence level including realized future inflation decreases from 38.1% in the first survey to 30.8% in the last survey.

I then take a closer look and examine the movement of the upper and lower bounds of the 50% confidence interval over the 12-survey period, and find that they converge towards the middle. Specifically, the 25th percentile forecasts increase, while the 75th percentile forecasts decrease. The changes in the confidence level boundaries in turn affect whether the range between them captures realized inflation. Analyses show that although both the upper and lower bounds become less likely to contain the one-year-ahead inflation rate, the bigger impact appears to come from adjustments made to the upper bound; that is, the reduction in the 75th percentile forecasts is more likely to exclude the actual inflation rate from the confidence interval.

Overconfidence essentially reflects the misalignment between a person's true ability and the person's confidence in their ability. In the context of this study, forecast uncertainty (i.e., an individual's confidence in their ability to predict future inflation) and forecast accuracy (i.e., the individual's true ability to predict future inflation) play an important role in determining whether realized inflation falls within the 50% confidence interval. I therefore conduct an examination of the roles of uncertainty and accuracy. Since forecast uncertainty and accuracy may change as a result of learning, I focus on survey participants' *initial* uncertainty and accuracy exhibited in the first round of the survey. I first show that individuals who are initially uncertain / inaccurate in their inflation forecasts behave very differently from those who are initially certain / accurate. The "learning-through-survey" effect documented in Kim and Binder (2023) is concentrated among individuals with high initial uncertainty or with low initial accuracy; these respondents experience a substantial decrease in uncertainty / improvement in accuracy over the 12 surveys. Individuals with high initial conviction become slightly more uncertain over time, and those with high initial accuracy see a deterioration in their subsequent forecast performance. Analyses on the joint effect of uncertainty and accuracy suggest that a majority of the main effect that learning leads to an increase in overconfidence comes from survey respondents who are initially uncertain but accurate. These individuals tend to substantially narrow their 50% confidence intervals over the course of the survey while their forecast accuracy suffers a decline.

To further explore the impact of learning on the evolution of overconfidence, I adopt another set of forecast estimates collected by the SCE, namely respondents' repeat forecasts on changes in future home prices. Similar to the analyses based on inflation forecast data, I examine the sequential change in survey participants' overconfidence, and find that, consistent with the main finding, individuals are on average overconfident in their home price estimates – their 50% confidence interval contains the realized home price percent change less than 25% of the time, and that their confidence intervals become less likely to contain realized home price changes over the course of the survey.

This study contributes to the literature in several important ways. First, not only do the documented results support the notion that people are generally overconfident in a novel setting (i.e., inflation forecast), but more importantly, they extend the extensive literature established by psychologists and financial economists by steering the conversation on overconfidence in a static sense to a dynamic one. That is, individuals do not necessarily remain overconfident at the same level. Instead, their own actions (e.g., learning), or perhaps mere exposure to certain information, may alter their overconfidence in a dynamic way. More broadly, the findings underscore the vast variability in human judgment (Kahneman, Sibony, and Sunstein, 2021).

In addition, this study adds to the discussion on the outcomes of learning. Undoubtedly, the ability to learn is a magnificent gift that humans have taken advantage of throughout history, and has played a tremendous role in the course of human evolution by allowing knowledge to be created and transmitted. Nonetheless, the process of learning may come with some limitations. For instance, it does not fully eliminate, and in some situations may even exacerbate, behavioral biases such as overconfidence (Boutros, Ben-David, Graham, Harvey, and Payne, 2020; Sanchez and Dunning, 2018). The evidence reported in this paper shines a fresh light on this front.

The unique setting employed in this study helps overcome some empirical challenges faced by prior research on the relation between learning and overconfidence, such as limited sample size, the lack of subjects that can represent the overall population, and the difficulty in observing behavioral biases of the same person over time. This work directly answers Sanchez and Dunning's (2018) call on future research to examine the issue using "a longitudinal analysis tracking the same respondents through time (page 24)."

This paper is also timely in that inflation has been one of the most worrying topics, not only to financial economists but to virtually everyone in society, in the post-COVID era.

According to a survey conducted by Pew Research in 2024, inflation is the top issue on the public's list of the biggest problems facing the country – 62% of Americans describe inflation as a very big problem.[2] A growing literature has identified various individual-level factors that can explain the formation of people's inflation expectations, such as exposure to price changes (D'Acunto, Malmendier, Ospina, and Weber, 2021), cognitive abilities (D'Acunto, Hoang, Paloviita, and Weber, 2023), and communication forms and strategies (Coibion, Gorodnichenko, and Weber, 2022; Coibion, Georgarakos, Gorodnichenko, and Weber, 2023). This study joins Kim and Binder (2023) in suggesting that consumers' beliefs and the conviction of their inflation expectations can change drastically in a repeat survey setting.

The remainder of this paper is organized as follows. Section 2 describes the sample, the construction of key variables, and summary statistics. Section 3 presents the empirical analyses. Concluding remarks are in Section 4.

2. Sample Selection, Key Variables, and Descriptive Statistics

The main source of data is the Survey of Consumer Expectations (SCE) administered by the Federal Reserve Bank of New York. Based on a nationally representative sample, the SCE gathers information on individuals' expectations on a wide range of economic issues, such as inflation, household finance, and labor and housing markets (Armantier, Topa, Van der Klaauw, and Zafar, 2017). The data also includes survey respondents' personal characteristics, such as gender, age, race, wealth, and education. Information on historical inflation rates is made available by the U.S. Bureau of Labor Statistics and is downloaded from the Federal Reserve Bank of St. Louis website. The sample period employed in this study ranges between June 2013 and May 2023. The start of the sample period marks the launch of the SCE, and the end allows for the observation

[2] https://www.pewresearch.org/politics/2024/05/23/top-problems-facing-the-u-s/

of the realized inflation as of May 2024, one year after the last inflation expectation record. To be included in the sample, a survey respondent must have twelve consecutive survey participation records. This requirement aims to facilitate the examination of the evolution of respondents' overconfidence over time. There are 75,444 observations in the sample, representing 6,287 unique survey respondents.[3] Detailed variable definitions are provided in Appendix 1.

2.1. Inflation Expectations

The survey on individuals' inflation expectations is conducted monthly, and respondents stay on the rotating panel for up to twelve months. Each month, the respondent is asked to assign a percent chance to ten inflation rate intervals, ranging from "the rate of deflation (opposite of inflation) will be 12% or higher" to "the rate of inflation will be 12% or higher." A sample of the question, obtained from the SCE questionnaire, can be found in Appendix 2. It is worth noting that the repeat responses make it possible to track the same individual over time, and is key to the analyses of the evolution of overconfidence as an individual learns more about inflation during the course of the survey. To ensure fair comparison and mitigate the concern that inflation expectations may be driven by some unobserved factors that are related to the duration of respondents' participation in the survey (for instance, respondents experiencing extenuating family issues may have a pessimistic view on future economic outlooks and are less likely to stay in the survey for the entire duration), similar to Kim and Binder (2023), I retain in the sample only respondents who have completed 12 consecutive waves of survey.

Critical to the empirical analyses employed in this study, the SCE inflation forecast data includes the mean, 25th percentile, and 75th percentile estimates of respondents' one-year-ahead

[3] The number of observations in some analyses may be slightly less than 75,444 because even though each respondent appears 12 times in the survey, a small number of them have missing values for their personal characteristics and inflation forecasts.

inflation forecasts. The mean forecasts reflect the central tendency of individuals' expectation distributions, or the "expected inflation rate." Meanwhile, the range between the 25th and 75th percentile estimates is the 50% confidence interval, which also captures respondents' uncertainty. In theory, if a survey participant is properly calibrated, realized future inflation should fall within the confidence interval 50% of the time.

Table 1 Panel A shows that during the sample period, survey respondents' inflation expectation over the next 12 months is 4.111% on average. Their 25th and 75th percentile estimates are 2.058%, and 6.136%, respectively, which results in an inter-quartile range (i.e., inflation expectation uncertainty) of 4.078%.

<Table 1 about here>

2.2. Realized Inflation

For every monthly inflation expectation record, the one-year-ahead realized inflation is calculated as the percent change of the consumer price index for all urban consumers (CPI) over the 12-month period following the month of the survey. By way of example, for a respondent who provided inflation estimates in June 2013, the corresponding one-year ahead realized inflation rate is based on the change in CPI from July 2013 to June 2014. As shown in Figure 1, from June 2013 to May 2023, the one-year-ahead realized inflation rate fluctuated below 3% during the first half of the sample period, spiked during the COVID-19 pandemic to a peak of around 9% in June 2021, and then leveled off in the last couple of years of the sample. Survey respondents' inflation expectations generally exceed realized actual inflation, which is consistent with Kim and Binder (2023), except for the COVID era. Unless noted otherwise, the realized inflation rates employed in this study are not seasonally adjusted because it is not clear that individuals rationally factor in seasonality when forming their inflation expectations. Nonetheless, I consider seasonally adjusted inflation in the robustness check section, and show that it does not affect the main finding documented in this study.

<Figure 1 about here>

2.3. Individual Characteristics

Panel B of Table 1 provides information on the demographics of respondents in the SCE. The personal characteristics reported in the table include gender, age, race, wealth, and education. Specifically, among the individuals included in the sample, 45.8% are female, 25.3% are young (under 40 years old), 85.6% are white, 28.4% are rich (with a household income over $100,000), and 57.3% have a college degree. These statistics are closely in line with those reported in Armantier, Topa, Van der Klaauw, and Zafar (2017).

3. Empirical Analyses and Results

This section presents the empirical analyses performed to address the key research questions. First, I document overconfidence behavior in the SCE data. I then demonstrate that the level of overconfidence increases over the twelve waves of the survey. The role of individuals' initial uncertainty and accuracy is explored next. Finally, respondents' expectations on future home price changes are used to confirm the findings based on inflation forecasts.

3.1. Are Individuals Overconfident in the SCE sample?

To gauge the extent of overconfidence, I create a dummy variable that is equal to one if the one-year-ahead realized inflation rate falls between the 25th percentile and 75th percentile inflation forecasts, and zero otherwise. Intuitively, when individuals are objective about their ability to forecast, the probability of a future event occurring should approximately match their confidence intervals (Soll and Klayman, 2004; Ben-David, Graham, and Harvey, 2013). In the context of this study, if survey respondents are properly calibrated, their confidence intervals (i.e., the range between their 25th percentile and 75th percentile estimates) should contain realized future inflation roughly 50% of the time, which is referred to as the "hit rate" thereafter.

As shown in Table 2 Panel A, the one-year-ahead hit rate (Hit) is 32.1%, which is much lower than the expected level of 50%. This result, based on a nationally representative sample, supports the popular notion that people are on average overconfident. The same panel also displays the likelihood of realized inflation landing outside the 50% confidence interval: 43.9% of the realized inflation outcomes fall below the lower bound (Miss_P25), and 24.0% exceed the upper bound (Miss_P75). The high missed rate below the lower bound suggests that the 25th percentile estimates are likely set too high when inflation in reality is quite low, which is consistent with Kim and Binder (2023), who report that individuals tend to over-estimate future inflation.

To better understand whether the overconfidence behavior observed in the full sample is concentrated within certain groups of individuals, Panel B of Table 2 breaks down the hit rate by survey participants' personal characteristics. It is evident that across all sub-groups based on gender, age, race, wealth, and education, no hit rate exceeds 40% for respondents' 50% confidence levels. As a result, it appears that overconfidence is a common phenomenon that is present in all population groups in the study. However, it is worth noting that the hit rate can differ substantially along some demographic dimensions. For instance, younger survey participants have a significantly higher hit rate compared to older participants (36.9% vs. 30.4%). Although not the focus of this study, these findings support the view that demographics can play a role in explaining human behavioral biases (e.g., Bhandari and Deaves, 2006).

<Table 2 about here>

3.2. Evidence of Learning

Two considerations determine whether survey participants successfully place realized inflation within their 50% confidence intervals. One is the level of uncertainty, as captured by the width of respondents' 50% confidence intervals. The other is the accuracy of their forecasts, which is measured as the absolute difference between the density mean of each individual's inflation expectation distribution and realized future inflation.[4] These two factors may collectively contribute to the presence of overconfidence. For instance, if the mean forecast happens to be very close to realized inflation, there is a good chance that even a narrow 50% confidence interval may include the actual inflation outcome. Meanwhile, a high level of uncertainty may compensate for a large deviation of realized inflation from the central location of the expectation distribution; that is, the effect of inaccurate forecasts can be mitigated if respondents widen their confidence intervals.

Kim and Binder (2023), also using the SCE data, discover that repeat survey respondents become more informed and less uncertain about future inflation, and coin the effect "learning-through-survey." Since the investigation of the main research question in this study crucially depends on the idea that respondents learn over the course of a repeat survey, it is worthwhile presenting respondents' learning pattern as documented in Kim and Binder (2023) before delving into the analyses of how learning affects overconfidence.

<Table 3 about here>

Column 2 of Table 3 displays the sequential change in survey respondents' inflation expectation uncertainty. There is a sharp decline in the level of uncertainty. The range between the 25^{th} percentile and 75^{th} percentile estimates of future inflation narrows by more than 30% over the course of the survey, from 5.414% down to 3.635%. In the meantime, Column 2 shows that survey

[4] The density mean is derived from the forecast density function based on each individual's responses to the probabilistic questions regarding future inflation and is considered the respondent's "expected inflation rate."

respondents' forecasts are far from accurate: the average error of one-year-ahead forecasts is 3.608%. Individuals' inflation expectations gradually approach realized inflation over the 12-month survey period; the last round of survey is associated with a forecast error that is about 12% lower than that in the first round. A visualization of the impact of learning (i.e., survey sequence) on forecast uncertainty accuracy is depicted in Figure 2.

<Figure 2 about here>

Collectively, these findings confirm the "learning-through-survey" effect documented in Kim and Binder (2023).[5] That is, as participants pay more attention to specific survey topics, they become more confident and accurate in their inflation forecasts. These sequential patterns, in conjunction with the finding that participants in the SCE are on average overconfident, also set the stage for the examination of the key question this paper aims to address: how does learning affect overconfidence?

3.3. Learning and Hit Rate

The results presented in the previous section do not automatically convey information regarding how survey participants' overconfidence evolves with learning, as the relation can be influenced by competing forces. On one hand, improved forecast accuracy may be indicative of enhanced forecasting skills, and thus may suggest that through learning, individuals become better at fitting future inflation into their 50% confidence intervals. On the other hand, it is also possible that respondents shrink their 50% confidence intervals (i.e., reduces their uncertainty) faster than they improve their forecast performance, leading to more miscalibration and greater overconfidence. This section formally analyzes how the hit rate changes with learning.

[5] Unreported regression analyses document consistent results.

Recall that the overall hit rate for one-year-ahead inflation forecasts is well below the expected 50% level (Table 2). Panel A of Table 4 presents survey participants' average hit rates for each of the 12 surveys. In the full sample (Column 1), the hit rate is 38.1% in the first round of the survey. It quickly drops to about 30% halfway through, and stabilizes afterwards. Put differently, the likelihood of respondents including realized inflation in their 50% confidence interval decreases substantially over time, signaling an increase in overconfidence. The difference in hit rate between the first and last surveys is statistically significant at the 1% level. A graphic demonstration of the change in hit rate based on the full sample is displayed in Figure 3.

<Figure 3 about here>

Columns 2 – 11 shows the sequential change in hit rate over the 12-survey period in different population groups, namely female vs. male, young vs. old, white vs. non-white, rich vs. poor, and college vs. non-college. The decreasing pattern of hit rate (i.e., increasing pattern of overconfidence) holds remarkably well in each column, suggesting that the positive relation between learning and overconfidence is applicable to all demographic groups.

Having documented a clear univariate trend, I next analyze the relation in a regression setting. To do so, I employ a linear probability model as follows:

$$Hit_{its} = \alpha + \sum_{n=2}^{12} \beta_n Survey_n + \theta_i + \mu_{ts} + \varepsilon_{its}, \qquad (1)$$

where Hit is a dummy variable indicating whether an individual's 50% confidence interval includes realized future inflation, and Survey is an indicator that represents the n^{th} survey an individual completes. User fixed effects, as captured by θ, are included in all regressions to remove the possibility that some inherent personal traits may affect the development of confidence over time.[6] In addition, I control for year-by-state fixed effects, μ, in order to isolate any macro-level factors that may vary by geographic location (e.g., the different state-level policies implemented in response to the COVID-19 pandemic). Standard errors are clustered at the respondent level.

<Table 4 about here>

As is shown in Column 1 of Panel B in Table 4, all the survey indicators carry a negative coefficient estimate that is statistically significant at the 1% level, suggesting that the one-year-ahead inflation forecast hit rate is significantly lower in subsequent surveys compared to the initial one. Furthermore, consistent with the univariate evidence presented in Panel A, the value of the coefficient estimates declines sharply initially and then stabilizes. The coefficient estimate on Survey 12 indicates that, all else equal, relative to the first round of the survey, the chances of respondents' 50% confidence intervals containing realized inflation are approximately 20.7% (-7.9%/38.1%) lower in the last round of the survey. This finding is consistent with the univariate results reported in Panel A, and supports prior research that shows a positive relation between learning and overconfidence (Sanchez and Dunning, 2018). The regressions for different sub-groups of the population (Columns 2 – 11 of Panel B) unveil a very similar pattern. Given the consistency between univariate and regression analysis results, in the interest of brevity, analyses beyond this section only report the univariate trends in overconfidence as in Panel A of Table 4; regression analysis results are available upon request.

3.4. Robustness Checks

Several additional analyses are performed to test the robustness of the main finding. First, in addition to inflation forecasts over the next 12 months, respondents in the SCE are also asked to provide inflation forecasts over the 12-month period between 24 months and 36 months after

[6] In untabulated analyses, instead of using user fixed effects, I include individual characteristics as control variables, and find consistent results.

the survey date (i.e., inflation expectations over the 12 months during year 3 in the future). I use the three-year-ahead inflation forecasts, coupled with realized inflation over the same time frame, to re-examine the relation between learning and overconfidence (Column 1 of Table 5). Second, to ensure that the pattern documented in Table 4 is not restricted to individuals who completed all 12 surveys, I employ a more comprehensive sample that includes both respondents who completed all 12 surveys and those who did not (Column 2).[7] Third, to account for the possibility that survey participants may consider seasonality in their inflation expectation forecasts, the raw CPI index is replaced with a seasonally adjusted CPI index in calculating the hit rate (Column 3). Fourth, because it is possible that people's inflation expectations are driven by items that tend to be volatile, which may bias the documented pattern of sequential changes in overconfidence, I replace the CPI index with CPI for core goods, with food and energy excluded (Column 4). The results presented in Table 5 show that the baseline finding of a decrease in hit rate (i.e., an increase in overconfidence) with learning is robust to these additional analyses.

<Table 5 about here>

3.5. The Upper and Lower Bounds of the Confidence Interval

Thus far the empirical results have established that as respondents acquire more information about inflation from repeat surveys, it becomes less likely that their 50% confidence intervals contain realized future inflation. In this section, I take a closer look at how the lower and upper bounds of the confidence interval evolve with learning, and which side (or both) is responsible for the decline in the hit rate.

<Table 6 about here>

[7] Since late 2019, there has been a small number of survey respondents who participated in more than 12 surveys (up to 16 rounds). This group represents a small portion (about 3%) of the sample. Including the observations with more than 12 survey responses does not change the main finding.

Panel A Table 6 shows that over the 12 surveys, the 25th percentile forecasts (i.e., the lower bound of the 50% confidence interval) increase from 1.656% to 2.359% (Column 1), while the 75th percentile forecasts (i.e., the upper bound of the 50% confidence interval) decrease from 7.070% to 5.994% (Column 2). The difference in the forecasts between the first and the final rounds of the survey is statistically significant for both the lower and upper bounds. This pattern not only confirms the notion that respondents become increasingly certain as a result of learning (Kim and Binder, 2023), but it also suggests that individuals' 50% confidence levels shrink towards the middle instead of skewing in either direction. However, it is notable that the lower bound forecasts increase gradually, whereas the upper bound forecasts drop initially and stay relatively steady after the second round of the survey. The narrowing of the confidence interval is depicted in Figure 4a.

<Figure 4 about here>

I then examine whether realized inflation is above the lower bound forecast and below the upper bound, respectively. Note that if realized inflation lies above the lower bound forecast and below the upper bound at the same time, that means the true inflation outcome falls within the 50% confidence interval. In Panel B Column 1, the variable Lower Bound Hit is the percent chance that realized one-year-ahead inflation is above or equal to the 25th percentile forecast. In Column 2, the variable Upper Bound Hit is the percent chance that realized one-year-ahead inflation is below or equal to the 75th percentile forecast. On the lower bound side, there is a 57.8% chance that the lower bound contains realized inflation in the first round of the survey, and it decreases slightly to 56.0% in the last round. Meanwhile, there is a more noticeable decline in the probability that the upper bound forecast is set above realized inflation over the 12-survey period, from 80.3% to 74.8%.

Considering both sets of results in Panel B, the decrease in hit rate on both ends of the 50% confidence interval (about 1.8% on the lower end and 5.5% on the upper end) adds up to the total decrease in hit rate of 7.3% in the full sample. This trend is displayed in Figure 4b. Taken together, the evidence presented in Table 6 indicates that both the lower and upper bounds of the 50% confidence interval contribute to the reduction in forecast uncertainty, with the decrease of the upper bound playing a bigger role in explaining the decline in hit rate over time.

3.6. The Role of Initial Uncertainty and Initial Accuracy

In the empirical setting employed in this study, the presence of overconfidence ultimately depends on individuals' uncertainty and accuracy when making inflation forecasts. Put differently, the width of the 50% confidence interval and the location of forecast distribution dictate the hit rate. This feature conveniently enables an investigation of the role of individuals' intrinsic confidence (i.e., uncertainty) and skill (i.e., accuracy) in the outcome of learning. Since survey participants substantially reduce their level of uncertainty and improve their accuracy over the 12-survey period (Table 3), I focus on the *initial* uncertainty and *initial* accuracy exhibited in the first round of the survey to gauge each individual's inherent confidence and skill before learning begins.

<Table 7 about here>

Specifically, the sample is split into a high initial uncertainty group and a low initial uncertainty group based on the median value of individuals' initial uncertainty in the sample. As shown in Columns 1 and 2 of Table 7, there are stark differences in respondents' initial uncertainty levels: the average difference between the 25th percentile and the 75th percentile of future inflation forecasts is merely 1.753% in the low uncertainty group, compared to 9.047% in the high uncertainty group. In comparing the subsequent uncertainty changes in these two groups, I find that the decline in uncertainty reported for the full sample (Table 3) is only present for the individuals who are initially uncertain about future inflation. For this group, the 50% confidence interval decreases from roughly 9.047% to 4.921%. Interestingly, for individuals who initially have low uncertainty (i.e., inherently confident people), their level of uncertainty increases to 2.539% in the second survey, and mildly declines to 2.348% by the last round of the survey, which still stands above the initial uncertainty level. This evidence indicates that respondents who start out with a higher level of conviction tend to remain confident throughout the 12-survey period, although they do appear to make adjustments when their initial confidence intervals are tight. Figure 5a displays the patterns of uncertainty changes for both groups.

<Figure 5 about here>

Similarly, the sample is divided into a high initial accuracy group and a low initial accuracy group based on the median value of individuals' initial forecast error in the sample. Column 3 shows that individuals who are initially accurate experience a steady increase in forecast error (i.e., a decline in forecast accuracy), from 1.109% to 2.565%, over the course of 12 surveys. In contrast, the low accuracy group experiences a sharp decline in forecast error from 6.791% to 4.396%. The visual depiction of these patterns can be found in Figure 5b. As a result, the positive impact of learning on accuracy as documented in Table 3 is primarily driven by individuals who are initially quite inaccurate in their forecasts.

3.7. Joint Effect of Uncertainty and Accuracy on Overconfidence

To gain a more nuanced understanding of the two determinants of overconfidence, in this section, I take uncertainty and accuracy into account simultaneously and examine their joint effects on changes in overconfidence in the context of learning. Combining the subsamples discussed previously results in the following four groups of observations: 1. high uncertainty + high accuracy, 2. high uncertainty + low accuracy, 3. low uncertainty + high accuracy, and 4. low uncertainty +

low accuracy. The hit rate over the 12 waves of survey for each of these groups is presented in Table 8, and the findings are summarized in Figure 6.

<Table 8 about here>

Individuals with high initial uncertainty and accurate forecasts (i.e., uncertain but good forecasters), presented in Column 1, experience a dramatic drop in hit rate over the 12-survey period: the hit rate in this group plummet from 94.4% (these individuals actually start out "underconfident") to 41.1%. This group represents the most significant decline in hit rate. Unconfident individuals with low accuracy (Column 2) experience a small increase in hit rates from 25.6% to 28.8%. Respondents in this group, according to Table 7, are subject to competing forces: their uncertainty falls, and in the meantime, they become more accurate in their forecasts. Individuals with low uncertainty and highly accurate forecasts (Column 3) see a modest decrease in hit rate from 39.9% to 32.3%. Similar to Column 2, respondents in this group also face two competing trends: their 50% confidence intervals widen while their forecasts become less accurate. It appears that the reduction in accuracy dominates and leads to a modest decrease in hit rate in this scenario. Finally, individuals with low uncertainty and low accuracy (Column 4) have a 0% hit rates in the first round of survey – these respondents are both confident and wrong. However, their hit rate improves to 21.3% by the last round of survey, suggesting that as these individuals expand their 50% confidence levels and increase their forecasting accuracy, they are able to make partial adjustments to their overconfidence bias.

<Figure 6 about here>

3.8. Evidence from Home Price Expectations

In addition to future inflation expectations, the Survey of Consumer Expectations (SCE) data also provides individuals' expectations on one-year-ahead future home prices. Similar to the inflation expectation questions, home price forecast distributions are constructed based on the participants' responses to the question "*And in your view, what would you say is the percent chance that, over the next 12 months, the average home price nationwide will…*" Ten home price change intervals are available for respondents to assign their estimated probabilities, ranging from "decreased by 12% or more" to "increased by 12% or more." This information enables the investigation of the sequential change in individuals' overconfidence in a different setting, and thereby obtain a deeper understanding of the learning-overconfidence relation.

Two measures are used to capture the realized changes in national housing prices. The first is the Purchase Only House Price Index for the United States (HPI) made available by the U.S. Federal Housing Finance Agency.[8] This index is based on millions of repeat sales transactions on single-family properties. The second is the Zillow Home Value Index (ZHVI), which measures the typical home value in the mid-tier across the U.S. Both indices are commonly used to measure movement in real estate asset prices. Figure 7 shows that over the sample period covered in this study, one-year-ahead home price changes measured by HPI and ZHVI follow very similar trends. The average respondent forecast is low compared to realized home price changes, but shares the same stable pattern until a major divergence occurs during the COVID period.

<Figure 7 about here>

To examine overconfidence in this setting, the hit rate is constructed in the same way as the inflation-based hit rate described previously: it is a dummy variable equal to one if the realized one-year-ahead change in housing price, as measured by either HPI or ZHVI, falls within the 50% confidence interval of home price change forecasts, and zero otherwise. Table 9 presents the results. On average the housing price change hit rate is about 24% in both columns, much lower than the

[8] Data is downloaded from https://fred.stlouisfed.org/series/HPIPONM226N.

50% level had individuals been properly calibrated about their forecasts. In addition, there is a familiar decline in hit rate over the survey period as observed previously. Using the HPI measure (Column 1) as an example, survey participants' hit rate starts at 27.3%, and subsequently drops to 22.1% over the course of 11 consecutive surveys.[9] This pattern is consistent with the hit rate change pattern documented in the main analyses; that is, individuals appear to become increasingly overconfident as they learn through the survey. In untabulated regression analyses in which user fixed effects and year by state fixed effects are included, the trend of decreasing hit rates over the survey period continues to hold. Overall, the results presented in Table 9 offer additional support to the dynamic overconfidence reported in this paper.

<Table 9 about here>

4. Conclusions

In his best-selling book, *Thinking, Fast and Slow* (2011), Nobel Laureate Daniel Kahneman expressed his pessimism regarding humans' ability to overcome overconfidence by learning. Consistent with this view, several recent studies either underscore the limitation of learning or show that learning can even escalate overconfidence (Gervais and Odean, 2001; Sanchez and Dunning, 2018; Boutros, Ben-David, Graham, Harvey, and Payne, 2020). This study extends the existing literature by employing a nationally representative sample of consumers and their expectations on future inflation to empirically test the relation between learning and the evolution of overconfidence.

The evidence documented in this study supports the notion that people are on average overconfident: when forecasting future inflation, their 50% overconfidence intervals contain the realized inflation only 32% of the time. More importantly, leveraging the unique feature of repeat

[9] Because this home price expectation question is seen by repeat respondents only, its responses become available from the second round of the survey.

forecasts in the SCE data, I am able to track the change in respondents' overconfidence, and find that as individuals acquire information with respect to inflation (Kim and Binder, 2023), they become increasingly overconfident during the course of the survey – their 50% confidence intervals are less likely to contain realized future inflation over time. This effect is present across different demographic groups, and is most pronounced in the first few rounds of the survey.

Respondents' initial forecast uncertainty and accuracy play an important role: initially uncertain participants significantly shrink their forecast uncertainty, while the initially inaccurate ones improve their forecast accuracy over time. Those who are initially certain or accurate, on the other hand, experience a modest increase in uncertainty or forecast error. The impact of learning on overconfidence is most salient in the group of respondents who are initially uncertain and accurate; these individuals dramatically narrow their confidence intervals, whereas their forecast accuracy deteriorates at the same time.

Using individuals' home price forecast data to re-examine the relation between learning and overconfidence, I find evidence that confirms individuals' overconfidence bias. Specifically, realized home price changes fall within respondents' 50% confidence intervals only about 24% of the time. Importantly, the pattern that overconfidence increases with learning continues to be present in this setting, which lends additional support to the main finding based on inflation forecast data.

This paper opens up different avenues for future research. For example, the psychology literature suggests that overconfidence can be manifested in three distinct ways, namely overestimation, overplacement, and overprecision (Moore and Healy, 2008). The empirical setting adopted in this study focuses on the overprecision type of overconfidence. It would be interesting to examine whether and how learning may affect overestimation and overplacement. Moreover, the 12-survey period in this study may not be sufficient to draw meaningful long-term inferences. More research is needed to gain a deeper understanding of the long-term implications of learning on overconfidence. It is my hope that researchers and policymakers will take the dynamic nature of overconfidence into account when designing tools to mitigate human behavioral biases.

References:

Anderson, Cameron, Sebastien Brion, Don A. Moore, and Jessica A. Kennedy, 2012, A status-enhancement account of overconfidence, *Journal of Personality and Social Psychology* 103, 718-735.

Armantier, Olivier, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar, 2017, An overview of the survey of consumer expectations, *Economic Policy Review* 51-72.

Ben-David, Itzhak, John R. Graham, and Campbell R. Harvey, 2013, Managerial miscalibration, *The Quarterly Journal of Economics* 128, 1547-1584.

Berner, Eta S., and Mark L. Graber, 2008, Overconfidence as a cause of diagnostic error in medicine, *The American Journal of Medicine* 121, S2-S23.

Bhandari, Gokul, and Richard Deaves, 2006, The demographics of overconfidence, *Journal of Behavioral Finance* 7, 5-11.

Boutros, Michael, Itzhak Ben-David, John R Graham, Campbell R Harvey, and John W Payne, 2020, The persistence of miscalibration, *Working Paper, National Bureau of Economic Research*.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, and Michael Weber, 2023, Forward guidance and household expectations, *Journal of the European Economic Association* 21, 2131-2171.

Coibion, Olivier, Yuriy Gorodnichenko, and Michael Weber, 2022, Monetary policy communications and their effects on household inflation expectations, *Journal of Political Economy* 130, 1537-1584.

Craig, Paul, 2013. *The killing zone: How and why pilots die* (McGraw-Hill Professional).

D'Acunto, Francesco, Daniel Hoang, Maritta Paloviita, and Michael Weber, 2023, IQ, expectations, and choice, *The Review of Economic Studies* 90, 2292-2325.

D'Acunto, Francesco, Ulrike Malmendier, Juan Ospina, and Michael Weber, 2020, Exposure to grocery prices and inflation expectations, *Journal of Political Economy* 129, 1615-1639.

Gervais, Simon, and Terrance Odean, 2001, Learning to be overconfident, *The Review of Financial Studies* 14, 1-27.

Hirshleifer, David, Angie Low, and Siew Hong Teoh, 2012, Are overconfident CEOs better innovators?, *The Journal of Finance* 67, 1457-1498.

Johnson, Dominic DP, 2009. *Overconfidence and war: The havoc and glory of positive illusions* (Harvard University Press).

Kahneman, Daniel, 2011. *Thinking, fast and slow* (Farrar, Straus and Giroux).

Kahneman, Daniel, Olivier Sibony, and Cass R Sunstein, 2021. *Noise: A flaw in human judgment* (Hachette UK).

Kim, Gwangmin, and Carola Binder, 2023, Learning-through-survey in inflation expectations, *American Economic Journal: Macroeconomics* 15, 254-78.

Koellinger, Philipp, Maria Minniti, and Christian Schade, 2007, "I think i can, i think i can": Overconfidence and entrepreneurial behavior, *Journal of Economic Psychology* 28, 502-527.

Malmendier, Ulrike, and Geoffrey Tate, 2005, CEO overconfidence and corporate investment, *The Journal of Finance* 60, 2661-2700.

Moore, Don A., and Paul J. Healy, 2008, The trouble with overconfidence, *Psychological Review* 115, 502-517.

Odean, Terrance, 1999, Do investors trade too much?, *American Economic Review* 89, 1279-1298.

Sanchez, Carmen, and David Dunning, 2018, Overconfidence among beginners: Is a little learning a dangerous thing?, *Journal of Personality and Social Psychology* 114, 10-28.

Soll, Jack B., and Joshua Klayman, 2004, Overconfidence in interval estimates, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 30, 299-314.

Table 1. Summary Statistics

Panel A reports summary statistics for one-year-ahead inflation expectation measures. Panel B reports summary statistics for survey respondents' personal characteristics. Detailed variable definitions are provided in Appendix 1.

Panel A: Expectation Measures

	N	Mean	Median	25th Percentile	75th Percentile	STD
Inf_Exp_Mean	74,151	4.111	3.000	1.530	5.666	5.182
Inf_Exp_P25	74,151	2.058	2.000	0.500	3.513	5.083
Inf_Exp_P75	74,151	6.136	4.116	2.852	7.496	6.190
Inf_Exp_Uncertain	74,151	4.078	2.279	1.188	4.752	4.490

Panel B: Survey Participant Characteristics

	N	Mean	Median	25th Percentile	75th Percentile	STD
Female	75,432	0.458	0.000	0.000	1.000	0.498
Young	75,444	0.253	0.000	0.000	1.000	0.435
White	75,444	0.856	1.000	1.000	1.000	0.351
Rich	75,444	0.284	0.000	0.000	1.000	0.451
College	75,444	0.573	1.000	0.000	1.000	0.495

Table 2: Are SCE Respondents Overconfident?

This table reports the hit rate for the full sample (Panel A) and for sub-groups based on survey respondents' personal characteristics (Panel B). Detailed variable definitions are provided in Appendix 1.

Panel A: Full Sample Hit and Miss Rate

N	Miss_P25	Hit	Miss_P75
74,151	0.439	0.321	0.240

Panel B: Hit Rate by Personal Characteristic

Female		Male	
N	Hit	N	Hit
33,755	0.327	40,384	0.315

Young		Old	
N	Hit	N	Hit
18,930	0.369	55,221	0.304

White		Non-White	
N	Hit	N	Hit
63,592	0.317	10,559	0.345

Rich		Poor	
N	Hit	N	Hit
21,293	0.319	52,858	0.321

College		Non-College	
N	Hit	N	Hit
42,800	0.321	31,351	0.321

Table 3: Evidence of Learning

This table reports forecast uncertainty (Column 1) and forecast error (Column 2) by survey order. Detailed variable definitions are provided in Appendix 1.

	(1) Forecast Uncertainty	(2) Forecast Error
Survey 1	5.414	3.936
Survey 2	4.752	3.673
Survey 3	4.349	3.631
Survey 4	4.137	3.631
Survey 5	3.967	3.553
Survey 6	3.912	3.609
Survey 7	3.881	3.609
Survey 8	3.786	3.558
Survey 9	3.752	3.558
Survey 10	3.703	3.505
Survey 11	3.721	3.567
Survey 12	3.635	3.481
Mean	4.078	3.608

Table 4: Learning and Hit Rate

This table reports the relation between learning and overconfidence. Panel A displays the hit rate by survey order. Panel B shows the estimates of linear probability regressions of whether a respondent's 50% confidence interval includes realized future inflation on indicators of survey order. User and year-by-state fixed effects are included in all regressions. Standard errors are clustered at the respondent level, and are displayed in parentheses. Detailed variable definitions are provided in Appendix 1. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Hit Rate by Survey Order

	(1) Full Sample	(2) Female	(3) Male	(4) Young	(5) Old	(6) White	(7) Non-White	(8) Rich	(9) Poor	(10) College	(11) Non-College
Survey 1	0.381	0.399	0.367	0.437	0.361	0.373	0.433	0.365	0.388	0.370	0.397
Survey 2	0.360	0.375	0.348	0.404	0.345	0.352	0.410	0.343	0.367	0.351	0.374
Survey 3	0.341	0.347	0.336	0.388	0.325	0.335	0.379	0.330	0.345	0.332	0.353
Survey 4	0.327	0.340	0.316	0.363	0.314	0.323	0.346	0.330	0.325	0.326	0.328
Survey 5	0.314	0.321	0.309	0.349	0.303	0.310	0.341	0.304	0.319	0.313	0.316
Survey 6	0.305	0.315	0.296	0.354	0.288	0.301	0.327	0.303	0.306	0.309	0.299
Survey 7	0.305	0.313	0.299	0.357	0.288	0.301	0.329	0.308	0.304	0.309	0.300
Survey 8	0.304	0.308	0.301	0.339	0.293	0.303	0.313	0.310	0.302	0.308	0.299
Survey 9	0.297	0.299	0.295	0.342	0.281	0.293	0.316	0.299	0.295	0.302	0.290
Survey 10	0.310	0.308	0.311	0.366	0.291	0.305	0.336	0.315	0.308	0.315	0.302
Survey 11	0.298	0.302	0.295	0.355	0.279	0.298	0.303	0.312	0.293	0.299	0.297
Survey 12	0.308	0.301	0.314	0.371	0.286	0.306	0.316	0.309	0.307	0.315	0.298

Panel B: Hit Rate by Survey Order Regressions

	(1) Full Sample	(2) Female	(3) Male	(4) Young	(5) Old	(6) White	(7) Non-White	(8) Rich	(9) Poor	(10) College	(11) Non-College
Survey 2	-0.019***	-0.021*	-0.016*	-0.032**	-0.014	-0.018**	-0.026	-0.022	-0.018**	-0.014	-0.026**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Survey 3	-0.038***	-0.050***	-0.029***	-0.044***	-0.035***	-0.035***	-0.054***	-0.035**	-0.039***	-0.033***	-0.046***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Survey 4	-0.052***	-0.056***	-0.048***	-0.070***	-0.045***	-0.046***	-0.087***	-0.033**	-0.059***	-0.039***	-0.072***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Survey 5	-0.064***	-0.073***	-0.056***	-0.086***	-0.055***	-0.059***	-0.091***	-0.059***	-0.066***	-0.052***	-0.084***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 6	-0.074***	-0.078***	-0.070***	-0.081***	-0.069***	-0.067***	-0.108***	-0.060***	-0.080***	-0.056***	-0.101***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 7	-0.074***	-0.082***	-0.066***	-0.082***	-0.070***	-0.067***	-0.108***	-0.056***	-0.081***	-0.055***	-0.102***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 8	-0.075***	-0.085***	-0.065***	-0.102***	-0.065***	-0.065***	-0.128***	-0.055***	-0.083***	-0.057***	-0.101***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 9	-0.085***	-0.097***	-0.075***	-0.103***	-0.078***	-0.077***	-0.131***	-0.070***	-0.090***	-0.067***	-0.112***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 10	-0.074***	-0.090***	-0.061***	-0.082***	-0.071***	-0.066***	-0.118***	-0.058***	-0.080***	-0.054***	-0.103***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 11	-0.086***	-0.097***	-0.078***	-0.093***	-0.084***	-0.074***	-0.155***	-0.065***	-0.095***	-0.072***	-0.108***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Survey 12	-0.079***	-0.101***	-0.061***	-0.082***	-0.078***	-0.068***	-0.146***	-0.069***	-0.083***	-0.059***	-0.109***
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.01)
User FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year x State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R-Squared	0.324	0.345	0.311	0.349	0.314	0.318	0.362	0.277	0.346	0.296	0.370

Table 5: Robustness Checks

This table reports robustness check results. Column 1 uses three-year-ahead inflation forecasts to estimate overconfidence. Column (2) includes respondents who completed 12 consecutive survey and those who did not. Column (3) is based on future inflation calculated using seasonally adjusted CPI. Column (4) employs CPI for core goods, with food and energy excluded. Detailed variable definitions are provided in Appendix 1.

	(1) Three-Year-Ahead Forecast	(2) Up To 12 Surveys	(3) Seasonally Adjusted	(4) Core Goods
Survey 1	0.373	0.395	0.379	0.425
Survey 2	0.352	0.373	0.359	0.409
Survey 3	0.348	0.355	0.340	0.397
Survey 4	0.325	0.344	0.328	0.376
Survey 5	0.331	0.326	0.313	0.368
Survey 6	0.320	0.316	0.305	0.360
Survey 7	0.314	0.315	0.306	0.364
Survey 8	0.306	0.312	0.303	0.353
Survey 9	0.306	0.311	0.298	0.348
Survey 10	0.308	0.316	0.312	0.356
Survey 11	0.298	0.300	0.299	0.349
Survey 12	0.301	0.300	0.308	0.352

Table 6: The Upper and Lower Bounds of the Confidence Interval

Panel A reports the development of the lower bound (25th percentile) and the upper bound (75th percentile) one-year-ahead forecasts over the 12-survey period. Panel B shows the lower bound and upper bound hit rates over the 12-survey period. Detailed variable definitions are provided in the Appendix 1.

Panel A: Lower & Upper Bound by Survey Order

	(1) Lower Bound One-Year-Ahead Forecast	(2) Upper Bound One-Year-Ahead Forecast
Survey 1	1.656	7.070
Survey 2	1.559	6.311
Survey 3	1.663	6.012
Survey 4	1.877	6.014
Survey 5	2.003	5.970
Survey 6	2.148	6.060
Survey 7	2.225	6.105
Survey 8	2.237	6.023
Survey 9	2.286	6.038
Survey 10	2.281	5.985
Survey 11	2.370	6.092
Survey 12	2.359	5.994

Panel B: Lower & Upper Bound Hit Rate by Survey Order

	(1) Lower Bound Hit	(2) Upper Bound Hit
Survey 1	0.578	0.803
Survey 2	0.584	0.777
Survey 3	0.579	0.762
Survey 4	0.573	0.754
Survey 5	0.560	0.755
Survey 6	0.548	0.757
Survey 7	0.545	0.761
Survey 8	0.546	0.758
Survey 9	0.543	0.753
Survey 10	0.558	0.752
Survey 11	0.555	0.743
Survey 12	0.560	0.748

Table 7: The Role of Initial Uncertainty and Initial Accuracy

This table reports the change in forecast uncertainty over the 12-survey period for individuals with high initial uncertainty (Column 1) and with low initial uncertainty (Column 2), and the change in forecast error over the 12-survey period for individuals with high initial accuracy (Column 3) and with low initial accuracy (Column 4). Detailed variable definitions are provided in the Appendix 1.

	(1)	(2)	(3)	(4)
	Forecast Uncertainty		Forecast Error	
	High Initial Uncertainty	Low Initial Uncertainty	High Initial Accuracy	Low Initial Accuracy
Survey 1	9.047	1.753	1.109	6.791
Survey 2	6.968	2.539	2.099	5.258
Survey 3	6.227	2.470	2.243	5.021
Survey 4	5.862	2.412	2.315	4.945
Survey 5	5.580	2.356	2.324	4.782
Survey 6	5.446	2.384	2.390	4.822
Survey 7	5.383	2.378	2.454	4.766
Survey 8	5.223	2.354	2.444	4.672
Survey 9	5.183	2.322	2.493	4.621
Survey 10	5.075	2.335	2.541	4.467
Survey 11	5.070	2.371	2.622	4.511
Survey 12	4.921	2.348	2.565	4.396

Table 8: Joint Effect of Uncertainty and Accuracy on Overconfidence

This table reports the evolution of the hit rate over the 12-survey period for individuals with high initial uncertainty and high initial accuracy (Column 1), with high initial uncertainty and low initial accuracy (Column 2), with low initial uncertainty and high initial accuracy (Column 3), and with low initial uncertainty and low initial accuracy (Column 4). Detailed variable definitions are provided in the Appendix 1.

	(1) High Uncertainty + High Accuracy	(2) High Uncertainty + Low Accuracy	(3) Low Uncertainty + High Accuracy	(4) Low Uncertainty + Low Accuracy
Survey 1	0.944	0.256	0.399	0.000
Survey 2	0.595	0.329	0.381	0.142
Survey 3	0.544	0.317	0.364	0.139
Survey 4	0.515	0.298	0.350	0.149
Survey 5	0.453	0.298	0.350	0.141
Survey 6	0.447	0.287	0.325	0.160
Survey 7	0.462	0.276	0.329	0.159
Survey 8	0.431	0.292	0.319	0.174
Survey 9	0.415	0.281	0.319	0.168
Survey 10	0.433	0.307	0.320	0.175
Survey 11	0.415	0.291	0.297	0.198
Survey 12	0.411	0.288	0.323	0.213

Table 9: Evidence from Home Price Changes

This table reports the relation between learning and overconfidence using data on home price change forecasts. The hit rate in Column 1 is calculated based on Purchase Only House Price Index for the United States (HPI). The hit rate in Column 2 is calculated based on the Zillow Home Value Index (ZHVI). Detailed variable definitions are provided in the Appendix 1.

	(1) HPI Hit	(2) ZHVI Hit
Survey 2	0.273	0.271
Survey 3	0.269	0.257
Survey 4	0.255	0.250
Survey 5	0.243	0.236
Survey 6	0.242	0.234
Survey 7	0.244	0.234
Survey 8	0.236	0.231
Survey 9	0.236	0.230
Survey 10	0.233	0.233
Survey 11	0.229	0.228
Survey 12	0.221	0.213
Mean	0.244	0.238

Appendix 1. Variable Definitions

Variable	Definition
Inf_Exp_Mean	The density mean of one-year-ahead inflation forecast distribution.
Inf_Exp_P25	The 25th percentile one-year-ahead inflation forecast.
Inf_Exp_P75	The 75th percentile one-year-ahead inflation forecast.
Inf_Exp_Uncertain	The difference between the 75th percentile and the 75th percentile one-year-ahead inflation forecasts.
Female	Dummy variable equal to one if the survey respondent is female, and zero otherwise.
Young	Dummy variable equal to one if the survey respondent is under 40-years old, and zero otherwise.
White	Dummy variable equal to one if the survey respondent is white, and zero otherwise.
Rich	Dummy variable equal to one if the survey respondent's household income is over $100,000, and zero otherwise.
College	Dummy variable equal to one if the survey respondent has a college degree, and zero otherwise.
Hit	The percent chance that realized one-year-ahead inflation falls within the 50% confidence interval.
Miss_P25	The percent chance that realized one-year-ahead inflation falls below the 25th percentile inflation forecast.
Miss_P75	The percent chance that realized one-year-ahead inflation exceeds the 75th percentile inflation forecast.
Forecast Uncertainty	The difference between the 75th percentile and the 75th percentile one-year-ahead inflation forecasts.
Forecast Error	The absolute difference between the density mean of each individual's inflation expectation distribution and realized future inflation.
Lower Bound Hit	The percent chance that realized one-year-ahead inflation is above or equal to the 25th percentile forecast.
Upper Bound Hit	The percent chance that realized one-year-ahead inflation is below or equal to the 75th percentile forecast.
HPI Hit	Hit rate calculated based on the Purchase Only House Price Index for the United States.
ZHVI Hit	Hit rate calculated based on the Zillow Home Value Index.

Appendix 2. SCE Questionnaire Example

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 (bin 1)	_____ percent chance
the rate of inflation will be between 8% and 12% (bin 2)	_____ percent chance
the rate of inflation will be between 4% and 8% (bin 3)	_____ percent chance
the rate of inflation will be between 2% and 4% (bin 4)	_____ percent chance
the rate of inflation will be between 0% and 2% (bin 5)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 0% and 2% (bin 6)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 2% and 4% (bin 7)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 4% and 8% (bin 8)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 8% and 12% (bin 9)	_____ percent chance
the rate of deflation (opposite of inflation) will be 12% or higher (bin 10)	_____ percent chance
TOTAL	**100**

If no response: error E1
If sum not equal to 100: "Your total adds up to XX" followed by error msg E3.

Figure 1: Inflation Expectation vs. Realization
This figure shows the mean one-year-ahead inflation expectations and realized one-year-ahead inflation over the sample period.

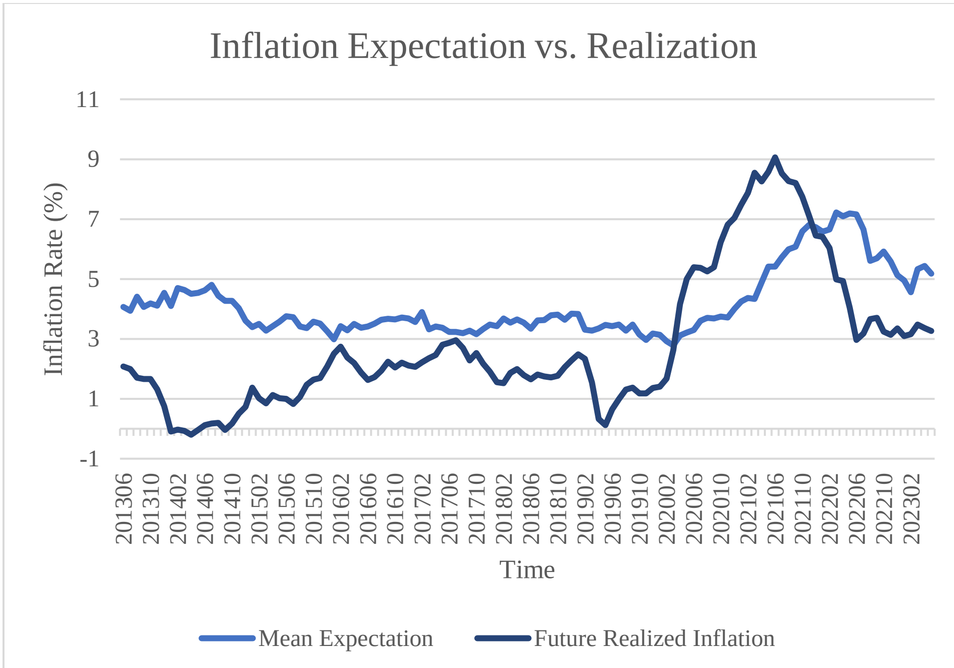

Figure 2: Evidence of Learning

This figure shows forecast uncertainty and forecast error by survey order.

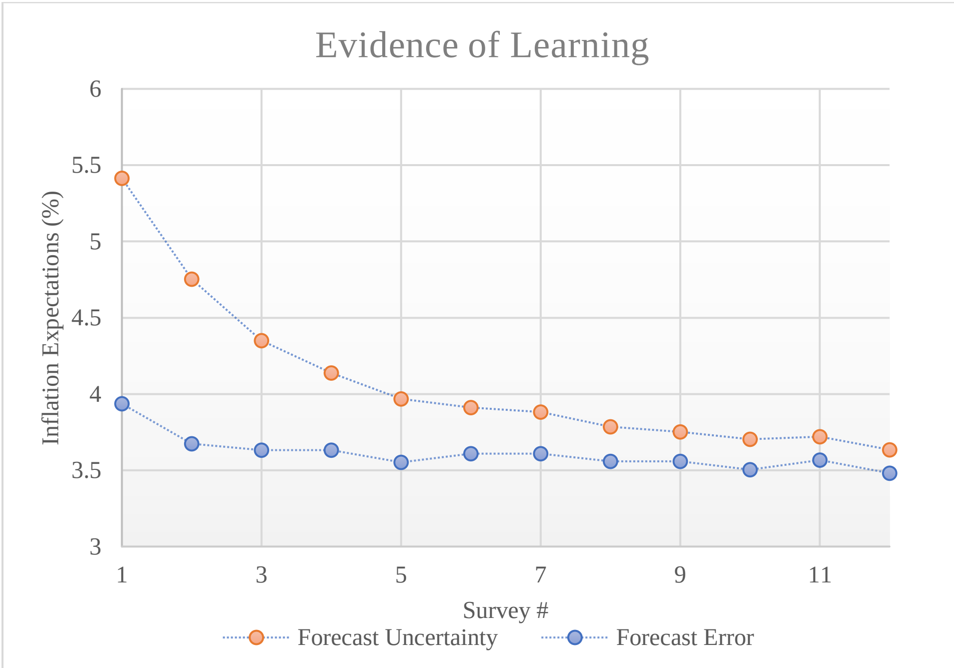

Figure 3: Learning and Hit Rate
This figure shows hit rate by survey order.

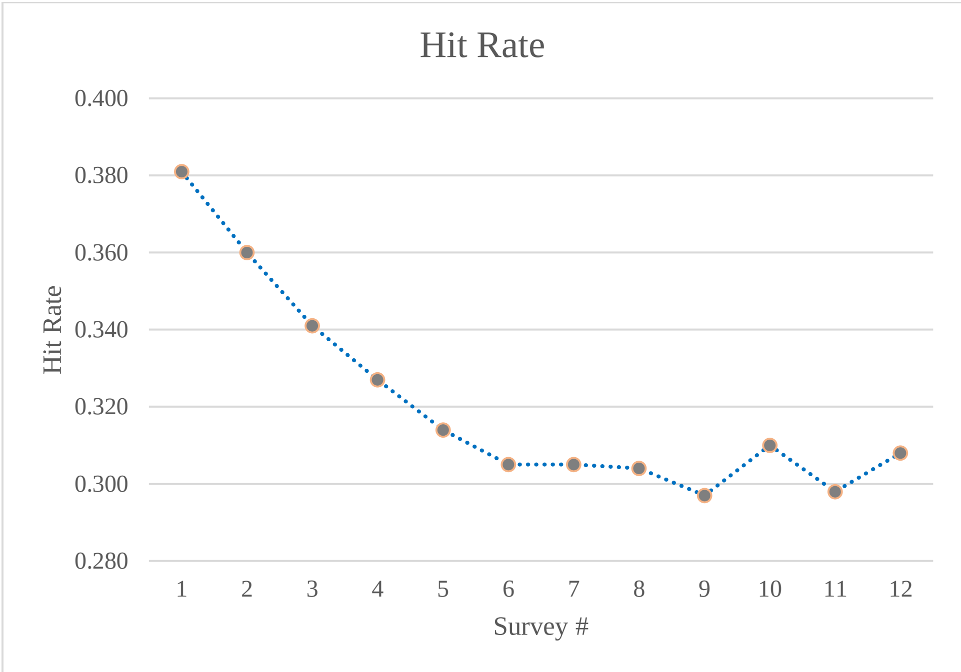

Figure 4: The Upper and Lower Bounds of the Confidence Interval
Figure 4a shows the development of the lower bound (25th percentile) and the upper bound (75th percentile) one-year-ahead forecasts by survey order. Panel B shows the lower bound and upper bound hit rates by survey order.

Figure 4a.

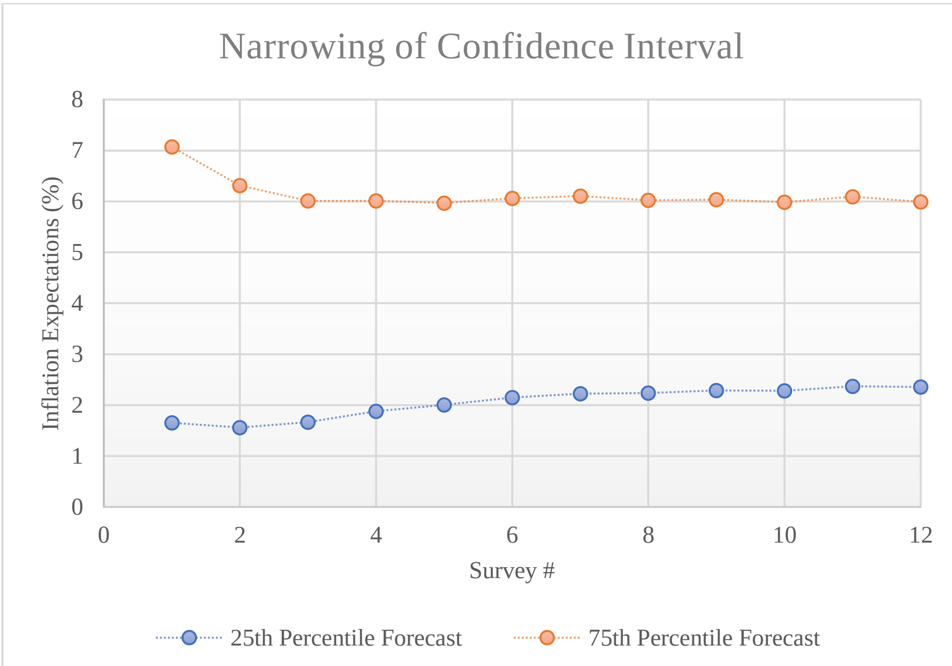

Figure 4b.

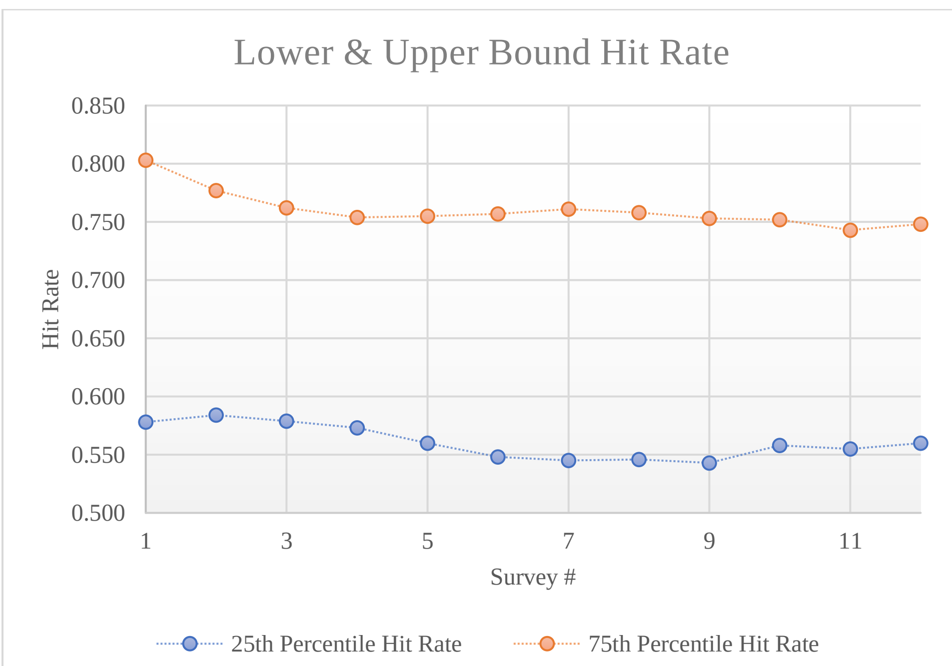

Figure 5. Role of Initial Uncertainty and Initial Accuracy
Figure 5a shows the change in forecast uncertainty by survey order for individuals with high initial uncertainty and with low initial uncertainty. Figure 5b shows the change in forecast error by survey order for individuals with high initial accuracy and with low initial accuracy.

Figure 5a.

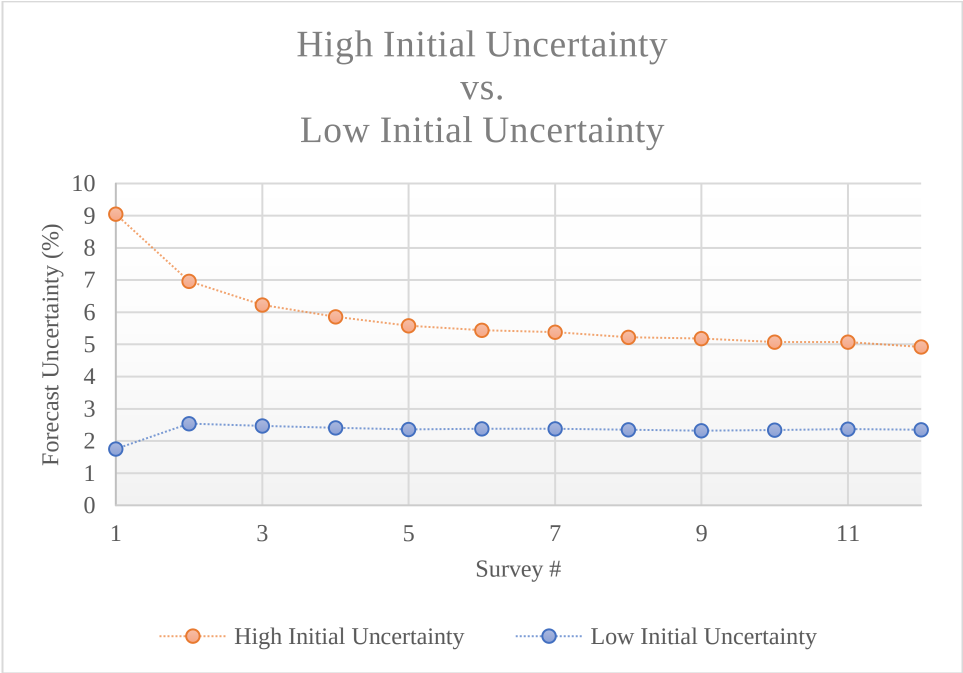

Figure 5b.

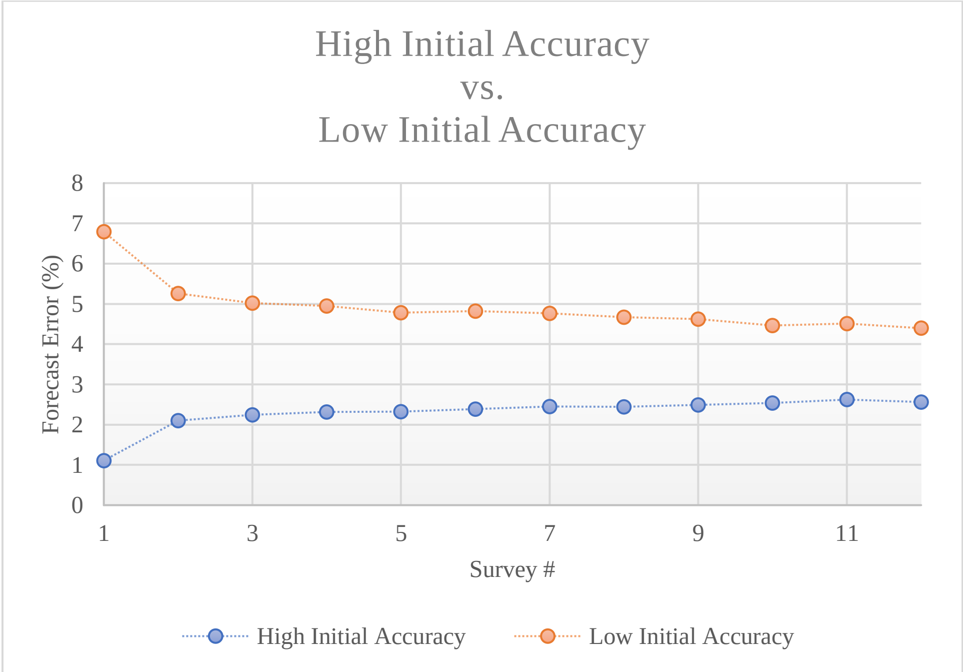

Figure 6. Joint Effect of Uncertainty and Accuracy on Overconfidence
This figure shows hit rate by survey order for individuals with high initial uncertainty and high initial accuracy, with high initial uncertainty and low initial accuracy, with low initial uncertainty and high initial accuracy, and with low initial uncertainty and low initial accuracy.

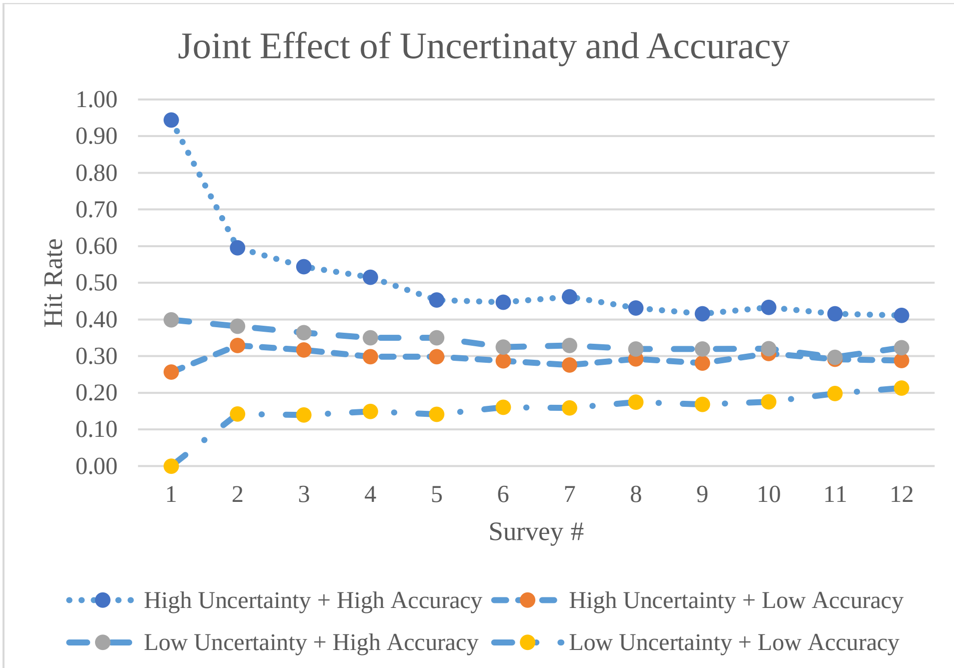

Figure 7. Home Price Expectation vs. Realization
This figure shows the mean one-year-ahead home price change expectations and realized one-year-ahead home price change (measured based on HPI and ZHVI) over the sample period.