Demand for Macroeconomic Data
Evidence from an Information-Acquisition Experiment∗

Andreas Fuster
Federal Reserve Bank of New York

Ricardo Perez-Truglia
University of California, Los Angeles

Basit Zafar
Arizona State University

This Draft: December 2017.
Preliminary. Please do not circulate.

Abstract
Information frictions—either in the form of costly information acquisition or constraints in information processing—are believed to play an important role in consumers’ formation of macroeconomic expectations. However, there is little direct empirical evidence about the causes and consequences of information acquisition. We conduct a survey experiment to study this question in the context of house price expectations. We let consumers buy different pieces of information that could be relevant for the formation of their expectation about the future median national home price. For this purpose, we use an incentive-compatible mechanism to elicit their maximum willingness to pay. In addition, we introduce exogenous variation in the value of information by randomly assigning individuals to rewards for the ex-post accuracy of their expectations. This setup generates several testable hypotheses. Consistent with rational inattention, individuals are willing to pay more for information when they stand to gain more from it. However, underscoring the importance of limits on information processing capacity, individuals disagree on which signal they prefer to buy. Less sophisticated individuals—those with lower education and financial numeracy—are less likely to demand information that has ex-ante higher predictive power, and this gap is not ameliorated when the stakes are higher. As a result, and in contrast to models of optimal information processing, random information provision or lowering of search costs does not decrease the cross-sectional dispersion of expectations. Our findings have implications for models of expectation information and for the design of information interventions.

JEL Classification: C81, C93, D80, D83, D84, E31, E58.

Keywords: expectations, survey, experiment, prices, housing, information frictions, rational inattention.

∗Fuster: andreas.fuster@ny.frb.org. Perez-Truglia: ricardo.truglia@anderson.ucla.edu. Zafar: basitak@gmail.com. We would like to thank Filip Matejka and seminar attendants at the Federal Reserve Bank of New York for useful comments. Felipe Montano Campos provided superb research assistance. The UCLA Ziman Center for Real Estate provided funding for this research. The views presented here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York, or the Federal Reserve System.
1 Introduction

Given that expectations play a major role in any decision made under uncertainty, there is a big literature in macroeconomics that tries to understand consumers’ expectation formation process. Studies have found considerable dispersion in consumers’ expectations (for, example, Mankiw, Reis, and Wolfers, 2003). The literature has tried to explain this as a result of rational inattention, that may arise due to costs of acquiring information (as in the sticky information models, by Mankiw and Reis (2002) and Reis (2006)), or constraints on individuals’ information processing capacity (as in Sims (2003) and Woodford (2003)). However, there is little direct empirical micro evidence on how individuals acquire information in the real world and process it.\footnote{In a recent survey article, Gabaix (2017), makes the case for more experimental evidence on the determinants of attention, and the consequences of inattention.} In this paper, we present a survey experiment to study the causes and consequences of these information acquisition and processing decisions.

We study information acquisition in the context of expectations about the evolution of national home prices. Our interest in home prices stems from the fact that home price expectations play an prominent role in many accounts of the housing boom that occurred during the mid-2000s in the US (e.g. Shiller, 2005; Glaeser and Nathanson, 2015). In addition, Armona et al. (2016) and Bailey et al. (2017) show that home price expectations tend to affect (intended and actual) housing-related behavior (such as buying a home, or making investments in a home). We designed a survey experiment, as part of the NY Fed’s 2017 SCE Housing Survey, an online annual survey on housing issues, conducted as part of the monthly Survey of Consumer Expectations (SCE).

The experiment has three main stages. In the first stage, respondents report their expectations for the national median home price for the end of 2017 (their “prior” belief). In the second stage, which comes much later in the survey, respondents are informed that their home price expectation would be re-elicited (the “posterior”), but that this time it would be incentivized: if their expectation ends up falling within 1% of the actual price, then they will be eligible to a cash reward. To allow for a sharp test of rational inattention, we create exogenous variation in the reward that respondents are offered at this stage: half of the subjects are randomly assigned to a reward that pays $100 with probability 10%, and the other half is assigned to a reward that pays $10 with probability 10%. Before respondents report their posterior belief that will determine if they get the reward, they are given the opportunity to see information that could potentially be useful for their forecast. They are told that they will be able to see a piece of information, which they have to choose from the following options: the average year-end home price forecast of a panel of housing experts; the home price change over the past year; the home price change over the past ten years; or no information at all. In the third and final stage, we use an incentive-compatible mechanism to elicit the maximum willingness to pay (WTP) for the information source that they preferred the most in the second stage. For this purpose, we use the multiple-price-list variation of the
Becker–DeGroot–Marschak method. Individuals must make choices in eleven scenarios, choosing between the information or a pay-off (that varies between $0.01 and $5) in each. A scenario is then randomly chosen, and the corresponding choice is implemented. This price elicitation method creates, indirectly, an information provision experiment: conditional on one’s WTP, whether the individual is shown information or not is determined by chance, through the random assignment to one of the eleven scenarios. The experiment concludes with the elicitation of posterior beliefs.

Our experimental set-up allows us to test features of several main models used in the literature. In sticky information models, such as Mankiw and Reis (2002) and Reis (2006), agents update their information sets infrequently due to costs. However, once they update their information set they process all the information and use it optimally to form expectations. The constraint in these models is primarily on the margin of updating the information set. Randomizing the incentive of the accuracy of the forecast in the second stage, and the variation in the price of information in the third stage, allows us to test for the role of information acquisition costs directly. These models contrast with noisy information models (as in Sims (2003), Woodford (2003), and Mackowiak and Wiederholt (2009)), where even if information is freely available and agents are updating their information continuously, agents may not use all of it because of limited information processing capacity. Thus, here the constraints are primarily on the margin of processing information. It is to test these models that we allow our respondents to choose between different information sources. The quasi-random variation in the provision of information then allows us to investigate how individuals learn from their favorite piece of information.

The three information sources from which the individuals have to choose contain different signals: the average experts’ forecast for home price change during 2017 was 3.6%, national home prices had increased by 6.8% during the last one year, and home prices had decreased by 0.9% in total over the last ten years. Under a given definition of informativeness, we can rank the informativeness of these three information sources. One reasonable criterion, although certainly not the only one, is their ex-ante predictive power in the evolution of US home prices during the years leading to the survey. Based on this criterion, experts’ forecast is the most informative, followed by past one-year changes, and then ten-year changes. This ranking is consistent with basic intuitions from the real estate literature. For instance, the fact that past one-year price changes are ranked higher than ten-year changes is consistent with the well-documented momentum in home prices over short horizons (Case and Shiller, 1989; Guren, 2016; Armona, Fuster, and Zafar, 2017).

Our first result, with regards to the preference over information sources, is that individuals disagree on which of the three pieces of information they want to buy: 45% chose forecasts of housing experts, 28% chose the last year home price change, and 22% chose the last ten year home price change (the remaining 5% reported to prefer no information at all). Thus, less than half of the sample chooses the source that, according to the the ex-ante predictive power, is most informative. It could certainly be the case that some of this heterogeneity in ranking of sources is
due to respondents using other criteria. However, that fact that more sophisticated respondents, as measured by education or numeracy, are substantially more likely to choose the expert forecast suggests that at least part of the variation is due to cognitive limitations in identifying more informative signals.

We find that individuals have significant willingness to pay for their favorite information: in the low-reward condition ($10 with 10% probability), the median individual is willing to pay $4.05 for her favorite information piece. This suggests that individuals expect to benefit from this information beyond the accuracy rewards provided in the survey. Furthermore, we find strong support for the rational inattention hypothesis: the median WTP is significantly higher in the $100-reward condition than in the $10-reward condition ($4.05 and $4.78, respectively). This difference is not only statistically significant (p-value<0.01), but also economically meaningful. However, the ranking of information sources does not vary by the size of the reward. That is, respondents do not seem to think more carefully about the usefulness of the potential information sources when the stakes are higher. This suggests constraints in consumers’ ability to decipher informative signals.

Our third result exploits the information-provision experiment, and deals with whether individuals learn from the information provided to them. Consistent with a genuine interest in the information, individuals use the information that they were willing to pay for in their expectation formation. The evidence suggests that individuals form posterior beliefs by putting 44.5% weight on the signal bought, and 55.5% on their prior belief. We conducted a follow-up survey four months after the baseline survey, and find a persistent effect of the information on beliefs. Also, as expected, we find that the rate of learning from the information is similar across all three information sources, providing further confirmation that the disagreement about the information ranking was meaningful. However, here again we see violations from fully rational behavior, in the sense that we do not find evidence of individuals who have more uncertain prior beliefs or individuals who pay more for the information putting more weight on the purchased information.

Our final result is about the evolution of cross-sectional dispersion in expectations. If individuals only face costs of acquiring information but process information optimally (as in some sticky information models), the dispersion in beliefs should go down when information acquisition is subsidized (be it through an increase in the incentive for the forecast, or lowering of the price of information). Exploiting the random provision of information, we explore the effects of information on dispersion of expectations. We show that being randomly assigned to a lower cost of information acquisition (or a higher reward for accuracy) does not cause lower cross-sectional dispersion in expectations.2 If individuals acquire more information when the cost is lower, how is it possible

---

2This finding has some parallels with the literature on media bias and polarization in political attitudes. Individuals may have a propensity to discount information if it is inconsistent with a prior belief. In that case, information may in fact lead to more dispersed and polarized beliefs (Lord, Ross, and Lepper, 1979; Glaeser, 2004; Mullainathan and Shleifer, 2005). This may also arise if individuals sort into information sources and there is strong media slant (Martin and Yurukoglu, 2017). Our setting differs from this literature in the sense that individuals may
that more information does not induce higher consensus? To understand the mechanisms behind this result, we can divide respondents into three groups, based on the information source that they prefer. On the one hand, exposure to information tends to reduce the dispersion in beliefs within each of these groups: for example, among individuals who prefer the expert forecast, exposure to it results in their posterior beliefs becoming more compressed around the signal of 3.6%. On the other hand, exposure to information tends to increase the dispersion in beliefs across these three groups. This is because each group acquires a different signal, and the different groups move in different directions. These two channels tend to cancel each other out: as a result, information acquisition does not lead to a decrease in the dispersion in beliefs.

Our results thus lend support to models of expectation formation in which individuals form expectations subject to information processing constraints. Importantly, our findings imply that even if the acquisition cost of information were zero, we would still observe substantial dispersion in consumers’ expectations, since they choose to acquire differing signals, which induces dispersion in beliefs. Our finding that (1) the ranking of information sources does not change with the incentive, and (2) there are systematic differences in the choices of information sources by education/numeracy of consumers, suggests that constraints on information processing are binding for a large part of the sample. Consumers simply do not know what information to pay attention to – this finding provides an explanation for the well-known fact that disagreement in expectations among consumers tends to be much larger than for other agents, even when the estimated degrees of information rigidity are not larger for them (Coibion and Gorodnichenko, 2012).

Our approach is related to a recent literature on information-provision experiments. Particularly relevant for our purposes are papers that employ information experiments in surveys to understand expectation formation in the context of inflation (Armantier et al., 2016; Cavallo et al., 2017; Coibion, Gorodnichenko, and Kumar, 2017) or housing (Armona et al., 2017). The experiments in the context of inflation find that when individuals are provided with official statistics, the dispersion in expectations is reduced substantially. However, these experiments exogenously expose respondents to different information and hence, by design, cannot shed light on the process of information acquisition. Our paper complements this literature by endogenizing the process of information acquisition, and provides sharper tests to distinguish between models of expectation formation.

Finally, our results have an important implication for the design of information interventions. Simply providing more information to consumers may not be sufficient. Either these interventions should be narrowly targeted in the sense that they provide consumers with limited but relevant information. Or consumers need to be guided to help them interpret and weigh the various pieces of information. This, of course, is far from straightforward, since the optimal weight that one ought to put on a certain piece of information may not be obvious in many applications.

get direct utility from their political attitudes/beliefs. On the other hand, it is hard to make that case for home price expectations.
The rest of the paper proceeds as follows. Section 2 introduces the research design and the survey, and outlines the testable hypotheses. Section 3 presents the results. The last section concludes.

2 Survey Design

The main data used in this paper come from a module added to the 2017 SCE Housing Survey. The housing survey has been fielded annually every February since 2014, and contains multiple blocks of questions, some differing between owners and renters. Among other things, respondents are asked about their perceptions of past local home price changes and expectations for future local home price changes, and past as well as future intended housing-related behavior (such as buying a home, and housing debt). Respondents also provide information about their location and many other demographic variables. When appropriate, questions had built-in logical checks (for instance, percent chances of an exhaustive set of events had to sum to 100). Item non-response is extremely rare, and almost never exceeds one percent for any question.

The SCE Housing Survey is fielded as part of the Federal Reserve Bank of New York’s Survey of Consumer Expectations (SCE), which is an internet-based survey of a rotating panel of approximately 1,400 household heads from across the US. The survey, as its name suggests, elicits expectations about a variety of economic variables, such as inflation and labor market conditions. Respondents participate in the panel for up to twelve months, with a roughly equal number rotating in and out of the panel each month.

Active panel members who had participated in a SCE monthly survey in the prior eleven months were invited to participate in the housing module. Out of a total sample of 1,489 household heads on the panel that were invited, 1,161 participated, implying a response rate of 78%.

2.1 Research Design

Appendix A provides screenshots of the relevant module. The broad organization of the module was as follows:

1. **Stage 1- Prior Belief**: This stage elicits individuals’ perceptions of past (one and ten year) national home price changes, as well as their expectations of national home prices at the end of the year.

---

3See https://www.newyorkfed.org/microeconomics/sce/housing#main

4The survey is conducted over the internet by the Demand Institute, a non-profit organization jointly operated by The Conference Board and Nielsen. The sampling frame for the SCE is based on that used for The Conference Board’s Consumer Confidence Survey (CCS). Respondents to the CCS, itself based on a representative national sample drawn from mailing addresses, are invited to join the SCE internet panel. The response rate for first-time invitees hovers around 55%. Respondents receive $15 for completing each survey. See Armantier et al. (2017) for additional information.
Respondents were informed that, according to Zillow, the price of a typical/median home in the US as of December 2016 was $193,800. They were then asked about how the price had changed over the last one year (December 2015) and last ten years (December 2016). They were also asked about their confidence (on a 5-point scale) in their recall.

Regarding expectations for home prices at the end of the year, individuals were first asked for a point forecast: “What do you think the value of the typical home in the U.S. will be at the end of this year (in December 2017)?” To minimize typos, after individuals entered the value, the survey environment would calculate and report the implied percent-change. Individuals could confirm the number and go on to the next screen, or go back to revise their guess. We refer to this as the respondent’s “prior” belief. The survey also elicited the respondent’s subjective belief distribution about home prices at the end of year: specifically, they were asked to assign probabilities to five intervals that future year-end home price changes may lie in (less than -10%; between -10% and -1%; between -1% and 1%; between 1% and 10%; more than 10%).

2. **Stage 2- Preferences over Information Sources**: A block of other housing-related questions taking roughly 15 minutes separated the first and second stages. In the second stage, respondents were notified that we would re-ask the question about future home prices that was asked earlier in the survey, except that this time the elicitation would be incentivized: “This time, we will reward the accuracy of your forecast: you will have a chance of receiving $X. There is roughly a 10% chance that you will be eligible to receive this prize: we will select at random 60 out of about 600 people answering this question. Then, those respondents whose forecast is within 1% of the actual value of a typical US home at the end of this year will receive $X.” We randomly assigned half of the respondents to X=$100 (“High Reward”) and the other half to X=$10 (“Low Reward”).

Before they provided their forecast, respondents were given an opportunity to see a potentially relevant source of information: “Before you report your forecast, you will have the opportunity to see only one of the following pieces of information that may help you with forecasting future year-ahead U.S. home prices. Please rank the following pieces of information on a 1 to 4 scale, where 1 is “Most Preferred” and 4 is the “Least Preferred”:

- Change in the value of a typical home in the U.S. over the last one year (2016).
- Change in the value of a typical home in the U.S. over the last ten years (2007-2016).
- Forecasts of a panel of housing experts about the change in U.S. home prices over this coming year (2017).
- None of the above – I would not like to see any information.”
To provide their rankings, respondents were asked to drag-and-drop each of these options into a table with the labels from “1=Most Preferred” to “4=Most Preferred.”

3. **Stage 3- Willingness-to-Pay for Information:** This stage followed right after the second stage, and elicited the respondent’s maximum willingness-to-pay (WTP) for their highest-ranked information type. This stage was obviously skipped for respondents who ranked “None of the above” as their most preferred information source in Stage 2.

To assess willingness to pay, we used the list price method (e.g., Andersen et al., 2006) with 11 scenarios. In each scenario, respondents had the choice of either seeing their favorite piece of information (that is the one they ranked highest in Stage 2) or receiving extra money with their compensation for completing the survey. The amount of money offered in these scenarios was predetermined, and varied in increments of $0.50 from $0.01 (in Scenario 1) to $5 (in Scenario 11).

Respondents were told that one of these 11 scenarios would be drawn at random, and the decision in that randomly chosen scenario would then be implemented.

4. **Stage 4- Posterior Belief:** In this stage, the respondent may have been presented with information from their highest ranked source; whether the respondent saw the information depended on the randomly chosen scenario in Stage 3, and their choice (of seeing information or not) in that scenario. Year-ahead home price expectations – that were elicited in Stage 1 – were re-elicited from all respondents.

We used the Zillow Home Value Index (ZHVI) as the source of information about the change of the median/typical US home price over the last one/ten years. According to the ZHVI, US home prices had decreased by 0.1% per year on average (or 0.9% in total) over the last ten years, and increased by 6.8% over the last one year. The Zillow Home Price Expectations Survey was the source of information about the experts’ forecast. This is a quarterly survey of 100 economists, real estate experts and market strategists. The average forecast of the experts was an increase of 3.6% in home prices during 2017. Note that these information sources are publicly available.

The paragraph which provided information, of course, depended on the type of information, but followed a similar structure in all three cases. First, it provided the raw information. Second, it offered a naive projection of home prices in December-2017 based on the annual growth rate implied by the information. For instance, in the case of the forecast of experts, respondents were presented with “The average forecast of a distinguished panel of housing

---

5 For more information on the construction of the ZHVI, see http://www.zillow.com/research/zhvi-methodology-6032/ (accessed on December 8, 2017). We used the ZHVI as of December 2016.
6 For details, see https://pulsenomics.com/Home-Price-Expectations.php. We used the average forecast as of the fourth quarter of 2016.
market experts who participate in the Zillow Home Price Expectations Survey is that home values in the US will increase by 3.6% over the next year. If home values were to increase at a pace of 3.6% next year, that would mean that the value of a typical home would be 200,777 dollars in December 2017.”

At the bottom of the same screen where the information was provided (or not provided), expectations about year-end home prices were re-elicited. Respondents were reminded about their prior belief. As in Stage 1, both the point estimate and subjective belief distribution were elicited. We refer to the point estimate from this stage as the “posterior” belief.

Respondents were picked at random with a 10% chance for being eligible for the incentive. They were informed at the very end of the survey about whether they had been picked and were eligible for the potential reward. If they were picked, they were told that they would be paid the potential incentive in January 2018 (since the December 2017 ZHVI was to be used to determine the incentive).

This summarizes the experimental setup. Four months after the initial survey, a short follow-up was fielded to active panelists in the June 2017 SCE monthly survey. As in Stages 1 and 4 of the main experiment, respondents were asked to report their expectations about year-end US median home prices. We kept the frame of reference identical in the follow-up survey: i.e., we provided individuals with the median US home price as of December 2016, and asked them to forecast this value in December 2017. Both the point estimate and subjective density were re-elicited. Of the 1,162 respondents who took the SCE Housing Survey, 762 were still in the panel in June, and hence eligible to take the follow-up survey. Of those, 573 did so, implying a response rate of 75.2%.

2.2 Discussion of the Experimental Design

Our design is supposed to mimic the process of information acquisition and processing in the real world, albeit in a stylized setting. Before turning to the empirical analysis, it is useful to discuss the various features of the experimental design, and to outline the main hypotheses the design allows us to test.

One of the key features of our setup is that respondents are presented with three possible pieces of information along with a no-information option, and then asked to rank them in terms of their preference, with the understanding that they would have the opportunity to see the top ranked source. This allows us to test whether individuals have some reasonable idea about the usefulness of the potential information sources and whether they agree on that.

Ideally, we would want to test the hypothesis that the demand for an information source increases with how informative the information source is. However, there is no single criterion to judge how informative each information source is. One reasonable metric of the usefulness of the information source is given by how well that source would have done in predicting the past
evolution of US home prices if one had simply used it to make a forecast of year-ahead home price changes.

Let $\hat{HPA}_t$ denote the predicted home price change during year $t$. And let $HPA^F_t$ be the mean forecast of experts about home price changes for year $t$, $HPA_{t-1}$ the annualized home price change over the past 10 years, and $HPA_{t-10}$ the annualized home price change over the past 10 years. For each source $I_t \in \{HPA^F_t, HPA_{t-1}, HPA_{t-10}\}$, we define its informativeness as the root mean squared error (RMSE) of a model $\hat{HPA}_t = I_t$. Thus, for the empirical analysis, we test a weaker version of the ideal hypothesis that is based on this specific metric of informativeness:

**Hypothesis 1 (Preference for Informative Signals):** The demand for an information source increases with its ex-ante predictive power.

To calculate the RMSE of each information source, we use CoreLogic home price data since it has a longer time series than Zillow (our qualitative conclusions are unchanged). This yields a RMSE of 4.6 for experts’ forecast, 7.83 for past one-year changes, and 8.37 for past ten-year changes. So based on this criterion, the experts’ forecast has been the most informative in predicting year-ahead home price changes historically, followed by past one-year changes, and then ten-year changes.

The above criterion for ranking the informativeness of the signals is broadly consistent with some basic insights from the real estate literature. First, the fact that the forecasts are ranked highest is consistent with the efficient market hypothesis: one would assume that forecasters use all available information in past home price changes optimally when providing a forecast. Additionally, this is consistent with existing information acquisition models, such as Carroll (2003), in which consumers periodically update their expectations from reports of expert forecasts, which are assumed to be rational. Second, the higher ranking of past one-year home price changes relative to past ten-year changes is consistent with the well-documented momentum in home prices over short horizons (Case and Shiller, 1989; Guren, 2016; Armona, Fuster, and Zafar, 2017). For instance, Armona et al. (2017) find that in a regression of one-year home price changes on lagged one-year home price changes at the zip code level, the average estimate (across the zip codes in the US) is 0.53 (statistically significant with $p < 0.01$). On the other hand, they find that the average estimate (across zip codes) of a regression of one-year home price changes on lagged five-year changes is 0.15, but indistinguishable from zero for the vast majority—more than 80%—of the zip codes.

Even though it is reasonable, our criterion is of course not the only one that one might use to determine the usefulness of information. For example, as of November 2017, according to the ZHVI, US home prices have increased by 6.5% over the past 12 months. Thus, based on ex post accuracy, using the past one-year change would have led to the most accurate expectation. By

---

7 The CoreLogic series starts in 1976. So we the RMSE for one-year (ten-year) changes is based on 38 (29) observations. The Zillow Home Price Expectations Survey started in 2009, and so the RMSE for experts’ forecast is based on 6 data points. Using the Zillow Home Price Value Index, the RMSE for past one-year (ten-year) changes is 5.72 (8.25). So the relative ranking of RMSEs remains unchanged.
this same ex-post metric, however, it is still hard to rationalize picking home price changes over the past 10 years over any of the other two information sources.

Rational inattention would predict that, in the absence of an incentive (such as the lack of a direct stake in the housing market), individuals in the real world may invest less resources in acquiring housing-relevant information and having more informed home price expectations.\(^8\) The goal of introducing variation in the incentive in Stage 2 for the posterior forecast is to test precisely this. That is, whether a higher expected incentive makes respondents willing to pay more for the information. In addition, rational inattention models with constraints on information processing capacity (Sims, 2003; Woodford, 2003; Mackowiak and Wiederholt, 2009) would predict that, when the stakes are higher, respondents think more carefully about the usefulness of the potential information sources and hence rank information differently (in particular, rank “None of the Above” lower) than their counterparts. This leads to our second hypothesis:

**Hypothesis 2 (Rational Inattention):** The willingness to pay for information increases with the expected incentive randomly assigned for the accuracy of the posterior belief. Also, individuals choose more informative sources when the stakes are higher.

Another aspect of the design that merits discussion is that, in Stage 4, some respondents may get to see information from one of the sources. As to whether a respondent sees the information (from the top ranked source) depends on her WTP and the scenario from Stage 3 that is randomly picked. This randomization generates random variation in the provision of information, since for two individuals with identical WTPs in Stage 3, whether the information is shown in Stage 4 or not is determined at random. We exploit this to investigate whether respondents incorporate the signal into their posterior belief, as would be expected if individuals were willing to pay for the information. In addition, rational updating would imply that individuals who have more uncertain prior beliefs and individuals who are willing to pay more for the information put more weight on the information that they receive. This leads to our third hypothesis:

**Hypothesis 3 (Rational Updating):** If individuals are willing to pay for a signal, they should incorporate that signal into their expectation formation once they get access to it. The weight on the signal should be higher for those who pay more for the signal and whose prior uncertainty is higher.

Finally, in models of costly search (such as Reis, 2006) where individuals process information optimally, expectations should be more likely to converge when the cost of acquiring information is lower (since more individuals should observe signals more often). We can test whether the price of information, which was randomly assigned through the randomization of Scenarios 1 through 11, leads to greater convergence in beliefs. In addition, under optimal processing and acquisition

\(^8\)This would follow from most sticky information models. For example, in the sticky updating model of Reis (2006), agents are modeled as maximizing utility subject to constraints, which also include costly information. Increasing the payoff for more informed expectations would lead more agents to incur the cost of acquiring housing-relevant information.
of information, beliefs should converge upon receipt of information. We can test for this since, conditional on one’s WTP, information provision in our setting is random. This leads to our final main hypotheses:

**Hypothesis 4 (Information-Acquisition and Dispersion of Expectations):**
A lowering of the cost of information reduces cross-sectional dispersion in expectations.
In addition, relative to the case with no information, information provision leads a decrease in the dispersion of beliefs.

### 2.3 Survey Implementation

Of the 1,162 valid responses, we trim the sample by dropping the top-5% and the bottom-5% responses to the prior belief. These 130 responses correspond to individuals who reported a prior belief of an annual growth rate below -1.96% (on average, -14.5%) or above 8.36% (on average, 26.5%). These extreme beliefs are probably due to typos, or people not paying attention to the survey question. Since the prior belief was reported before the treatments, dropping these extreme prior beliefs should not contaminate the experimental analysis. For the posterior beliefs, we cannot just drop individuals because that would contaminate the experimental analysis. Instead, we winsorize the post-treatment outcomes to minimize the sensitivity to outliers.\(^9\) In any case, we use graphical analysis whenever possible, to certify that the results are not driven by outliers.

Column (1) of Table 1 shows characteristics of the sample for the main survey. The sample aligns well with average demographic characteristics of the United States along most dimensions. For instance, the average age of our respondents is 50.8 years, and 46.7% are females, while the corresponding numbers among US household heads, according to the 2016 ACS survey, are 45.5 years and 48.0%. 74.6% of respondents are homeowners, compared to a national homeownership rate in 2017:Q1 of 63.6% according to the Census. Our sample, however, has significantly more education and income: 56% of our respondents have at least a Bachelors’ degree, while only 37% of the US household heads fall in this category. Likewise, the household income of the median respondent in the sample is $67,500, substantially higher than the US 2016 median of $57,600. This may partly be a result of differential internet access and computer literacy across income and education groups in the US population. We see that respondents, on average, expect national home prices to increase by 2.1% over the next year.

Columns (2)-(3) of the table show the characteristics for the subsamples assigned to the low and high reward treatments, respectively. Column (4) shows that the characteristics (except for the proportion of Whites) do not statistically differ between the two treatments. This should not be surprising, since random assignment should have largely preserved balance between the two groups. Importantly, we see that the follow-up response rate does not differ according to the

---

\(^9\)For the posterior beliefs from the baseline survey and the beliefs from the follow-up survey, we winsorize the outcomes to take values in the range from -1% to +8%.
reward treatment. Appendix Table B.1 provides the same randomization balance analysis, but for the assignment of price-elicitation scenario in Stage 3 instead of the assignment of the reward size. Again, the evidence suggests that the randomization was successful.

Table 2 provides additional information on how the follow-up sample compares with the initial sample. Columns (2)-(4) show that the subsample that was eligible to be invited to the follow-up survey was similar to the ineligible group (respondents who have been phased out of the panel). The last three columns of the table are also reassuring, since we see no evidence of selection into the follow-up survey based on observables.

3 Empirical Analysis

3.1 Hypothesis 1: Preferences over Informative Signals

To understand how respondents acquire information, it is useful to describe the distribution of expectations prior to the information acquisition. Figure 1.a shows a histogram of the point estimates provided by respondents. In terms of the implied annual growth rates, the mean (median) value is 2.1% (1.7%), with substantial disagreement across respondents: the cross-sectional standard deviation of prior beliefs is 2.0%. To assess if individuals felt confident about their expectations, Figure 1.b shows the probability distribution of beliefs around the individual's own point estimate, averaged over all individuals. On average, individuals are aware that their guesses may turn out to be wrong: they think there is only a 54% chance that the true price will turn out to fall within 1% of their guess. Moreover, there is a lot of dispersion in the degree of certainty: e.g., while 10.5 percent of the sample thinks there is a 90% chance or more of getting it right (that is, within a percent of their guess), 12 percent of the sample thinks there is a 20% chance or less.

What happens when these individuals, with uncertain beliefs, are offered to acquire information? Figure 2.a shows the distribution of ranking order for the different types of information over the whole population. Individuals disagree on which of the three pieces of information they want to buy: 45% chose forecasts of housing experts, 28% chose the last year home price change, and 22% chose the last ten year home price change (the remaining 5% reported to prefer no information at all). The past predictive power criterion indicated that the forecasts were most informative, followed by the last year home price change and then the last ten year home price change. Thus, the popularity of the choices is increasing in the informativeness of the information source. However, this correlation is far from perfect: even though the forecasts are the most popular choice, it was chosen by less than half of the sample.

This heterogeneity in ranking of sources could be driven by lack of knowledge on part of consumers about the relative informativeness of the signals, or respondents using different criteria to determine the informativeness of the signals. However, systematic differences in ranking by education or numeracy of respondents – reasonable proxies for ability to filter signals – would
suggest evidence of the former. This is investigated in Figure 2.c and 2.d, which break down the information choices by numeracy and education, respectively. We see that individuals with more education or with higher numeracy are substantially more likely to choose the “best” source: e.g., the expert forecast was chosen 50% of the time by college graduates relative to 40% of the time by non-graduates (p-value<0.01).

The heterogeneity is explored further in Table 3. The table reports both univariate and multivariate relationships between the choice of each information source and various individual- and location-specific characteristics. Columns (1)-(3) of the table report correlates of choosing each of the information sources in a univariate framework. Besides numeracy and education of respondents, only a handful of variables are significant, suggesting that observable characteristics (at the individual or location level) are otherwise unable to explain most of the heterogeneity in how individuals rank information. Homeowners are more likely to choose the year-ahead forecast, and less likely to choose experts’ forecast. We see that respondents who reported checking external sources during the survey (about 14% of the sample) are 15 percentage points more likely to choose experts’ forecast. Consistent with a selection story, this group of respondents who are savvy enough to look up other information sources are perhaps also more sophisticated and adept at screening informative signals. One might expect respondents who have a high level of confidence in their perception of past home price changes to be more likely to choose the forecast of experts. In fact, we see the opposite: these respondents are more likely to choose the past-year information source, and less likely to choose experts’ forecast. This finding would be consistent with a selection channel, where individuals who are more certain about past home price changes are likely the ones who have looked for that information in the past and thus revealed a “preference” for that type of information. Likewise, one might expect respondents residing in states with more volatile housing prices (as measured by the standard deviation in monthly home prices over the past 24 months) to be less likely to choose past home price changes. We do not find evidence of that.

Our conclusions remain unchanged in the columns (5)-(7) of the table, which reports estimates from a multivariate regression. This leads to our first result:

**Result 1:** While the information source with the highest ex-ante predictive power — experts’ forecast — is the modal choice, there is considerable disagreement across households on the relative ranking of information sources. The ranking of sources is systematically related with measures of ability, which suggests that the heterogeneity is driven, at least in part, by cognitive limitations in deciphering informative signals.

---

10 Similarly Burke and Manz (2014) find that respondents with higher levels of economic literacy choose more relevant information when forming inflation forecasts.
3.2 Hypothesis 2: Rational Inattention

Before we can test if higher stakes changes the willingness to pay for information, it is useful to understand the distribution for the whole sample. Using the responses to the eleven hypothetical scenarios, we can identify the range of an individual’s WTP: e.g., if an individual chose information over any amount up to $3, and then chose the amount from $3.50 on, it would mean that the individual’s WTP must be in the range from $3, to $3.5. Around 5% of respondents provided inconsistent responses to scenarios: e.g., they chose information over $3 but then chose $2.5 over information. This share of inconsistent respondents is within the range of other studies using this list method for elicitation of WTP for information. For instance, the share of inconsistent respondents was about 2% in Allcott and Kessler (2015), and 15% in Cullen and Perez-Truglia (2017).

Figure 3.a shows the histogram of these WTP based on this approach. We find that individuals have significant willingness to pay for their favorite information. This figure indicates that the median maximum willingness to pay for their favorite information is between $4.5 and $5.\textsuperscript{11} This is a fairly high level of WTP given that the information we provide is publicly available – for example, one can find these pieces of information using a search tool like Google in a few minutes. This indicates that most individuals are either unaware of the availability of this information, or they expect a high search cost. Also, the median MWP ($4.5-$5) is quite high compared to the expected reward in case individuals are perfectly accurate ($1 for half of the sample and $10 for the other half). This evidence suggests that individuals value the information beyond the context of the survey – for instance, they may want to use this information in the formation of beliefs relevant for their real-world housing decisions. In this context, having incorrect expectations about house prices can translate in expected loses in the order of thousands of dollars, relative to which the experimental incentive pales in comparison. Additionally, we can compare the median WTP in our study ($4.5-$5) with the results from a few other papers that elicited willingness to pay for information using similar methods. Those studies find valuations that are lower: $0.40 for travel information (Khattak, Yim and Prokopy, 2003), $0.80 for food certification information (Angulo, Gil and Tamburo, 2005), and $3 for home energy reports (Allcott and Kessler, 2015).

We next test Hypothesis 2, i.e., whether the WTP and ranking of information systematically varies with reward size. Figure 3.b conducts a non-parametric test of this hypothesis, by comparing the distribution of WTP between the two reward groups. This figure suggests that, consistent with the rational inattention hypothesis, individuals are willing to pay more when in the higher reward treatment. The Epps-Singleton test suggests that this difference is statistically significant (p-value=0.02).\textsuperscript{12}

\textsuperscript{11} An alternative estimate is given by means of an interval regression model. This is a maximum likelihood model that assumes that the latent WTP is normally distributed. The median valuation, which is given by the constant in this model, is estimated to be $4.41 (95% CI from 4.18 to 4.65).

\textsuperscript{12} The Epps-Singleton test is a version of the Kolmogorov–Smirnov test of equality of distributions that is valid
To better understand the economic magnitude of this difference, column (4) of Table 4 presents the rational inattention test in regression form. The constant reported in column (4) can be interpreted as the median WTP for the low-reward condition ($10 with 10% probability). This median valuation is estimated to be $4.05 (95% CI from 3.73 to 4.38).\footnote{The fact that even under an expected reward of just $1 individuals are willing to pay $4.05 for the information suggests that individuals value the information beyond the context of the survey (i.e., in their natural context).} The coefficient on High Reward indicates that, relative to the $10 reward, individuals assigned to the $100 reward were willing to pay an additional $0.73 for their favorite information (or 18% more). Note that the expected reward goes from $1 to $10, because the reward is given only with 10% probability. The $0.73 difference in WTP then implies that for each additional dollar of expected reward, the WTP for info goes up by 8.11 cents.

Another way to interpret the result is in the following terms:

\[
WTP_i = U_{Info} + 0.1 \cdot \text{Reward}_i \cdot [P_i(\text{Accurate}|\text{Info}) - P_i(\text{Accurate}|\text{NoInfo})] + \varepsilon_i.
\]

The first term, \(U_{Info}\), represents the expected value of the benefits from the information in the real world (e.g., because the individual expects to make better choices when deciding whether to buy a house). The second term, reflects the benefits from the information through the survey reward, under the simplifying assumption that the respondent is risk-neutral over small amounts. We can infer the value of \(P_i(\text{Accurate}|\text{Info}) - P_i(\text{Accurate}|\text{NoInfo})\) from a regression of \(WTP_i\) on \(0.1 \cdot \text{Reward}_i\). Indeed, we do not even need to run a new regression – we can recover that parameter from the coefficients on column (4) of Table 4.\footnote{The coefficient on the High Reward dummy indicates that increasing \(0.1 \cdot \text{Reward}_i\) by 9 (i.e., \(0.1 \cdot 100 - 0.1 \cdot 10\)) increases the WTP by $0.73. Thus, increasing \(0.1 \cdot \text{Reward}_i\) by 1 would increase the WTP by 0.081 (= \(0.73 / 9\)).} This estimator suggests that \(P_i(\text{Accurate}|\text{Info}) - P_i(\text{Accurate}|\text{NoInfo}) = 0.081\). In other words, by acquiring the information, the median individual expects her probability of being accurate (i.e., being within 1% of the realization) will increase by 8.1 percentage points (or, equivalently, 15% of the baseline probability).\footnote{The median individual responded that there was a 54% chance that their guess fell within 1% of the true price. We use this as an estimate of the median \(P_i(\text{Accurate}|\text{NoInfo})\). Thus, the 8.1 percentage point effect translates into a 15\% (=8.1/54) effect.}

It is worth asking whether the level of attention varies systematically with respondents’ abilities. Column (5) and (6) of Table 4 investigate whether higher numeracy and higher educated individuals are more rationally attentive (i.e., react more to the higher reward). In column (5), the High-Reward dummy is interacted with a measure of numeracy (from 0 to 5). In column (6), the High-Reward dummy is interacted with a dummy for college graduate. In column (5), the effect of high reward is 80\% larger for individuals with highest numeracy with respect to the effect for individuals with lowest numeracy, with the difference being statistically significant at the 5\% level.
In column (6), the effect of high reward is 134% larger for college graduates with respect to non-graduates, although the difference is imprecisely estimated and thus statistically insignificant.\textsuperscript{16} We had already seen that higher-educated and higher-numeracy respondents were more likely to choose expert’s forecast. So not only are lower education/numeracy less likely to rank information sources optimally, they are also less responsiveness to the higher rewards.

Given that individuals pay more for information when the stakes are higher, the next question is whether individuals choose information types differently when the stakes are higher. Figure 2.b breaks down the information choice by reward type. The choices are almost identical across both groups; the p-value of the difference is 0.88. Column (1) of Table 4 presents this same test in regression form. It corresponds to a linear probability model where the dependent variable is whether the individual chose the forecast (i.e., the “best” information type according to the past predictive power). Column (1) suggests that individuals were not more likely to choose forecasts under the higher reward. In addition, Columns (2) and (3) of Table 4 show that the effect of the larger reward on choosing the expert forecast does not differ by numeracy or education.

This leads to our second result:

\textbf{Result 2:} Consistent with rational inattention, the WTP for information is higher when the incentive is higher, with greater responsiveness by more sophisticated respondents. However, the ranking of the information types does not systematically differ by reward size.

We next investigate the drivers of the heterogeneity in the WTP. Column (4) from Table 3 uses the interval regression model to estimate the effect of a set of factors on the WTP, with the impact of each factor investigated one at a time. Gender or education and numeracy are not systematically related with the WTP. We see that owners and individuals residing in states with higher median home prices – the group of respondents who arguably have higher stakes in the housing market – have (economically and statistically) significantly higher WTPs. For example, increasing the median house value in state by 1 SD increases the WTP by 22 cents.

Ex ante one would expect that individuals with higher confidence in their recall of past home price changes and those who have looked for housing-related information in the past would value the information presented in the experiment less; the two variables are positively correlated (correlation coefficient of 0.18). However, we find the opposite: for example, individuals who looked for housing-related information in the past are willing to pay an additional 59 cents, relative to those who have not. This suggests that the “selection” channel dominates. That is, individuals who have looked for information in the past are also those who intrinsically value information more and have a higher demand for it.

\textsuperscript{16}Note the the estimate for the High Reward dummy is no longer significant. On the other hand, the impact of the High Reward for college-educated respondents, which is the sum of the two estimates is a precisely estimated $0.96. Thus, the impact of the higher rewards on the WTP is primarily driven by college-educated respondents.
We would expect individuals with more uncertain prior beliefs to be willing to pay more for information, because they stand to learn more from it. To measure uncertainty at the individual level, we use the responses to the probability bins. We fit these binned responses to a normal distribution for each individual, and use the estimated standard deviation of the fitted distribution as a measure of individual-level uncertainty, with higher values denoting higher uncertainty. For instance, an individual with a 2% expectation has an uncertainty of 1pp, it means that the individual’s 95% confidence interval for her expectation is [0.04%, 3.96%] (= [2 − (1 * 1.96), 2 + (1 * 1.96)]).

When looking at the relationship between uncertainty of prior beliefs and willingness to pay, we find the opposite—individuals with higher prior uncertainty are, on average, willing to pay $.29 less. This again could be consistent with a “selection” story, where individuals who have more precise priors are the ones who value information more.17

The effect of local volatility in home prices on the WTP is ex ante less clear. On one hand, updating more often is more valuable for such respondents, and hence they should value information more. On the other hand, past changes in home prices would be less informative. We have already seen that respondents in these areas do not choose experts’ forecast more often. Here, we see that these respondents in fact value information more: increasing the volatility in state by 1 SD increases the WTP by 31 cents.

Column (8) of Table 3 is similar to column (1), except that it presents estimates from a multivariate regression. We see that qualitatively estimates continue to be similar.

We know that historically experts’ forecasts have done better (relative to the other two information pieces) in terms of predicting home price changes. Then, under this metric, individuals should be willing to pay more for experts’ forecasts. Panels c and d of Figure 3 break down the WTP by information type, showing how the WTP for experts’ forecast compares with that for past 1 year and 10 year home price changes, respectively. To establish whether these pairwise differences are statistically significant or not, these figures report the results from an Epps–Singleton (ES) two-sample test using the empirical characteristic function. These pairwise differences are statistically insignificant; that is, individuals in each of these three groups seem to be equally willing to pay for the information.

3.3 Hypothesis 3: Rational Updating

Recall that our design generated quasi-random variation in whether a respondent gets to see information. For two individuals with identical WTP (and conditional on top-ranked information), whether information is shown or not is determined by chance. We use this random variation in the information provision to estimate the rate at which individuals absorbed the signal. Furthermore, we can calculate this learning rate for different sub-populations, particularly for sub-groups

\[ \text{Note that the correlation of prior uncertainty with education/numeracy as well as with looking up housing-related information in the past is negative.} \]
choosing the different information sources.

We use a simple learning model that naturally separates learning from the signal shown from other sources of signal-reversion.\textsuperscript{18} Let $b_{\text{prior}}$ denote the mean of the prior belief, $b_{\text{signal}}$ the signal, and $b_{\text{posterior}}$ the mean of the corresponding posterior belief. When priors and signals are normally distributed, then Bayesian learning implies that the mean of the posterior belief should be a weighted average between the signal and the mean of the prior belief:

$$b_{\text{posterior}} = \alpha \cdot b_{\text{signal}} + (1 - \alpha) \cdot b_{\text{prior}}.$$ 

The degree of learning can be summarized by the weight parameter $\alpha$. This parameter can take a value from 0 (individuals ignore the signal) to 1 (individuals fully adjust to the signal). Rearranging this expression, we get:

$$b_{\text{posterior}} - b_{\text{prior}} = \alpha \cdot (b_{\text{signal}} - b_{\text{prior}}).$$

That is, the slope between the perception gaps ($b_{\text{signal}}^k - b_{\text{prior}}^k$) and revisions ($b_{\text{posterior}}^k - b_{\text{prior}}^k$) can be used to estimate the learning rate. However, it is possible that individuals would still revise their beliefs towards the signal even if they were not provided with the signal. For instance, consider someone who makes a typo when entering her prior belief and reports an estimate that differs significantly from the signals. If she does not commit the typo again when reporting her posterior belief, it will look like as if she is reverting to the signal even though she is not shown information. Also, it is possible that individuals think harder the second time they are asked about their home price expectation, especially since the posterior belief was incentivized but the prior belief was not. It is certainly plausible that some individuals search for more housing-related information online during the survey. At the end of the survey, we asked respondents whether they had searched for information online during the duration of the survey, making it clear that doing so was permitted. 14.4\% reported doing so. Interestingly, the search rate does not differ by whether respondents saw information in the experiment (the search rate is 14.6\% for respondents who saw the information, and 14.1\% for those who did not). Also note that the simple act of taking a survey about housing may also make respondents think more carefully about their responses, and may lead them to revise their home price expectations even if they are not provided with any new information (see Zwane et al. 2011 for a discussion of how surveying people may change their subsequent behavior). Thus, we need to use the random variation in information provision to separate true learning from mean-reversion. Consider the dummy $S_i$ that takes the value 1 if the individual was shown the signal. Let $WTP_i$ be a set of dummies corresponding to the “threshold price” chosen by the individual in the scenarios. Conditional on this threshold, whether the individual gets the information or not ($S_i$) is determined entirely on the scenario that is executed, which is randomly chosen. Thus, we use the following regression specification:

\textsuperscript{18}Similar learning models are used in Armantier et al. (2016) and Cavallo et al. (2017).
\[ b_{i}^{\text{posterior}} - b_{i}^{\text{prior}} = \alpha \cdot (b_{i}^{\text{signal}} - b_{i}^{\text{prior}}) \cdot S_i + \beta \cdot (b_{i}^{\text{signal}} - b_{i}^{\text{prior}}) + WTP_i \delta + \varepsilon_i, \]

The parameter of interest is still \( \alpha \), which measures the true learning rate: i.e., the effect of being randomly shown information on the updates. On the other hand, \( \beta \) reflects the degree of spurious mean-reversion.

The results from this regression are presented in Figure 4.a. The y-axis indicates the revision in national home price beliefs, i.e., posterior belief minus prior belief. The x-axis shows the “gap” between the signal and the prior belief, interacted by the treatment assignment dummy. For instance, if the respondent had a prior belief of 1% and was shown the forecast of experts (which was 3.6%), the x-axis would take the value of 2.6%. Intuitively, the x-axis shows how much potential for revision there is, and the y-axis shows the actual revision. If individuals fully incorporated the signals, we would expect all dots to lie on the 45-degree line. If individuals did not incorporate any of the information, we would expect the dots to lie in a horizontal line. The slope of the line is 0.445, which is not only highly statistically significant (p-value<0.001), but also economically substantial: it is midways between the case where individuals fully react to the information (slope of 1) than the case where individuals fully ignore the information (slope of 0). In other words, the average individual puts 44.5% weight on the signal and 55.5% on their prior belief.

A natural question to ask is what the medium-term impacts of information are on mean beliefs as well as cross-sectional dispersion. To test whether the effect on beliefs was persistent, Figure 4.b reproduces Figure 4.a but instead of using \( b_{i}^{\text{posterior}} - b_{i}^{\text{prior}} \) as the y-axis, we use: \( b_{i}^{\text{follow-up}} - b_{i}^{\text{prior}} \), where \( b_{i}^{\text{follow-up}} \) is the belief reported four months later, in the follow-up survey. There is good reason to believe that the slope will be lower in the medium-term horizon because individuals may be exposed to additional signals during the interim four months, thus gradually watering down the effect of the signal provided during our experiment. The result, shown in Figure 4.b, supports the hypothesis of persistent learning: the estimated slope (0.169) is smaller than the short-term equivalent (0.445), but it is still economically significant and statistically significant at the 10% level.

Figure 5.a. investigates whether the leaning rates differ across the three information sources. Ex ante, there is little reason for why the rates should differ: once a respondent has revealed their preference for a certain information source, they should be equally responsive to it. Panels b and c of Figure 5 investigate whether the learning rate differs by the WTP for information or uncertainty in prior belief. Under Bayesian updating, respondents who are more uncertain should put more weight on the signal. Likewise, individuals who value the information more should arguably put more weight on it. We fail to find evidence of differential learning based on either prior uncertainty or the WTP.

Our next result is as follows:
**Result 3:** Subjects incorporate information that they bought and the weight that respondents put on the information does not vary by information source. However, contrary to “rational” updating, the weight does not differ by one’s WTP for the information or prior uncertainty.

### 3.4 Hypothesis 4: Information-Acquisition and Dispersion of Expectations

In a setting where individuals weight the signals optimally, more access to information (either due to a reduction in the cost of information, or an increase in the incentive for the accuracy of the beliefs) should reduce the cross-sectional dispersion in beliefs. We next investigate this directly.

Since, in Stage 3, a scenario is picked at random, the experimental set-up induces exogenous variation in the cost of information. We exploit this next, and compare how beliefs evolve in the case where “low price” scenarios (prices in the range $0.01-$1.5) were picked at random, versus “high price” scenarios. Table 5 shows how the beliefs in the full sample evolve for the low- and high-price groups. As expected (due to the scenario being picked at random), the distribution of prior beliefs for the two groups is very similar. At the final stage, the mean posterior belief for the two groups is significantly different, a result of different exposure to information for the two groups. However, we do not find evidence of dispersion, as measured the the cross-sectional standard deviation, going down more for the low-price group. The average standard deviation of the posterior is similar for both groups, and in fact is higher than that for the prior.

How is it that more information does not induce higher consensus? This becomes clearer in Figures 6 and 7. Figure 6 shows the distribution of prior beliefs for individuals who would not be shown the information (Figure 6.a) versus individuals who would eventually be shown the information (Figure 6.b). Comparing the two indicates that these two groups started with similar distribution of beliefs. Figure 7 shows the comparison of posterior beliefs between individuals who were not shown information (7.a) versus individuals who were shown the information (Figure 7.b). Figure 7.a shows that, among individuals who were not shown any information, the distribution of posterior beliefs is the same regardless of whether the individuals preferred the experts’ forecast, past 1-year home price changes, or past 10-year changes.19 In contrast, Figure 7.b shows that, for individuals who saw the information, posterior beliefs ended up dramatically different across the three information types. In each group, posterior beliefs moved towards the values of the respective signals: that is, 0.1% for the ten years price change, 3.6% for the expert forecast and 6.8% for the one year price change. As a result, within a certain information type, the revelation of information tended to decrease dispersion in expectations. However, since those groups moved towards differing

---

19 Consistent with the evidence discussed above that subjects in the no-information group may have searched for information or thought harder about the question, a comparison of Figure 6.a versus Figure 7.a indicates that the distribution of beliefs changed from prior to posterior beliefs even for individuals who were not shown information.
signals, the dispersion of beliefs across those three groups increased. The net effect of information acquisition on belief dispersion depends on the combination of these two channels.

Table 6 summarizes the results from Figures 6 and 7. The first two columns of Table 6 provide statistics of how the average level, the dispersion, and uncertainty of beliefs evolve, conditional on seeing the information and the most-preferred information source. We are primarily interested in how the dispersion in beliefs changes. As a measure of belief dispersion, we focus on the standard deviation (SD) of beliefs. The first thing that we can corroborate is that, within information types, information provision tended to reduce belief dispersion. For instance, for individuals who preferred the experts’ forecast but did not get to see the information, the SD in beliefs went up, from 2.0 percentage points (pp) for prior beliefs to 2.3 pp for posterior beliefs. In contrast, for individuals who preferred the forecast and were shown the information, the SD in beliefs went down (from 2.0 pp to 1.4 pp). The same is true for individuals who chose the past 10 years change, but not for individuals who chose 1 year change (for whom, probably due to the extreme value of the signal, the dispersion went up).

To see the net effect of information provision, we can compare the evolution of belief dispersion for the entire sample. For the group that did not see the information, dispersion went up (from 2.0 pp to 2.3 pp) from prior to posterior. In comparison, the dispersion went up by exactly the same amount for individuals to which information was shown. This evidence suggests that the information acquisition opportunity did not lead to a decrease in the dispersion in beliefs. However, information leads to significantly different average posterior beliefs for the two groups.

This leads to the following result:

**Result 4**: A lower cost of information does not lead to a decrease in the cross-sectional dispersion of beliefs. Likewise, endogenous information-acquisition and information-processing does not lead to lower dispersion.

We can also see how the cross-sectional dispersion and individual-level uncertainty evolves in the medium-term. Ex-ante, the medium-term impact on dispersion is not entirely clear. In the interim four months, individuals may receive various signals, and depending on the heterogeneity in the signals the respondents receive, the cross-sectional dispersion may go up or down. Individual-level uncertainty should, however, not go up since individuals are being asked about year-end home prices, and at least some uncertainty could be resolved in the interim four months. The last column in Table 6 shows how these statistics evolve for the information-shown and not-shown groups. Consistent with persistent learning, the mean expectation for the information-shown group continues to be different from that of the not-shown group. However, comparing the follow-up belief with the posterior belief, we see that the cross-sectional dispersion has gone up for both the information-shown and not-shown groups. Notably, uncertainty is unchanged for the information-shown group, but declines quite a bit for the not-shown group, for whom the average uncertainty is now similar to the group that was shown information in our experiment. This suggests that
some uncertainty gets resolved even for the not-shown group in the interim four months.

It is also worth commenting on how uncertainty at the individual-level evolves. This is reported in the last row in each panel in Table 6. We see that mean uncertainty in posterior beliefs is lower than that in prior beliefs for both groups (those who saw the information as well as those who did not). However, consistent with the notion that information should make individuals more certain, we see that uncertainty declines more so for the group that is shown information: the mean uncertainty for the information-shown group declines from 4.9 pp to 4.4 pp (a decline of 0.50 pp), versus a decline from 5.3 pp to 4.9 pp for the not-shown group.

### 4 Conclusions

Using an innovative experimental setup that makes the information acquisition process endogenous, this paper attempts to understand the role of information frictions in explaining the heterogeneity in consumers’ expectations about home price changes. Consumers exhibit substantial demand for information and, consistent with rational inattention, the demand for information is higher when the stakes are higher. While information acquisition costs do seem to matter, our findings indicate that the main driver of the heterogeneity in consumers’ expectations are constraints on information processing. Consumers disagree on what sources of information are most informative, with less sophisticated agents less likely to choose “informative” signals. Importantly, as a result of endogenous information acquisition, we see that the cross-sectional variance of the expectations distribution does not decrease (as would have been expected in a setting with rational acquisition and processing of formation). This is because, while individuals respond to the different information sources, they do not weigh the different pieces of information optimally. This has implications for both the modeling of expectation formation, and for the practical design of information interventions.

On the modeling front, most models in the literature with information frictions assume that individuals, once they acquire information, process it in a rational way. Our results suggest that this may not be a great assumption. Instead our results lend support to models where consumers have limited information processing capacity, and may process information only at a finite rate (for example, as in Sims, 2003). Our findings suggest that consumers may simply not know which pieces of information to pay attention to. In fact, we find that less sophisticated individuals (as proxied by education or numeracy) are less likely to pick more informative signals. As a result, after the information experiment, the posterior beliefs of such consumers are more dispersed, relative to their sophisticated counterparts. In fact, our results also provide an explanation for the puzzle for why consumers tend to have so much disagreement in their expectations.

On the practical front, our results underscore the need for information campaigns to better-designed. Our findings imply that it is not sufficient to simply provide people with more information. Instead, individuals need to be guided in terms of helping them navigate how to interpret
and weigh the different information sources. Policy-makers may want to act in a paternalistic way, by only disclosing the “good signals” (or they make disclose all signals but make the best signals more salient).
References


Figure 1: Prior Beliefs: Expectations about Median House Price

a. Point Estimate

b. Uncertainty

Notes: Panel (a) shows the distribution of the expected value of the typical home in the U.S. 1 year forward (i.e., from December 2016 to December 2017). The green line corresponds to the median house value in U.S. in December 2006. The histogram is censored at $190,000 and $210,000. Panel (b) corresponds to the distribution of the confidence about the forecast made in Panel (a) by individuals.
Notes: Panel (a) shows the distribution of the type of information most preferred by individuals that may help them with forecasting future year-ahead U.S. home prices. Panel (b) provides the same information according the size of the reward. Panel (c) according to the level of numeracy. And, Panel (d) according to the level of education. P-value of difference tests the joint significance of the estimates of a multinomial logit regression.
Figure 3: Willingness to Pay for Favorite Information

a. All

b. By Reward Size

c. Forecast vs. 1yr

d. Forecast vs. 10yr

Notes: Panel (a) shows the distribution of maximum willingness to pay for favorite information in the whole sample. Panel (b) shows the distribution of maximum willingness to pay for information according the size of the reward. Panel (c) compares the distribution of MWP between individuals who preferred forecasts information and individuals who preferred information over the last one year. Panel (d) compares the distribution of MWP between individuals who preferred forecasts information and individuals who preferred information over the last ten years. P-value of difference refers to the Epps-Singleton characteristic function that tests the equality of two distributions.
Figure 4: Short and Medium-Term Learning Rates

a. Short Term

b. Medium-Term

Notes: The dots correspond to the binned-scatterplot based on 20 bins. Slopes, robust standard errors (in parentheses) and $R^2$ are based on a linear regression.
Figure 5: Learning from Feedback

a. By Info Chosen

b. By WTP

c. By Uncertainty in Priors

Notes: The dots correspond to the binned-scatterplot based on 20 bins. Slopes, robust standard errors (in parentheses) and $R^2$ are based on a linear regression.
Figure 6: Prior Beliefs: Individuals Who Will not be Shown Information vs. Individuals Who Will

a. Information not to be Shown

b. Information to be Shown

Notes: The distribution of the prior beliefs according the type of information most preferred. Panel (a) shows the distribution when individuals will not be shown information. Panel (b) shows the distribution when individuals will be shown information.
Figure 7: Posterior Beliefs: Individuals Who Were Shown Information vs. Individuals Who Were Not

a. Information Not Shown

b. Information Shown

Notes: The distribution of the posterior beliefs according the type of information most preferred. Panel (a) shows the distribution of individuals who were not shown the information. Panel (b) shows the distribution of individuals who were shown the information.
Table 1: Descriptive Statistics and Randomization Balance by Reward Size

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Low Reward</th>
<th>High Reward</th>
<th>(4) F-test P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior Belief ($1,000s)</td>
<td>197.8</td>
<td>197.9</td>
<td>197.7</td>
<td>0.408</td>
</tr>
<tr>
<td></td>
<td>(0.121)</td>
<td>(0.172)</td>
<td>(0.170)</td>
<td></td>
</tr>
<tr>
<td>Prior Belief (% change)</td>
<td>0.0210</td>
<td>0.0210</td>
<td>0.0200</td>
<td>0.408</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Income &gt; $60,000</td>
<td>0.570</td>
<td>0.593</td>
<td>0.546</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>College Graduate</td>
<td>0.564</td>
<td>0.566</td>
<td>0.562</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>50.82</td>
<td>51.09</td>
<td>50.55</td>
<td>0.578</td>
</tr>
<tr>
<td></td>
<td>(0.485)</td>
<td>(0.694)</td>
<td>(0.677)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.467</td>
<td>0.459</td>
<td>0.475</td>
<td>0.612</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.640</td>
<td>0.660</td>
<td>0.620</td>
<td>0.183</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.021)</td>
<td>(0.021)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.819</td>
<td>0.788</td>
<td>0.849</td>
<td>0.0100</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.018)</td>
<td>(0.016)</td>
<td></td>
</tr>
<tr>
<td>Own Residency</td>
<td>0.746</td>
<td>0.751</td>
<td>0.741</td>
<td>0.722</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.019)</td>
<td>(0.019)</td>
<td></td>
</tr>
<tr>
<td>Responded to follow-up survey</td>
<td>0.555</td>
<td>0.554</td>
<td>0.556</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.022)</td>
<td>(0.022)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>1,032</td>
<td>514</td>
<td>518</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Individual characteristics obtained from baseline survey. Column (1) corresponds to all respondents, columns (2) and (3) correspond to each of the treatment of the size of reward. Column (4) present p-value for the test of the null hypothesis that the mean characteristic is equal across all treatment groups. All variables constructed from the survey data.
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>No</td>
<td>Yes</td>
<td>P-value</td>
<td>No</td>
<td>Yes</td>
<td>P-value</td>
</tr>
<tr>
<td>Prior Belief ($1,000s)</td>
<td>197.8 (0.121)</td>
<td>198.0 (0.213)</td>
<td>197.7 (0.147)</td>
<td>0.284</td>
<td>197.7 (0.378)</td>
<td>197.7 (0.159)</td>
<td>0.932</td>
</tr>
<tr>
<td>Prior Belief (% change)</td>
<td>0.0210 (0.001)</td>
<td>0.0220 (0.001)</td>
<td>0.0200 (0.001)</td>
<td>0.284</td>
<td>0.0200 (0.002)</td>
<td>0.0200 (0.001)</td>
<td>0.932</td>
</tr>
<tr>
<td>Income &gt; $60,000</td>
<td>0.570 (0.015)</td>
<td>0.546 (0.026)</td>
<td>0.582 (0.019)</td>
<td>0.276</td>
<td>0.654 (0.047)</td>
<td>0.569 (0.021)</td>
<td>0.0970</td>
</tr>
<tr>
<td>College Graduate</td>
<td>0.564 (0.015)</td>
<td>0.555 (0.026)</td>
<td>0.569 (0.019)</td>
<td>0.673</td>
<td>0.615 (0.048)</td>
<td>0.560 (0.021)</td>
<td>0.290</td>
</tr>
<tr>
<td>Age</td>
<td>50.82 (0.485)</td>
<td>51.52 (0.861)</td>
<td>50.45 (0.585)</td>
<td>0.302</td>
<td>48.60 (1.363)</td>
<td>50.78 (0.645)</td>
<td>0.146</td>
</tr>
<tr>
<td>Female</td>
<td>0.467 (0.016)</td>
<td>0.499 (0.027)</td>
<td>0.451 (0.019)</td>
<td>0.142</td>
<td>0.529 (0.049)</td>
<td>0.436 (0.021)</td>
<td>0.0830</td>
</tr>
<tr>
<td>Married</td>
<td>0.640 (0.015)</td>
<td>0.639 (0.026)</td>
<td>0.640 (0.018)</td>
<td>0.996</td>
<td>0.606 (0.048)</td>
<td>0.646 (0.020)</td>
<td>0.442</td>
</tr>
<tr>
<td>White</td>
<td>0.819 (0.012)</td>
<td>0.820 (0.020)</td>
<td>0.818 (0.015)</td>
<td>0.956</td>
<td>0.837 (0.036)</td>
<td>0.815 (0.016)</td>
<td>0.589</td>
</tr>
<tr>
<td>Own Residency</td>
<td>0.746 (0.014)</td>
<td>0.755 (0.023)</td>
<td>0.742 (0.017)</td>
<td>0.636</td>
<td>0.750 (0.043)</td>
<td>0.740 (0.018)</td>
<td>0.829</td>
</tr>
<tr>
<td>Observations</td>
<td>1,032</td>
<td>355</td>
<td>677</td>
<td>104</td>
<td>573</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Individual characteristics obtained from baseline survey. Column (1) corresponds to all respondents, column (2) corresponds to individuals who were not invited to the follow-up survey, column (3) corresponds to individuals who where invited to the follow-up survey. Column (4) presents p-value for the test of the null hypothesis that the mean characteristic is equal across (2) and (3). Column (5) corresponds to individuals who were invited to the follow-up survey but did not respond. Column (6) corresponds to individuals who were invited to the follow-up survey and responded. Finally, column (7) presents p-value for the test of the null hypothesis that the mean characteristic is equal across (5) and (6). All variables constructed from the survey data.
Table 3: Factors Associated to Information Choice and Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dummy chose regressions</td>
<td>Dummy chose Multivariate regressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Forecast</td>
<td>1-yr</td>
<td>10yr</td>
<td>MWP</td>
<td>Forecast</td>
<td>1-yr</td>
<td>10-yrs</td>
<td>MWP</td>
</tr>
<tr>
<td>Female</td>
<td>0.022</td>
<td>-0.012</td>
<td>-0.010</td>
<td>-0.293</td>
<td>0.033</td>
<td>-0.015</td>
<td>-0.019</td>
<td>-0.171</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.029)</td>
<td>(0.027)</td>
<td>(0.243)</td>
<td>(0.033)</td>
<td>(0.030)</td>
<td>(0.028)</td>
<td>(0.251)</td>
</tr>
<tr>
<td>College</td>
<td>0.106***</td>
<td>-0.075**</td>
<td>-0.031</td>
<td>0.079</td>
<td>0.073**</td>
<td>-0.050</td>
<td>-0.023</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.030)</td>
<td>(0.027)</td>
<td>(0.250)</td>
<td>(0.033)</td>
<td>(0.031)</td>
<td>(0.029)</td>
<td>(0.257)</td>
</tr>
<tr>
<td>High Numeracy</td>
<td>0.119***</td>
<td>-0.087***</td>
<td>-0.032</td>
<td>-0.050</td>
<td>0.112***</td>
<td>-0.083***</td>
<td>-0.030</td>
<td>-0.295</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.029)</td>
<td>(0.027)</td>
<td>(0.240)</td>
<td>(0.034)</td>
<td>(0.030)</td>
<td>(0.029)</td>
<td>(0.254)</td>
</tr>
<tr>
<td>Looked for Outside Info during survey</td>
<td>0.150***</td>
<td>-0.122***</td>
<td>-0.028</td>
<td>0.048</td>
<td>0.123***</td>
<td>-0.102***</td>
<td>-0.021</td>
<td>-0.065</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.036)</td>
<td>(0.037)</td>
<td>(0.335)</td>
<td>(0.046)</td>
<td>(0.038)</td>
<td>(0.038)</td>
<td>(0.340)</td>
</tr>
<tr>
<td>Uncertainty in Prior Belief (Std)</td>
<td>0.005</td>
<td>-0.004</td>
<td>-0.000</td>
<td>-0.291**</td>
<td>0.005</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.262**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.123)</td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.124)</td>
</tr>
<tr>
<td>Median House Value in State (Std)</td>
<td>0.029*</td>
<td>-0.021</td>
<td>-0.008</td>
<td>0.216*</td>
<td>0.021</td>
<td>-0.011</td>
<td>-0.010</td>
<td>0.173</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.013)</td>
<td>(0.125)</td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.132)</td>
</tr>
<tr>
<td>House Value Volatility in State (Std)</td>
<td>0.005</td>
<td>-0.014</td>
<td>0.009</td>
<td>0.305***</td>
<td>-0.007</td>
<td>-0.007</td>
<td>0.014</td>
<td>0.258**</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.118)</td>
<td>(0.017)</td>
<td>(0.014)</td>
<td>(0.015)</td>
<td>(0.121)</td>
</tr>
<tr>
<td>Dummy Looked for Info in Past</td>
<td>-0.005</td>
<td>0.015</td>
<td>-0.010</td>
<td>0.589**</td>
<td>-0.031</td>
<td>0.031</td>
<td>-0.000</td>
<td>0.445*</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.029)</td>
<td>(0.027)</td>
<td>(0.246)</td>
<td>(0.033)</td>
<td>(0.030)</td>
<td>(0.029)</td>
<td>(0.263)</td>
</tr>
<tr>
<td>Dummy Own House</td>
<td>-0.082**</td>
<td>0.091***</td>
<td>-0.010</td>
<td>0.658**</td>
<td>-0.071*</td>
<td>0.084**</td>
<td>-0.013</td>
<td>0.538*</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.032)</td>
<td>(0.031)</td>
<td>(0.279)</td>
<td>(0.039)</td>
<td>(0.033)</td>
<td>(0.032)</td>
<td>(0.291)</td>
</tr>
<tr>
<td>Conf. in past Recall</td>
<td>-0.027*</td>
<td>0.026*</td>
<td>0.001</td>
<td>0.258**</td>
<td>-0.013</td>
<td>0.014</td>
<td>-0.001</td>
<td>0.125</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.126)</td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.134)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.47</td>
<td>0.29</td>
<td>0.23</td>
<td>4.41</td>
<td>0.47</td>
<td>0.29</td>
<td>0.23</td>
<td>4.41</td>
</tr>
<tr>
<td>Observations</td>
<td>989</td>
<td>989</td>
<td>989</td>
<td>938</td>
<td>989</td>
<td>989</td>
<td>989</td>
<td>938</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.039</td>
<td>0.031</td>
<td>0.004</td>
<td></td>
<td>0.039</td>
<td>0.031</td>
<td>0.004</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. An interval regression is estimated in columns (4) and (8) using maximum willingness to pay as the dependent variable. In columns (1), (2), (3), (5), (6) and (7) an OLS regression is estimated using a dummy variable (=1) if the individual preferred the forecast information, 1 year information, or 10 years information as the dependent variable. Column (1) to (4) present the coefficients from the univariate regression between the dependent variable and each independent variable. Columns (5) to (8) present the coefficients from the multivariate regression.
Table 4: Effect of Reward Size on Information Choice and Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>Dummy Chose Forecast</th>
<th>Max. Willingness to Pay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Dummy High Reward</td>
<td>0.00449</td>
<td>0.00314</td>
</tr>
<tr>
<td></td>
<td>(0.0326)</td>
<td>(0.0325)</td>
</tr>
<tr>
<td>High Reward*College</td>
<td></td>
<td>0.0263</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0657)</td>
</tr>
<tr>
<td>High Reward*Std. Numeracy</td>
<td>0.0201</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0330)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.471***</td>
<td>0.469***</td>
</tr>
<tr>
<td></td>
<td>(0.0231)</td>
<td>(0.0231)</td>
</tr>
</tbody>
</table>

Notes: Heteroskedasticity-robust standard errors in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. In columns (1) to (3) an OLS regression is estimated using a dummy variable (=1) if the individual preferred the forecast information as the dependent variable. An interval regression is estimated in columns (4) to (6) using maximum willingness to pay as the dependent variable. The variable Dummy Less than College equals 1 if the level of the education of the individual is less than a Bachelor degree. The variable Std. Numeracy is the standardized variable of Numeracy Score with mean=0 and Std=1. It indicates the level of ability in numeracy (i.e., the greater the value of the variable, the greater the abilities in numeracy for the individual).
Table 5: Cost of Information and Dispersion of Expectations

<table>
<thead>
<tr>
<th></th>
<th>Low Price</th>
<th>High Price</th>
<th>P-value Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prior Mean</td>
<td>2.05 (0.088)</td>
<td>2.08 (0.099)</td>
<td>0.78</td>
</tr>
<tr>
<td>SD</td>
<td>1.95 (0.062)</td>
<td>2.08 (0.070)</td>
<td>0.16</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>3.85 (0.129)</td>
<td>3.62 (0.131)</td>
<td>0.37</td>
</tr>
<tr>
<td>Posterior Mean</td>
<td>3.15 (0.103)</td>
<td>2.84 (0.107)</td>
<td>0.04</td>
</tr>
<tr>
<td>SD</td>
<td>2.30 (0.073)</td>
<td>2.26 (0.076)</td>
<td>0.70</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>2.78 (0.111)</td>
<td>2.90 (0.121)</td>
<td>0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>495</td>
<td>443</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
Table 6: Net Effect of Information-Acquisition on Dispersion of Expectations

<table>
<thead>
<tr>
<th>Information Shown</th>
<th>Baseline Sample</th>
<th>Follow-Up Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Prior (1)</td>
<td>Posterior (2)</td>
</tr>
<tr>
<td>All</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=747 (419)</td>
<td>2.12 (0.073)</td>
<td>2.01 (0.052)</td>
</tr>
<tr>
<td></td>
<td>2.15 (0.084)</td>
<td>2.29 (0.059)</td>
</tr>
<tr>
<td></td>
<td>3.08 (0.127)</td>
<td>2.60 (0.090)</td>
</tr>
<tr>
<td>Forecast</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=354 (189)</td>
<td>2.19 (0.110)</td>
<td>2.07 (0.078)</td>
</tr>
<tr>
<td></td>
<td>3.26 (0.073)</td>
<td>1.36 (0.051)</td>
</tr>
<tr>
<td></td>
<td>3.35 (0.186)</td>
<td>2.56 (0.132)</td>
</tr>
<tr>
<td>1 Year Change</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=209 (124)</td>
<td>2.23 (0.139)</td>
<td>2.01 (0.099)</td>
</tr>
<tr>
<td></td>
<td>5.01 (0.161)</td>
<td>2.32 (0.114)</td>
</tr>
<tr>
<td></td>
<td>3.57 (0.254)</td>
<td>2.83 (0.181)</td>
</tr>
<tr>
<td>10 Year Change</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=184 (106)</td>
<td>1.86 (0.138)</td>
<td>1.87 (0.098)</td>
</tr>
<tr>
<td></td>
<td>0.85 (0.113)</td>
<td>1.53 (0.080)</td>
</tr>
<tr>
<td></td>
<td>2.02 (0.200)</td>
<td>2.06 (0.143)</td>
</tr>
<tr>
<td>Information Not Shown</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=242 (134)</td>
<td>1.93 (0.135)</td>
<td>2.10 (0.096)</td>
</tr>
<tr>
<td></td>
<td>2.48 (0.149)</td>
<td>2.32 (0.106)</td>
</tr>
<tr>
<td></td>
<td>2.79 (0.224)</td>
<td>2.60 (0.160)</td>
</tr>
<tr>
<td>Forecast</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=114 (69)</td>
<td>2.05 (0.192)</td>
<td>2.05 (0.137)</td>
</tr>
<tr>
<td></td>
<td>2.75 (0.217)</td>
<td>2.32 (0.155)</td>
</tr>
<tr>
<td></td>
<td>2.77 (0.309)</td>
<td>2.56 (0.221)</td>
</tr>
<tr>
<td>1 Year Change</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=81 (41)</td>
<td>1.77 (0.222)</td>
<td>2.00 (0.159)</td>
</tr>
<tr>
<td></td>
<td>2.27 (0.251)</td>
<td>2.26 (0.179)</td>
</tr>
<tr>
<td></td>
<td>3.03 (0.430)</td>
<td>2.76 (0.312)</td>
</tr>
<tr>
<td>10 Year Change</td>
<td>Mean</td>
<td>S.D.</td>
</tr>
<tr>
<td>N=47 (24)</td>
<td>1.90 (0.349)</td>
<td>2.39 (0.252)</td>
</tr>
<tr>
<td></td>
<td>2.17 (0.348)</td>
<td>2.39 (0.251)</td>
</tr>
<tr>
<td></td>
<td>2.42 (0.505)</td>
<td>2.47 (0.373)</td>
</tr>
</tbody>
</table>

Notes: The first number in N corresponds to the number observations in the baseline survey. The number in parentheses corresponds to the number of observations in the Follow-Up survey. In columns (1) and (2) we present the results for the Baseline Sample. In columns (3), the sample includes individuals who were invited and responded the Follow-up survey.
We will next be asking you about your expectations of nationwide home price changes.
As of December 2016, the value of the median or "typical" home in the US was **193,800** dollars (according to Zillow.com). Now, think about how the value of the typical home in the US has changed over time. (By value, we mean how much that typical home would approximately sell for.)

What do you think the value of such a home was 

*Please provide your best guess in each box below.*

**one year earlier** (in December 2015)? 193000 dollars

**ten years earlier** (in December 2006)? 190000 dollars

How confident are you in your answers?

*Please select only one.*

<table>
<thead>
<tr>
<th>Not at all confident</th>
<th>Somewhat confident</th>
<th>Very confident</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

© 2017 nielsen
We would now like you to think about the future value of the typical home in the US. As mentioned earlier, according to Zillow.com, the value of the typical home in the US was 193,800 dollars as of December 2016.

What do you think the value of the typical home in the US will be at the end of this year (in December 2017)?

*Please enter a number in the box below.*

dollars
We would now like you to think about the **future** value of the typical home in the US. As mentioned earlier, according to Zillow.com, the value of the typical home in the US was **193,800** dollars as of December 2016.

What do you think the value of the typical home in the US will be **at the end of this year** (in December 2017)?

*Please enter a number in the box below.*

![194000 dollars](image)

You said that you expect the value of a typical home in the US to be $194,000 at the end of this year. That is, you expect home prices to change by **0.10%** over the **course of the year 2017**.

*If not, please change your answer.*
You estimated the value of the typical home in the US to be 194,000 dollars at the end of this year. Now we want to ask you about how confident you are about this forecast.

What do you think is the percent chance (or chances out of 100) that the value of such a home at the end of this year (in December 2017) will be...

(Please note: The numbers need to add up to 100.)

<table>
<thead>
<tr>
<th>Price Range</th>
<th>Percent Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 174,600 dollars</td>
<td></td>
</tr>
<tr>
<td>Between 174,600 and 192,100 dollars</td>
<td></td>
</tr>
<tr>
<td>Between 192,100 and 195,900 dollars</td>
<td></td>
</tr>
<tr>
<td>Between 195,900 and 213,400 dollars</td>
<td></td>
</tr>
<tr>
<td>More than 213,400 dollars</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

© 2017 nielsen
Earlier in the survey, we asked you to forecast the value of a typical home in the US at the end of this year. Later in this survey, we will ask you to do so again.

This time, we will reward the accuracy of your forecast: you will have a chance of receiving $100. There is roughly a 10% chance that you will be eligible to receive this prize: we will select at random 60 out of about 600 people answering this question. Then, those respondents whose forecast is within 1% of the actual value of a typical US home at the end of this year will receive $100.

Your payment will depend on your answer, so consider this question carefully. You will be informed at the end of the survey if you have been chosen for this potential prize.
Before you report your forecast, you will have the opportunity to see only one of the following pieces of information that may help you with forecasting future year-ahead US home prices. Please rank the following pieces of information on a 1-4 scale, where 1 is "Highest ranked/Most Preferred" and 4 is the "Least Preferred".

*Please click on each piece of information on the left, and drag it to the right hand side of the screen.*

| Change in the value of a typical home in the US over the last one year (2016). | 1=Most preferred |
| Change in the value of a typical home in the US over the last ten years (2007-2016). | 2 |
| Forecasts of a panel of housing experts about the change in US home prices over this coming year (2017). | 3 |
| None of the above -- I would not like to see any information | 4=Least preferred |
You said that you would most prefer seeing information on the change in the value of a typical home in the US over the last one year (2016). Now we want to assess how much you would value this information.

You will next be presented with 11 scenarios. In each scenario, you will be given the choice of either seeing information about the change in the value of a typical home in the US over the last one year (2016) OR receiving extra money with the check that you will be getting for completing this survey. The amount of money that you will be offered in these scenarios is pre-determined, and goes from $0.01 to $5. For instance, in Scenario 1, you will need to choose between seeing information or receiving $0.01; and in Scenario 11, you will need to choose between seeing information or receiving $5.

We will draw one of these 11 scenarios at random for you. Your choice in the randomly chosen scenario will then be implemented. That is, you will have to make 11 choices, but only one of those choices will be implemented.

Since one scenario will be picked at random, your choices will not affect which scenario will be chosen.
You will now be asked to make a decision for each of the 11 scenarios.

**Scenario 1:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $0.01?

Note: if this scenario is chosen for you, your choice will be implemented. If you choose the information, you will see it on the next page. Instead if you choose the money, you will receive $0.01 in your check.

- [ ] see information  - [ ] receive $0.01

**Scenario 2:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $0.50?

- [ ] see information  - [ ] receive $0.50

**Scenario 3:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $1?

- [ ] see information  - [ ] receive $1

**Scenario 4:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $1.50?

- [ ] see information  - [ ] receive $1.50

**Scenario 5:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $2?

- [ ] see information  - [ ] receive $2

**Scenario 6:**
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $2.50?

- [ ] see information  - [ ] receive $2.50
 Scenario 7:  
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $3?

- see information  
- receive $3

 Scenario 8:  
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $3.50?

- see information  
- receive $3.50

 Scenario 9:  
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $4?

- see information  
- receive $4

 Scenario 10:  
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $4.50?

- see information  
- receive $4.50

 Scenario 11:  
Would you like to see information about the change in the value of a typical home in the US over the last one year (2016) OR receive $5?

- see information  
- receive $5
We would now like to ask you again about the future value of a typical home in the US at the end of this year.

Remember you will now have a chance of receiving $100 for the accuracy of your forecast. There is roughly a 10% chance that you will be eligible to receive this prize. About 600 people are answering this question, of whom 60 will be randomly picked for this potential prize.

If you are picked, you will receive $100 if your forecast is within 1 percent of the actual median home value in the US in December 2017 (according to the Zillow Home Value Index).

Your payment will depend on your answer, so consider this question carefully. You will be informed at the end of the survey if you have been chosen for this potential prize.
Scenario 1 was picked at random for you.

You had chosen to receive information about the change in the value of a typical home in the US over the last one year (2016).

According to the Zillow Home Value Index, the value of a typical home in the US increased by 6.8% over the last one year (December 2015 - December 2016 ). That means a typical home in the US that currently has a value of 193,800 dollars would have had a value of 181,500 dollars in December 2015. If home values were to increase at a pace of 6.8% next year, that would mean that the value of a typical home would be 206,978 dollars in December 2017.

Earlier in the survey, you reported that you thought the value of the typical home in the US at the end of this year (in December 2017) would be 194,000 dollars.

We would now like to ask you again about the future value of a typical home in the US at the end of this year.

© 2017 nielsen |  
According to the Zillow Home Value Index, the value of a typical home in the US increased by 6.8% over the last one year (December 2015 - December 2016). That means a typical home in the US that currently has a value of 193,800 dollars would have had a value of 181,500 dollars in December 2015. If home values were to increase at a pace of 6.8% next year, that would mean that the value of a typical home would be 206,978 dollars in December 2017.

Earlier in the survey, you reported that you thought the value of the typical home in the US at the end of this year (in December 2017) would be 194,000 dollars.

We would now like to ask you again about the future value of a typical home in the US at the end of this year.

What do you think the value of the typical home in the US will be at the end of this year (in December 2017)?

*Please enter a number in the box below.*

[Textbox for input]
According to the Zillow Home Value Index, the value of a typical home in the US increased by 6.8% over the last one year (December 2015 - December 2016). That means a typical home in the US that currently has a value of 193,800 dollars would have had a value of 181,500 dollars in December 2015. If home values were to increase at a pace of 6.8% next year, that would mean that the value of a typical home would be 206,978 dollars in December 2017.

Earlier in the survey, you reported that you thought the value of the typical home in the US at the end of this year (in December 2017) would be 194,000 dollars.

We would now like to ask you again about the future value of a typical home in the US at the end of this year.

What do you think the value of the typical home in the US will be at the end of this year (in December 2017)?

Please enter a number in the box below.

200000 dollars

You said that you expect the value of a typical home in the US to be $200,000 at the end of this year. That is, you expect home prices to change by 3.20% over the course of the year 2017.

If not, please change your answer.
You estimated the value of the typical home in the US to be 200,000 at the end of this year (in December 2017). Now we want to ask you about how confident you are about this forecast.

What do you think is the percent chance (or chances out of 100) that the value of such a home at the end of this year (in December 2017) will be...

*(Please note: The numbers need to add up to 100.)*

<table>
<thead>
<tr>
<th>Less than 180,000 dollars</th>
<th>percent chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between 180,000 and 198,000 dollars</td>
<td>percent chance</td>
</tr>
<tr>
<td>Between 198,000 and 202,000 dollars</td>
<td>percent chance</td>
</tr>
<tr>
<td>Between 202,000 and 220,000 dollars</td>
<td>percent chance</td>
</tr>
<tr>
<td>More than 220,000 dollars</td>
<td>percent chance</td>
</tr>
</tbody>
</table>

**TOTAL**

0
It was ok to refer to other sources (such as Google, Zillow, etc.) when taking the survey. Did you use any such sources when answering any question in the survey?

*Please select only one.*

- Yes
- No