Understanding Rationality and Disagreement in House Price Expectations

Zigang Li* Stijn Van Nieuwerburgh† Wang Renxuan‡

March 25, 2025

Abstract

Professional house price forecast data are consistent with a rational model where agents must learn about the parameters of the house price growth process and the underlying state of the housing market. Slow learning about the long-run mean generates overreaction to forecast revisions and a modest response of forecasts to lagged realizations. Heterogeneity in signals and priors about the long-run mean helps the model account for cross-sectional dispersion in forecasts. Introducing behavioral biases helps improve the model's predictions for short-horizon overreaction and dispersion. Using a cross-section of forecasters and a term structure of forecasts are crucial for inference.

JEL: C32, D84, G41, R32

*Rotman School of Management, University of Toronto, 105 St George St, Toronto, ON M5S 3E6; Email: zigang.li@rotman.utoronto.ca.

†Columbia University Graduate School of Business, NBER, and CEPR, 665 West 130th Street, New York, NY 10027; Email: svnieuwe@gsb.columbia.edu; Web: https://www0.gsb.columbia.edu/faculty/svannieuwerburgh/.

‡CEIBS, 699 Hongfeng Rd, Pudong, Shanghai, China, 201203; Email: rxwang@ceibs.edu.

The authors thank Pulsenomics for making their micro survey data available for academic research. One of the authors is a participant in the Pulsenomics survey. He has not received any compensation for this participation nor has Pulsenomics had any say in the content of this article. The authors thank Alina Arifeva, Nick Barberis, Pedro Bordalo, Kent Daniel, Hongye Guo (discussant), Zhiguo He, Lawrence Jin (discussant), Theresa Kuchler (discussant), Spencer Kwon (discussant), William Larson, Haoyang Liu (discussant), Sean Myers, Hoonsuk Park (discussant), Laura Veldkamp, Liyan Yang, Anthony Yezer, conference participants at the AREUEA National Meeting, the Sixth CEIBS Finance and Accounting Symposium, the NBER behavioral finance meeting, the Hoyt conference, the Finance Down Under Conference, the Adam Smith Workshop, WFA 2024 and seminar participants at the Federal Reserve Board, Frankfurt Business School, FISF, and Columbia Business School finance and macroeconomics. The authors thank Zheyang Zhu for excellent research assistance.

1 Introduction

Propelled by the increasing availability of survey data, an active literature studies how economic agents form expectations and how those expectations affect outcomes. Several features of these forecasts, their extrapolative nature or the large revisions to news, are at odds with full-information rational models.

The housing market is an important laboratory for this inquiry given the prominence of housing in households' portfolios and in the broader functioning of the macroeconomy, and given the role that expectations may play in driving aggregate mortgage credit extension and house price fluctuations.[1]

One interesting difference between house prices and many other prediction targets is their persistence. Even house price growth rates display substantial serial correlation, much more so than for example stock returns. This feature complicates inference on the rationality of house price beliefs from moments such as the dependence of beliefs on lagged house price growth.

Compared to forecasts of key macroeconomic and financial variables, house price expectations remain under-studied.[2] This paper uses a new and unique Pulsenomics survey data set that sheds new light on the features of professional forecasters' beliefs on house price growth. Such professional forecast data underpins decision making by mortgage lenders, brokers, asset managers, and builders.[3] Each quarter, a panel of around one hundred forecasters predict house price growth in the current and each of the next four calendar years. Our main finding is that the data on professional forecasts of house price appreciation is broadly consistent with the predictions of a rational learning model.

Our baseline model features a house price appreciation (HPA) process that contains a long-run mean, a persistent state, and a temporary shock. Forecasters do not know the state of the housing market nor the parameters that govern the HPA dynamics. They combine prior beliefs, data on past HPA, and private signals about current HPA to learn about both the latent state and the parameters. Forecasters are heterogeneous in the signals they receive and in their prior beliefs about long-run mean growth. We estimate the model parameters using Simulated Method of Moments and find that the combination of learning about the long-run mean and forecaster heterogeneity is crucial for explaining several important stylized facts

[1] Weber et al. (2022); D'Acunto and Weber (2024) make a strong case for the relevance of survey data in general. Bailey et al. (2018) show that house price expectations affect home purchase decisions, while Bailey et al. (2019) show that they affect mortgage leverage choice and mortgage rate. Chopra, Roth and Wohlfart (2023) find long-run home price expectations impact renters' spending. Favilukis, Ludvigson and Van Nieuwerburgh (2017); Landvoigt (2017); Kaplan, Mitman and Violante (2020); Greenwald and Guren (2021) debate the role of expectations and credit constraints in accounting for the boom and bust in U.S. house prices. Jacobson (2022) endogenizes beliefs in the Kaplan et al. (2020) model via adaptive expectations. Andersen et al. (2022) studies how own, past house purchases shape the willingness to transact. A large literature establishes that realized house price growth affects household consumption (e.g., Aladangady, 2017; Guren et al., 2021; Sodini et al., 2023; Vestman et al., 2023).

[2] Research on house price expectations using household surveys goes back to Case and Shiller (1988), but the scope of the data continues to improve (Kuchler, Piazzesi and Stroebel, 2023). We provide a detailed comparison of professional and household forecast data in Appendix E. This includes a discussion of Armona, Fuster and Zafar (2019).

[3] The Pulsenomics data is distributed via Zillow, which reaches 138 million unique visitors each month. In addition, Fannie Mae uses the survey results in its quartely updates on housing market trends, and Bank of America uses he data to create inernal housing models. The Pulsenomics survey receives substantial media coverage (`https://pulsenomics.com/media/`).

about HPA.

A first key moment for inference is the slope of a regression of forecast errors on the lagged forecast revision. Coibion and Gorodnichenko (2015, CG) estimate this slope using consensus, i.e., average forecast data on inflation. They argue that the positive coefficient estimate, which suggests underreaction to news, is evidence against full-information rational expectations (FIRE) and for limited- and sticky-information models. Bordalo et al. (2020) discover that for most macroeconomic and financial time-series, the slope of the CG regression is negative when estimated on individual—as opposed to consensus—forecast data. They propose diagnostic expectations to reconcile overreaction at the individual level with underreaction at the aggregate level. Ours is the first paper to study individual and consensus forecasts jointly in the housing context. We find the same sign reversal of the CG slope when using individual versus consensus forecast data. Furthermore, we show that the negative CG slope at the individual level is larger in absolute value for longer-horizon forecasts.[4]

The main result from our model is that slow learning about the unknown long-run mean of HPA can generate overreaction, i.e., a negative CG slope at the individual level. When forecasters have sufficiently diffuse priors about that long-run mean, they assign more weight to their signals when making forecasts. The resulting forecasts appear excessively volatile or to be "overreacting" to news about house price growth. This mechanism operates most clearly at forecast horizons beyond one year, but can manifest itself for some parameters even at the one-year horizon. These results echo findings in Collin-Dufresne, Johannes and Lochstoer (2016) that learning about long-run mean consumption growth is slow when the state of the economy is persistent.[5,6] Forecasters who simultaneously have to learn about multiple parameters and the state of the housing market face an identification problem of having to attribute noisy signals to various sources of uncertainty (Renxuan, 2020).[7] For the learning about the mean to be active, prior uncertainty about the long-run mean must be high enough. Indeed, in a special case of our model with no prior uncertainty about the long-run mean, a model akin to Bordalo, Gennaioli, Ma, and Shleifer (2020), the longer-run

[4] The experimental evidence in Afrouzi, Kwon, Landier, Ma, and Thesmar (2023) and the empirical evidence in Bordalo, Gennaioli, La Porta, and Shleifer (2019) also shows stronger overreaction for longer-horizon forecasts.

[5] Other notable models with parameter learning are Friedman (1979); Timmermann (1993); Lewellen and Shanken (2002); Cogley and Sargent (2008); Croce, Lettau and Ludvigson (2015); Kozlowski, Veldkamp and Venkateswaran (2020); Singleton (2021); Farmer, Nakamura and Steinsson (2023).

[6] Adam, Kuang and Xie (2024) decompose forecast revisions from the Survey of Professional Forecasters on 14 macroeconomic variables into components driven by prior beliefs, public news, and private signals. They find that forecasters tend to underreact to priors and public news but overreact to private signals. Their Bayesian learning model, in which forecasters learn about the latent state, attributes these patterns to overconfidence in the accuracy of private signals, aligning with our results on overconfidence. Their study focuses primarily on short-term forecasts and excludes parameter learning, whereas our study highlights that learning about the long-run mean is a critical mechanism for explaining overreaction in longer-horizon forecasts.

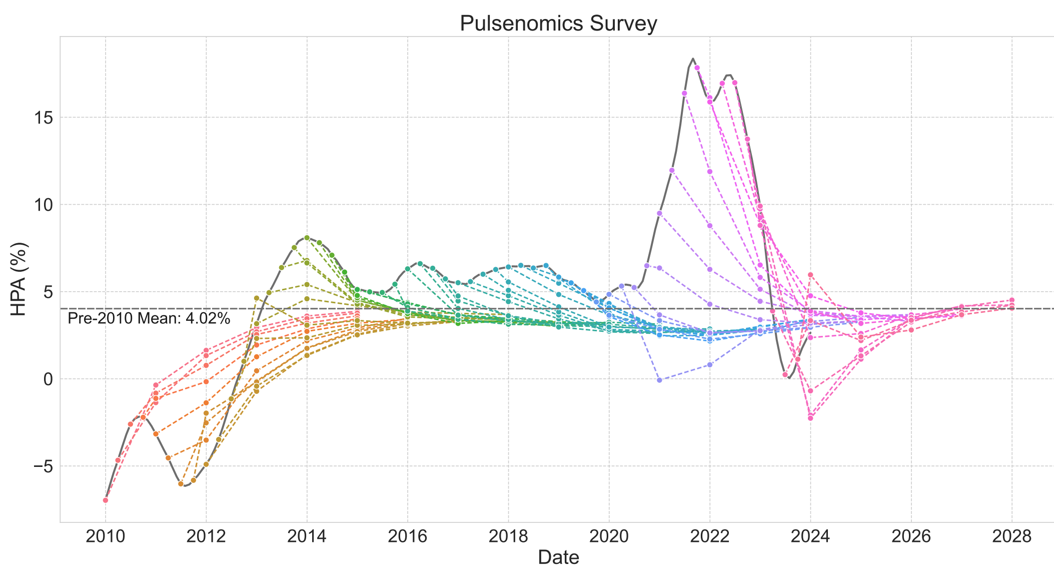
[7] An emerging literature investigates the sources of slow learning. Nagel and Xu (2022) propose that investors' fading memory when learning about long-run fundamentals can generate slow-learning, leading to a subjective equity premium. We return to the Nagel and Xu analysis below. Barberis and Jin (2023) draw on the psychology literature, suggesting "model-free" learning, where individuals infer from personal experiences (Malmendier and Nagel, 2011), even from their distant past, leading to slow-learning. Afrouzi et al. (2023) generate over-weighting of recent information that could result from memory decay or from frictions in information processing.

individual CG slopes are no longer negative.

A second key moment for inference is the sensitivity of forecasts to lagged HPA. Our professional forecasters increase their forecast in response to higher past house price growth. This evidence is consistent with that from household surveys.[8] The question whether extrapolation represents an optimal use of information is complicated by the fact that HPA displays substantial serial correlation. We find that forecasters under-extrapolate: their forecasts of HPA are less sensitive to lagged HPA than the sensitivity of actual realizations of HPA.[9]

Figure 1 visualizes this under-extrapolation. It plots the consensus HPA forecast for each horizon in the colored dashed line segments. The solid black line shows the realized house price growth. At each date, the forecast for annual HPA four years hence is close to the unconditional mean, indicating that the average forecaster expects full mean-reversion by year four. The actual data exhibit less mean-reversion.

Figure 1: Zillow/Pulsenomics Survey Forecasts of House Price Growth



Notes: The black solid line shows the realized annual HPA. Each short colored dashed line with six small dots shows the actual annual HPA (the first circle), the average survey forecast of annual HPA for the current calendar year (the second circle), and the average survey forecasts of annual HPA for the following four calendar years (the subsequent four circles).
Alt text: A graph showing the realized annual house price growth (solid black line) alongside the average survey forecasts for house price growth over different horizons (colored dashed lines), with each colored line segment representing forecasts for the current year and the next four years.

This second set of moments places tight restrictions on the model parameters. In particular, it points to

[8]Evidence that house price growth expectations are positively correlated with past realized house price growth goes back to Case and Shiller (1988). Recent contributions on extrapolation in household survey data includes Case, Shiller and Thompson (2012); Fuster, Laibson and Mendel (2010); Greenwood and Shleifer (2014); Barberis et al. (2015); Armona, Fuster and Zafar (2019); Glaeser and Nathanson (2017); Liu and Palmer (2021); Giglio et al. (2021*a*,*b*); Kuchler, Piazzesi and Stroebel (2023); Afrouzi et al. (2023).

[9]Regressions of forecast errors on lagged HPA generate the same conclusion. Kohlhas and Walther (2021) estimate similar regressions for GDP growth and inflation.

forecasters who have a tight prior around a modest value of persistence in the state of the housing market. If forecasters either had a higher prior for persistence or a more diffuse prior, the model-implied forecasts would be too sensitive to lagged HPA compared to the observed forecasts. As a result, the survey data leave little scope for learning about the persistence to play an important role in our model.

This result sets our paper apart from the mechanism in Farmer, Nakamura and Steinsson (2023), where downward-biased beliefs about the persistence parameter play a crucial role in accounting for the underreaction of consensus forecasts to news about interest rates.[10] We use our model and data to show that upward-biased beliefs about the persistence parameter can generate overreaction of individual forecasts to news about HPA, but only at the expense of forecasts that are overly sensitive to lagged HPA. That is, the Farmer et al. (2023) mechanism generates overreaction and over-extrapolation. It cannot generate the pattern in the HPA forecast data of overreaction and under-extrapolation.

The third key moment is the cross-sectional dispersion in forecasts. The dispersion in forecasts of HPA is large; it exceeds the time-series variation in the average forecast. A novel stylized fact is that the dispersion shrinks in the horizon of the forecast.

Our model accounts for the dispersion facts by allowing for both signal heterogeneity and heterogeneity in the prior about the long-run mean HPA.[11] Signal heterogeneity helps, but only to a point. It has limited bite in a world where rational forecasters are aware that their signals contain substantial noise and discount them. Heterogeneity in prior beliefs about the long-run mean has a first-order impact on forecast dispersion, as long as the prior uncertainty about the long-run mean is not too high. If it is too high, forecasters assign little weight to their priors when forming posteriors and heterogeneity in the prior fades away.[12]

In addition to these three key moments, the model matches the average forecast error in HPA expectations, as well as the weak sensitivity of HPA forecasts to contemporaneous HPA realizations. The latter soundly rejects the full-information rational expectations (FIRE) prediction that the forecast and the truth should move one-for-one on average (Mincer and Zarnowitz, 1969).

Taken together, the rational learning model accounts for a comprehensive set of moments of HPA fore-

[10]Famer et al. (2023) do not model individual forecasters, but study consensus forecasts. Hence, their model cannot speak to the cross-sectional dispersion in forecasts. Our model nests their setup. Gabaix (2019) and Angeletos, Huo and Sastry (2021) also study biases arising from forecasters using an incorrect value of the persistence parameter.

[11]A literature studies the (term structure of) disagreement among forecasters (e.g., Mankiw, Reis and Wolfers, 2003; Lorenzoni, 2009; Patton and Timmermann, 2010; Giacoletti, Laursen and Singleton, 2021). Andrade et al. (2016) document that disagreement persists for forecasting horizons over five years and that the term structure of disagreement takes different shapes for different macro variables. They propose a sticky information model in a multi-variate setting. Kindermann et al. (2021) show that Bayesian learning with heterogeneous information between renters and owners can explain the dispersion of cross-sectional forecasts during the recent house price boom in Germany. Goldstein and Gorodnichenko (2023) consider signal heterogeneity about future information, which differentially affects the dispersion of forecasts at different horizons.

[12]Signal heterogeneity can be micro-founded from models of optimal information acquisition (Van Nieuwerburgh and Veldkamp, 2009, 2010; Kacperczyk, Van Nieuwerburgh and Veldkamp, 2016; Mackowiak and Wiederholt, 2009) or result from heterogeneous exposure to information (e.g., Kuchler and Zafar, 2019, on local house prices), how information is delivered (Malmendier and Veldkamp, 2022), social network ties (Bailey et al., 2018, 2019), or lived experience (Malmendier and Nagel, 2011, 2016).

casts which have traditionally been interpreted as *prima facie* evidence against the rational expectations model. These moments, which exploit the time-series and cross-sectional dimensions of our survey data, tightly identify the parameters that govern the expectations formation process. They point to the role of learning about an unknown long-run mean of HPA, low perceived persistence of the housing market state, and significant heterogeneity in signals about future HPA and especially in priors about the long-run mean.

Four additional pieces of evidence from our Pulsenomics panel data bolster the case for our learning explanation. Since learning is slow and forecasters have substantial prior heterogeneity, the model predicts that forecast errors should be persistent over time at the individual forecaster level. We find that they are. Second, we find that the most pessimistic (optimistic) forecasters in the data make similar forecasts to their counterparts in the model. This is evidence that the amount of prior heterogeneity we estimate for the baseline model is not excessive. Third, we find evidence that among a wide array of forecaster characteristics, hand-collected from Linked-IN, the length of professional experience is associated with lower forecast errors. We show the same is true in the model, where experience is defined as the number of periods of learning. Fourth, we find lower forecast dispersion among more-experienced than among less-experienced forecasters, consistent with the model.

The baseline learning model falls short in generating (i) a negative enough CG coefficient at the one-year forecast horizon and (ii) enough forecast dispersion at the one-year horizon. To remedy these two shortcomings, we enrich the baseline model with one of two seminal concepts from behavioral economics: over-confidence and diagnostic expectations. An overconfident forecaster has a subjective signal precision that exceeds the objective one. Following Bordalo et al. (2020), a forecaster with diagnostic expectations forms distorted estimates for the mean-reverting state of the housing market. We show that either over-confidence or diagnostic expectations help close the gap between model and data for the one-year individual GC slope and the one-year forecast dispersion. In contrast, the effect from these behavioral frictions on longer-horizon forecasts is small. Rational learning about the long-run mean is what helps the model match the longer-horizon CG slope and forecast dispersion. The enriched model is able to account for the full set of expectations moments. It helps clarify which moments are most revealing of behavioral frictions and quantify those frictions. We arrive at a compelling synthesis of the extent of rationality and disagreement in house price forecasts. In contrast, a version of our model that shuts down learning but keeps the behavioral frictions is unable to account for the facts.

To provide additional intuition for parameter identification and to highlight the uniqueness of our data, we estimate our model on two subsets of moments: one where the only available data is on consensus forecasts and one where there only is data on short-horizon forecasts. Indeed, most survey data sets either lack multiple forecasts horizons or a cross-section of forecasters. Unsurprisingly, in both cases we obtain a worse fit to the data. But the nature of the failure is instructive. The model that is only confronted with consensus data estimates a much lower prior uncertainty parameter for long-run mean HPA. This effectively shuts down the main mechanism of learning about the long-run mean. As a result, this model fails to generate the negative CG slopes at the individual level. The model that is only confronted with the one-year ahead forecast moments similarly estimates lower prior uncertainty for the long-run mean HPA but also higher prior uncertainty for the persistence of the state. The latter uncertainty activates learning about persistence and results in a higher posterior estimate for that persistence, leading to forecasts that are excessively sensitive to lagged HPA. Reliably inference requires rich survey data.

Zooming out from HPA forecast data, we explore how common the combination of overreaction and under-extrapolation is among a broad set of macroeconomic variables using Survey of Professional Forecasts data. We find it is the most common combination, showing that the empirical pattern in the HPA data is of broader interest to the belief formation literature. Equally importantly, our learning model can generate not only the co-occurrence of overreaction and under-extrapolation but also of overreaction and over-extrapolation, and of underreaction and under-extrapolation, for different parameter values. The empirical pattern is valuable for parameter identification.

The rest of this paper is organized as follows. Section 2 introduces the Pulsenomics house price forecast data and documents key empirical facts on professional house price growth forecasts across time and forecasters. Section 3 sets up the model and describes the estimation algorithm. The results for the benchmark learning model are in Section 4. Section 5 provides direct support from the microdata for the learning explanation. Section 6 incorporates overconfidence and diagnostic expectations into the benchmark model. Section 7 explains the importance of having a panel dimension and a term structure of forecasts in order to correctly identify the model parameters and account for the key moments of the survey data. Section 8 studies extrapolation and reaction coefficients in other macro and finance series. Section 9 concludes. The Appendix contains additional empirical results, model derivations, and model results.

2 Facts About Professional House Price Growth Forecasts

Using a unique data set of professional forecasters collected by the forecasting firm Pulsenomics, we document several important empirical facts about HPA forecasts, including time series variation, cross-sectional dispersion, sensitivity to current and past HPA realizations, and average bias.

2.1 Data

The forecast data we use come from the Home Price Expectations Survey conducted by Pulsenomics, an independent research and index product development firm.[13] The sample period of our survey data is 2010Q1 to 2023Q4 (56 quarters). Every quarter, more than 100 professionals are surveyed to predict the year-by-year home price growth over a five-year horizon. This diverse expert panel consists of professionals and economists from industry, policy, and academia. The survey for each quarter is conducted in the middle of the second month of that quarter, e.g., the 2023Q1 survey is conducted in the middle of Feb 2023.[14]

The survey data is unique in the following ways. First, it is a panel data set. Each individual forecaster can be tracked over time. This allows us to study belief formation at both the individual and the consensus levels. Second, it contains a term structure of forecasts with relatively long horizons (from current year to 4 years ahead). Information from both the cross-section and the term structure of expectations turns out to be very informative for estimating the belief formation process. Third, the number of forecasters each period is large compared to other surveys such as the Survey of Professional Forecasters.

Appendix E compares our professional survey expectations with two common household survey expectations measures. We show the two household HPA series are highly correlated with our professional HPA series. Household forecasts show similar "pulling the longer-run forecast to the long-run mean" as in Figure 1. We confirm the under-extrapolation using regression analysis, and compare anomaly moments discussed below. The belief formation process of professionals is of independent interest given (i) their expertise, (ii) the prevalence of professional survey data in economics and finance, and (iii) the fact that many individuals rely on professionals.[15]

The data we use for the house price index is the Zillow Home Value Index (ZHVI). This index is calculated by Zillow based on more than 100 million U.S. homes. Since forecasters in the Zillow Home Price Expectations Survey are asked to predict the changes in the ZHVI, we naturally use this index as our data source for the house price index. The index is available monthly and data for a given month is published on the third Thursday of the following month. This data series starts in 1996.[16] House prices display large swings over our 1996–2023 sample, with booms from the late 1990s until 2006, from 2012 until 2019, and from the middle of 2020 until 2023, and busts from 2007 until 2011 and in 2020. This range of outcomes should make this a good period to learn about the parameters of the data-generating process for HPA.

[13] The Wall Street Journal runs a similar survey, but we are not aware of any analysis like ours with that data set.

[14] More details about the timing of the survey and the data availability can be found in Appendix A.

[15] Indeed, Fuster et al. (2022) show that, when households are given the opportunity to pick among different sources of information to help predict future house price changes, about half chose the forecast of housing experts.

[16] In the last quarter of our sample, 2023Q4, the forecast target changes to the Fannie Mae non-seasonally adjusted House Price Index. Since the old and new target are extremely highly correlated, we ignore this minor change.

2.2 Time-series and Cross-section of Forecasts

Consistent with the findings from household surveys (Kuchler, Piazzesi and Stroebel, 2023), there is significant time-series variation in professional HPA forecasts. The magnitude of the variation in forecasts is smaller than the variation in realizations of HPA and the variation in long-term forecasts is smaller than the variation in short-term forecasts.[17] This finding is consistent with Figure 1, which showed that long-term forecasts barely fluctuate around the long-run mean.

There also is substantial dispersion across forecasters. The cross-sectional variation in forecasts exceeds the time-series variation in average forecasts of house price growth. The dispersion is higher for short-run than for long-run HPA forecasts.[18] The literature makes clear that there is a lot of unexplained heterogeneity in individual forecasts. This opens the door for unobservable heterogeneous signals and priors across forecasters. Below, we will use the individual forecasts to build additional support for the learning mechanism, taking advantage of our panel data.

2.3 Forecast Anomalies

As motivation for the model, we document several key empirical facts in our HPA forecast data. These empirical patterns are typically interpreted as evidence that expectations depart from full-information rationality (FIRE).

First, we show that HPA forecasts are sensitive to past HPA realizations. We estimate the sensitivity coefficients from:

$$E_{it}[y_{t+h}] = a + by_{t,t-j} + \varepsilon_{i,t+h}, \tag{1}$$

where y_{t+h} denotes one-year HPA h years from now,[19] $y_{t,t-j}$ is lagged annualized house price growth over the past j years, and $E_{it}[y_{t+h}]$ the individual forecast of that future HPA. Panel A of Table 1 shows the results. Specifically, short-term forecasts are strongly positively correlated with past HPA realizations, while long-term forecasts are insensitive to past HPA. Panel B shows the same correlations for realized HPA in the same period 2010–2022, replacing $E_{it}[y_{t+h}]$ on the left-hand side of (1) with the forecast error $y_{t+h} - E_{it}[y_{t+h}]$.[20] The forecast error also shows strong positive sensitivity to past HPA realizations. This indicates that professional forecasters "under-extrapolate" at all future horizons $h = 1, 2, 3, 4$. Put differently, they

[17] See Appendix Figure A2.

[18] Appendix Figure A3 shows the cross-sectional dispersion of HPA in the 2023.Q4 survey for 2024 HPA (left panel) and for 2027 HPA (right panel).

[19] The Pulsenomics Survey asks the participants to make forecasts for calendar-year price growth. Here we simplify the notation for ease of discussion. Appendix A details how we compute the forecasts that respect the exact structure and timing of the Pulsenomics survey instrument.

[20] The sum of the coefficients in panels A and B are close to the persistence of house price appreciation, once one considers the correct average forecast horizon in each of the columns. For example, the average forecast horizon for the current year (1-year ahead) forecast is 2.5 (6.5) quarters.

Table 1: Effect of Past HPA on Forecasts and Forecast Errors

	Panel A: HPA Forecasts									
	Current Year		1-Year Ahead		2-Years Ahead		3-Years Ahead		4-Years Ahead	
Past 1Y HPA	0.538***		0.165***		0.046**		0.015		0.002	
	(0.048)		(0.042)		(0.018)		(0.011)		(0.011)	
Past 4Y Ann. HPA		0.405***		0.102**		0.022		0.010		0.007
		(0.080)		(0.049)		(0.021)		(0.014)		(0.013)
Num. of Obs.	5,868	5,868	5,868	5,868	5,868	5,868	5,868	5,868	5,868	5,868
R-Squared	0.523	0.273	0.101	0.036	0.013	0.003	0.002	0.001	0.000	0.000
	Panel B: HPA Forecast Errors									
	Current Year		1-Year Ahead		2-Years Ahead		3-Years Ahead		4-Years Ahead	
Past 1Y HPA	0.184***		0.081		-0.060		0.143*		0.266***	
	(0.063)		(0.106)		(0.091)		(0.085)		(0.078)	
Past 4Y Ann. HPA		0.167**		0.181		0.144		0.240**		0.346***
		(0.066)		(0.112)		(0.105)		(0.097)		(0.109)
Num. of Obs.	5,868	5,868	5,503	5,503	5,117	5,117	4,737	4,737	4,352	4,352
R-Squared	0.090	0.069	0.010	0.040	0.005	0.029	0.026	0.085	0.089	0.170

Notes: Panel A shows the relationship between HPA forecasts and past HPA. Panel B shows the relationship between forecast errors and past HPA. The coefficients are estimated from individual-level panel regressions. All past HPA variables are annualized. Standard errors are clustered at the quarter and forecaster level. Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

over-estimate the degree of mean reversion in HPA. The long-run forecasts are "pulled to the mean" too strongly compared to the actual HPA data, as also shown in Figure 1.

Second, individual HPA forecasts have persistent errors, $e_{i,t+h|t} \equiv y_{t+h} - E_{it}[y_{t+h}]$. The average bias in the panel data set of forecasters can be estimated from the regression:

$$e_{i,t+h|t} = a^{fb} + \varepsilon_{i,t+h}, \tag{2}$$

Panel A of Table 2 shows that forecasts are lower than average HPA realizations during our 2010–2023 sample, so that the bias a is large and positive.

Third, we estimate (inverse) Mincer-Zarnowitz (MZ) regressions of forecasts on contemporaneous realizations:

$$E_{it}[y_{t+h}] = a^{mz} + b^{mz} y_{t+h} + \varepsilon_{i,t+h}. \tag{3}$$

FIRE implies that the truth and the forecast move one-for-one, i.e., $b^{mz} = 1$ in the above regression.[21] We find that forecasts move less than one-for-one with the truth. Panel B of Table 2 shows slope estimates close

[21] The standard Mincer and Zarnowitz (1969) regression is defined as: $y_{t+h} = \widetilde{a}^{mz} + \widetilde{b}^{mz} E^c_t[y_{t+h}] + \widetilde{\varepsilon}_{i,t+h}$. In an individual-level panel regression, the dependent variable y_{t+h} is identical for every individual in a single cross-section. This causes the error terms to be correlated with the regressors. The inverse MZ regression does not have this issue. Appendix F.1 provides a detailed comparison between the estimation of MZ and inverse MZ regressions in the panel regression context.

Table 2: Other Forecast Anomalies

	Forecast Horizon				
	Current Year	**1-Year Ahead**	**2-Years Ahead**	**3-Years Ahead**	**4-Years Ahead**
Panel A: Bias					
a^{fb}	1.293***	2.952***	4.025***	4.299***	4.021***
	(0.362)	(0.584)	(0.550)	(0.556)	(0.607)
Num. of Obs.	5,868	5,503	5,117	4,737	4,352
Panel B: Inverse MZ Regressions					
b^{mz}	0.622***	0.130***	-0.064***	-0.051***	-0.057***
	(0.062)	(0.050)	(0.018)	(0.013)	(0.013)
Num. of Obs.	5,868	5,503	5,117	4,737	4,352
R-squared	0.511	0.039	0.010	0.009	0.014
Panel C: CG Regressions					
b^{cg}	-0.250**	-0.323*	-0.570***	-0.576***	-0.411***
	(0.121)	(0.178)	(0.066)	(0.013)	(0.061)
Num. of Obs.	5,050	4,728	4,407	4,065	2,881
R-squared	0.032	0.017	0.036	0.033	0.015

Notes: This table shows the regression estimates of other forecast anomalies at different horizons. Panel A shows the estimates of bias in Equation 2. Panel B shows the estimates of inverse MZ regression specified in Equation 3. Stars indicate the level of statistical significance in testing the hypothesis: $b = 1$. Panel C shows the estimates of CG regression in Equation 4. All coefficients are estimated from individual-level panel regressions. Standard errors are in parentheses. Standard errors are clustered at the quarter and forecaster level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

to zero for $h = 1$ and even negative for $h = 2, 3, 4$.

Fourth, and most importantly, individual forecast revisions negatively predict future forecast errors. We estimate the CG regressions (Coibion and Gorodnichenko, 2015) at the individual forecaster level:

$$e_{i,t+h|t} = a^{cg} + b^{cg} \left(E_{it}[y_{t+h}] - E_{it-1}[y_{t+h}] \right) + \varepsilon_{i,t+h}, \tag{4}$$

These regressions test the null of FIRE ($b^{cg} = 0$) against the alternative of limited- or sticky-information models. If $b^{cg} > 0$, the forecasts suffer from "underreaction," in the sense that an increase in the forecast leaves the new forecast still too low on average. If $b^{cg} < 0$, the forecasts suffer from "overreaction." Panel C of Table 2 finds that professional forecast errors are negatively correlated with forecast revisions, and more so at longer horizons. Professional house price forecasts suffer from overreaction. Below, we revisit the CG regressions estimated at the consensus level, and observe a sign reversal relative to the individual-level data. Our paper is the first to estimate the CG regressions for house prices, and to do so at multiple forecast horizons. We observe "overreaction" to HPA news at the individual level which is substantially stronger at longer horizons than at the (more commonly-investigated) short horizon.

Appendix Table A4 shows that the forecast anomalies are similar when including forecaster fixed effects.

3 Model

We develop a model to account for and shed light on these empirical results. The model aims to simultaneously match the cross-section of individual house price forecasts and the heterogeneity across forecasters. It provides a rich laboratory to quantitatively assess the importance of imperfect information, learning, and behavioral biases in shaping expectations.

3.1 Setup

The key assumption of the model is that forecasters do not know and must learn the underlying state of the housing market and the parameters that govern the data generating process (DGP) of the HPA process. Our model nests that of Bordalo et al. (2020) and Farmer, Nakamura and Steinsson (2023) who study macroeconomic and financial time series. Unlike Bordalo et al. (2020), who also assume the presence of a latent state, we assume that forecasters additionally do not know the underlying parameters of the DGP and in particular the long-run mean. They need to estimate the parameters in order to forecast HPA. Unlike Farmer, Nakamura and Steinsson (2023), who also feature both latent state and parameter learning, we introduce forecasts at multiple horizons and disagreement among forecasters by allowing different forecasters to receive heterogeneous noisy signals. Farmer et al. (2023) work with consensus, not individual forecast data. Our estimation favors a different mechanism from that in Farmer et al.

House Price Growth Process We assume that HPA y_t, measured by log changes in the ZHVI, contains a long-term mean, $\bar{\mu}$, a mean-reverting component x_t, and a noise term e_t:

$$y_t = \overline{\mu} + x_t + \sqrt{\overline{\gamma}\overline{\sigma}^2} e_t \quad e_t \sim N(0,1) \tag{5}$$

$$x_t = \overline{\rho} x_{t-1} + \sqrt{(1-\overline{\gamma})\,\overline{\sigma}^2}\eta_t \quad \eta_t \sim N(0,1) \tag{6}$$

where $\overline{\gamma}$ governs the proportion of variance in total house price growth, $\overline{\sigma}^2$, that comes from the noise component e_t. The noise e_t term captures the fact that aggregate house price growth data from Zillow is a noisy measure of the true HPA. Indeed, Zillow revises its own historical past realized HPA as new data becomes available.[22] Through the persistent state variable x, the model captures the fact that HPA is

[22] As one example, Zillow began reporting the "Raw/MidTier" as its "headline" index series as of August 2022, instead of "smoothed, seasonally adjusted" series. While the correlation between these two series is above 99%, they are not the same.

serially correlated, unlike price growth in many other financial time series, and may exhibit mean-reversion. The lines over the parameters indicate that these are the true parameters that govern the objective DGP, unknown to the forecasters.

Forecasters' Heterogeneous Signal Extraction Process Each quarter, the Pulsenomics Survey asks its 100 participants to make forecasts for HPA over the current calendar year and each of the next four calendar years. Let T denote the current quarter during which forecasters are asked to fill out the survey. Each forecaster observes (is shown) the realized HPA up to previous quarter ($T-1$) when making forecasts for future HPA, but does not know the HPA in the current quarter T.

Each forecaster i at time T has information set $S_{i,T}$, defined as

$$S_{i,T} := (s_{i,1}, \ldots, s_{i,T-1}, s_{i,T})' \tag{7}$$

The information set contains all historical realized HPA and a private signal about current and future HPA, or

$$s_{i,t} = \begin{cases} y_t & t \leqslant T-1 \\ y_T + \alpha_{i,T}, \quad \alpha_{i,T} \sim N(0, \overline{\Gamma}^2 \overline{\sigma}^2) & t = T \end{cases} \tag{8}$$

The parameter $\overline{\Gamma}$ denotes *objectively* how much noise (variance of $\alpha_{i,T}$) is in forecasters' information signal. The signal noise $\alpha_{i,T}$ is assumed to be i.i.d. across investors and time.[23] The parameter Γ (without the $\bar{\cdot}$) captures forecasters' subjective beliefs about the amount of noise in their signals. When $\Gamma = \overline{\Gamma}$, which is our benchmark case, forecasters correctly understand the true amount of noise that their signals contain.

Different information sets $S_{i,T}$ lead forecasters to disagree about the current and future HPA, as reflected in their house price forecasts. One can interpret the differences in signals as different forecasters using different predictors to estimate HPA or paying different amount of attention to the latest developments in housing markets.[24]

Forecasters' Subjective Expectation Formation Process Based on their priors, past realized HPA, and their own individual signals, the forecasters make inference about the values of the parameters and the latent state of the housing markets. They correctly understand the underlying structure of the D.G.P. in Equation (5) and (6), and the presence of noise in their signals as in Equation (8). However, they do not

[23] $\overline{\Gamma}$ is unrelated to $\overline{\gamma}$. The latter captures the proportion of the noise in the realized HPA that is short-run.
[24] See footnote 12 for different possible underlying sources of signal heterogeneity.

know the true values of the underlying parameters and states. They need to make inference about the parameters ($\overline{\mu}$, $\overline{\rho}$, $\overline{\sigma}$, $\overline{\gamma}$) and the latent state $\{x_t\}_{t=0}^T$ in order to form expectations about future HPA based on their information set $S_{i,T}$. They are Bayesian learners. We assume in our benchmark case that the forecasters know the value of $\overline{\Gamma}$, the objective signal volatility.

From the perspective of the forecasters, the learning process takes the following form:

$$s_{i,t} = \mu_{i,t} + \widetilde{x}_{i,t} + \widetilde{e}_{i,t}, \qquad \widetilde{e}_{i,t} \sim N(0, H_{i,t}) \tag{9}$$

$$\widetilde{x}_{i,t} = \rho_{i,t}\widetilde{x}_{i,t-1} + \sqrt{1-\gamma_{i,t}}\sigma_{i,t}\widetilde{\eta}_{i,t}, \qquad \widetilde{\eta}_{i,t} \sim N(0,1) \tag{10}$$

where

$$H_{i,t} = \begin{cases} \gamma_{i,t}\sigma_{i,t}^2, & t \leqslant T-1 \\ (\gamma_{i,t} + \Gamma^2)\sigma_{i,t}^2, & t = T \end{cases} \tag{11}$$

Because of the heterogeneity in forecaster information, the parameter estimates ($\mu_{i,t}$, $\rho_{i,t}$, $\sigma_{i,t}$, $\gamma_{i,t}$) are heterogeneous across forecasters, as indicated by the subscript "i" (Patton and Timmermann, 2010). The state estimates, $\widetilde{x}_{i,t}$, are also heterogeneous across forecasts and subjective, as indicated by " $\widetilde{\cdot}$ ". The disagreement across forecasters is bounded as all of the forecasters observe the same history of realized HPA up until the previous quarter, or $s_{i,t} = y_t, \forall t = 1, ..., T-1$.

Due to parameter uncertainty and learning (Collin-Dufresne, Johannes and Lochstoer, 2016), the parameter estimates ($\mu_{i,t}$, $\rho_{i,t}$, $\sigma_{i,t}$, $\gamma_{i,t}$) are time-varying, because they are the posterior mean estimates given the realized data up until that point, signals, and prior distributions.

Forecasters' Prior Beliefs We assume the following prior distributions for the parameters:

$$\mu_0 \sim N(\mu_{\mu,i}, \sigma_\mu^2), \quad \rho_0 \sim N(\mu_\rho, \sigma_\rho^2), \quad \sigma_0^2 \sim IG(A, B), \quad \gamma_0 \sim Beta(C, D)$$

These priors are common across forecasters, except for the prior belief about the long-run mean HPA. Forecasters prior long-run mean HPA are uniformly distributed over the range $\mu_{\mu,i} \sim [\mu_\mu - \frac{1}{2}\lambda, \mu_\mu + \frac{1}{2}\lambda]$, where the parameter λ captures the degree of prior heterogeneity.

The forecasters in the model know the parameters that govern the prior distributions: $(\mu_\mu, \sigma_\mu^2, \lambda, \mu_\rho, \sigma_\rho^2, A, B, C, D)$. We start the forecasters off in 1996.Q1 with these priors. By the time the survey data comes in in 2010.Q1, they have already learned for 56 quarters.[25]

[25] In an exercise detailed in Section 5, we allow each forecaster to learn for a different number of periods (i) prior to entering the

Upon receiving new signals, forecasters update their prior beliefs to arrive at posterior beliefs for the parameters ($\mu_{i,t}$, $\rho_{i,t}$, $\sigma_{i,t}$, $\gamma_{i,t}$) using Bayes Rule. Given these new estimates for parameters, they estimate the underlying mean-reverting state $\tilde{x}_{i,t}$ using Kalman filtering and smoothing. We derive analytical expressions for the posterior means based on the marginal distributions of the parameters. We then perform Gibbs-Sampling to combine the marginal posteriors to arrive at the posterior mean of the forecasts for HPA. Appendix B contains the details.

Given their posterior parameter and state estimates, the forecasters form their forecasts for the h-quarter ahead house price growth $E_{i,t}(y_{t+h})$ given their long-run mean HPA estimate ($\mu_{i,t}$), persistent parameter estimate ($\rho_{i,t}$), and estimate of the state ($\tilde{x}_{i,t}$). Appendix D details how these expectations are formed.

Incorporating Behavioral Biases to the Expectation Formation Process Following a literature on overconfidence (Daniel and Hirshleifer, 2015), we introduce an additional parameter $\phi := \overline{\Gamma}/\Gamma$. When $\phi > 1$, the objective signal noise is larger than the subjective signal noise. Forecasters' perceived signal precision is greater than the actual signal precision, i.e. they are overconfident about the precision of their signals. Conversely, when $\phi < 1$, forecasters are under-confident. The benchmark model sets $\phi = 1$.

Following the literature on diagnostic expectations (Bordalo et al., 2020), we allow forecasters to have potentially distorted estimates of the state of the housing market, $\hat{x}_{i,t}$, when applying the Kalman filter, given their posterior estimates for the parameters. The degree of distortion is governed by the parameter θ, with a positive (negative) θ leading to an overreaction (underreaction) to news. When $\theta = 0$, forecasters apply the optimal Kalman gain and we recover our benchmark rational learning model. The technical details on how to incorporate diagnostic expectations in our learning model are in Appendix C. We incorporate overconfidence and diagnostic expectations into our learning model and examine their impact on the model's fit to the data in Section 6.

Forecasts of Others The model assumes that forecasters do not use information on the forecasts of others. Appendix G provides empirical support for this assumption, by showing that the lagged consensus forecast does add explanatory power to the individual forecasts, and additional discussion.

3.2 Benchmark Model Estimation

We, the econometricians, do not know the parameters of the prior distributions and must estimate them. We estimate the six key parameters, $\delta = (\mu_\mu, \sigma^2_\mu, \mu_\rho, \sigma^2_\rho, \Gamma, \lambda)$ that govern our benchmark rational learning

survey, based on the length of their career prior to entering the survey, and (ii) since entering the survey, based on their survey entry and exit dates. We assign each forecaster heterogeneous priors based on the HPA realizations experienced prior to entering the survey. The results are similar to those in the benchmark model.

model using Simulated Method of Moments (SMM). We calibrate the parameters A, B, C, D that govern the prior distributions of σ_0^2 and γ. In the extensions, we vary the parameters ϕ and θ to evaluate the impact of behavioral biases in forecasts. The latter parameters are set to one and zero, respectively, for our baseline analysis. Our estimation aims to match salient features of the individual forecasters' panel data set as well as forecast dispersion, 36 moments in all. The technical details of the SMM estimation are relegated to Appendix D. Given its computational complexity, the model is estimated on a supercomputer using Microsoft Azure.

Given a candidate parameter vector δ, we proceed as follows

1. Obtain N forecasters' heterogeneous signals by
   (a) simulating normally distributed private signals $\alpha_{i,t} \sim N(0, \bar{\Gamma}^2\bar{\sigma}^2), \quad i = 1, ..., N$
   (b) and adding them to the current quarter's realized HPA: $s_{i,t} := y_t + \alpha_{i,t}$.
2. Perform Bayesian learning for each forecaster i, using $s_{i,t}$ as inputs, and prior parameters $(\mu_\mu, \sigma_\mu^2, \mu_\rho, \sigma_\rho^2)$. This leads to individual forecasters' HPA prediction, $E_{i,t}(y_{t+h})$, for different horizons h and periods $t = 1, ...T$.
3. Construct the 36 model-implied moments based on these forecasts $E_{i,t}(y_{t+h})$ as follows:
   (a) Estimate panel regressions based on individual forecasts $E_{i,t}(y_{t+h})$ and actual realized HPA y_{t+h}. We use four forecasting horizons for each regression moment: $h = 1, 2, 3, 4$.
      i. Bias: Regress forecast error $y_{t+h} - E_{i,t}(y_{t+h})$ on a constant; collect mean estimates (1 moment per horizon).
      ii. Coibion-Gorodnichenko regression: Regress forecast error $y_{t+h} - E_{i,t}(y_{t+h})$ on a constant and the forecast revision $E_{i,t}(y_{t+h}) - E_{i,t-1}(y_{t+h})$; collect intercept and slope (2 moments per horizon).
      iii. Inverse Mincer-Zarnowitz regression: Regress forecast $E_{i,t}(y_{t+h})$ on a constant and the realization y_{t+h}; collect intercept and slope (2 moments per horizon).
      iv. Forecast on lagged realization: Regress forecast $E_{i,t}(y_{t+h})$ on constant and lagged realized HPA y_{t-j} for $j = 0.25, 2, 4$; collect slope (3 moments per horizon).
   (b) Forecast dispersion for each $h = 1, 2, 3, 4$
      i. We compute the cross-sectional variance of the individual forecasts in each quarter, and average over quarters:
$$Disp_h = \frac{1}{T}\sum_{t=1}^{T} Var_t(E_{i,t}(y_{t+h}))$$

4. Minimize the distance between the model-implied and the empirical moments through searching over the parameters δ:

$$\hat{\delta} = \underset{\delta}{\operatorname{argmin}}\, Q(\delta) = \underset{\delta}{\operatorname{argmin}}\, \boldsymbol{g}(\delta)'\hat{W}\boldsymbol{g}(\delta), \tag{12}$$

where the moment function $\boldsymbol{g}(\delta)$ is defined as a vector of the differences between the empirical moments $\boldsymbol{m}$ and the model-implied moments $\hat{\boldsymbol{m}}(\delta)$:

$$\boldsymbol{g}(\delta) = \boldsymbol{m} - \hat{\boldsymbol{m}}(\delta). \tag{13}$$

$\hat{W} = \mathrm{diag}(\hat{\Omega})^{-1}$ is a diagonal weighting matrix.[26] $\hat{\Omega}$ is the covariance matrix of the empirical moments $\boldsymbol{m}$, calculated using the influence function technique from Erickson and Whited (2002).

We start our estimation in 1996.Q1, when the ZHVI time series starts. This means that the prior distribution refers to 1996.Q1 and that the forecasters in our model start their learning process in 1996.Q1. The model produces quarterly model-implied house price growth forecasts from 1996.Q1 until 2023.Q4. Since Pulsenomics collects survey data from 2010.Q1 onward, we use the model-implied expert forecasts from 2010.Q1 until 2023.Q4 to match the empirical moments.

Given individual forecasts, we can compute the model-implied consensus forecast as the average forecast among the N forecasters in the model:

$$E_t^c(y_{t+h}) = \frac{1}{N}\sum_{i=1}^{N} E_{i,t}(y_{t+h}).$$

4 Results: Benchmark Learning Model

4.1 Parameter Estimates

Table 3 shows the estimated parameters in $\hat{\delta}$ that govern the prior distributions of μ_0 and ρ_0, as well as the signal dispersion parameter $\overline{\Gamma}$. Standard errors on the parameter estimates are reported in parentheses below the point estimates.

The prior mean long-run growth rate of μ_μ =0.66% per quarter is consistent with the historical sample average prior to 1996. Indeed, the Case-Shiller house price index data, which start in 1987, show a quarterly house price growth rate of 0.69% for the period 1987—1995, prior to the start of our estimation in 1996.

The standard deviation of the prior is $\sigma_\mu = 0.185\%$, which reveals substantial prior uncertainty about

[26] We increase the weights of four CG slope moments and four forecast dispersion moments by a factor of twenty, due to their significance in understanding individual forecasters' expectations. Using the optimal weight matrix $\hat{W}^{opt} = \hat{\Omega}^{-1}$ yields similar parameter estimates.

Table 3: Parameter Estimation for Different Estimation Procedures

	Estimation Procedure		
	Benchmark	**No LR.**	**No Disp.**
Prior About Long-run Mean			
μ_μ	0.660	0.777	0.889
	(0.040)	(0.217)	(0.034)
σ_μ	0.185	0.024	0.009
	(0.034)	(2.452)	(1.667)
Heterogeneity of Prior About Long-run Mean			
λ	2.662	3.140	
	(0.439)	(2.367)	
Prior About Persistent Parameter			
μ_ρ	0.675	0.504	0.769
	(0.271)	(1.035)	(0.215)
σ_ρ	0.020	0.054	0.007
	(0.057)	(0.093)	(0.144)
Information Heteorgeneity			
$\bar{\Gamma}$	1.419	0.849	
	(0.496)	(0.280)	

Notes: This table presents the estimates for the key parameters from our optimization process. The first column is for the benchmark model. The second column (“No LR.") excludes empirical moments based on the 2-year, 3-year and 4-year ahead forecasts in the optimization. The third column (“No Disp.") is for a simpler model that shuts down the heterogeneity in prior beliefs and signals and estimates the model by matching moments based on consensus forecasts: $\widehat{\delta}_{cs} = \text{argmin}_\delta\, Q_{cs}(\delta) = \text{argmin}_\delta\, \boldsymbol{g}_{cs}(\delta)'\hat{W}_{cs}\boldsymbol{g}_{cs}(\delta)$, where $\hat{W}_{cs} = \text{diag}(\hat{\Omega}_{cs})^{-1}$ with the weights of four CG slope moments increased by a factor of twenty. $\boldsymbol{g}_{cs}$ and $\hat{\Omega}_{cs}$ are the consensus moment function and the covariance matrix of consensus moments, respectively. Appendix D provides more details on the optimization procedure.

the long-run mean house price growth rate. This is the key parameter that activates the “learning about the long-run mean" mechanism. Uncertainty about long-run mean HPA is the key feature that helps the model account for various anomalies, notably the negative individual CG slopes. Long horizon forecasts are especially informative for identifying σ_μ. This is evidenced by the large difference between the σ_μ estimates of “Benchmark" and those of“No-LR", which does not include long-horizon forecast moments. We return to this discussion in Section 7.

The parameter $\lambda = 2.662\%$ captures heterogeneity in the prior mean μ_μ across forecasters. The data insist on a large amount of cross-sectional heterogeneity in beliefs about long-run HPA. As we shall see, learning about the long-run mean is very difficult (slow), so that much of the prior uncertainty persists over time and in the cross-section. This is the key parameter to help the model account for the dispersion in forecasts.

The parameter ρ_0 governs the prior over the persistence of the latent state of the housing market, x. We estimate a tight prior distribution centered around a mean of $\mu_\rho = 0.675$ (quarterly value) with a standard deviation of $\sigma_\rho = 0.020$. The small degree of uncertainty about ρ_0 (tight prior) limits the role of learning about $\bar{\rho}$.

Finally, we estimate a substantial amount of signal dispersion across forecasters. Recall that forecasters receive an i.i.d. draw of a signal about current HPA, $\alpha_{i,T}$, which is normally distributed with mean zero and signal precision $(\bar{\sigma}\bar{\Gamma})^{-2}$. Our estimate for $\bar{\Gamma}$ of 1.419% indicates that a forecaster with one-standard deviation higher signal predicts a 1.419% point higher quarterly HPA.

To keep the model estimation manageable, we calibrate the prior distribution of $\sigma_0^2 \sim IG(2.82, 3.73)$. This calibration is equivalent to setting the mean and volatility of the prior for house price growth volatility in 1996.Q1 equal to the observed mean and volatility of annual house price growth volatility from 1975 to 1995 from the Freddie Mac house price index data. We also calibrate γ at 0.002, which implies that 4% of total price growth volatility, on average, is noise.[27]

Figure A4 plots the prior distributions of μ_0, ρ_0, and σ_0 implied by these parameter estimates. Figure A5 shows the time series of the posterior beliefs after learning has taken place.

4.2 Model Fit

Table 4 column (1) shows the 36 empirically-observed moments that the model aims to fit. Column (2) reports the moments implied by the Benchmark model. Note that we only have 6 parameters to match these 36 moments, so that a close fit is far from guaranteed. The rational learning model goes a long way in explaining the key empirical moments we documented. This includes many moments commonly interpreted as providing *prima facie* evidence of deviations from rationality.

Importantly, Panel A shows that the benchmark model generates a negative slope of the individual CG regressions at horizons of 2, 3, and 4 years that is large in absolute value. The observed slopes lie within the 95% confidence interval of the model. The negative slope is usually interpreted as evidence that forecasters "overreact" to news that leads them to update their forecast. We discuss an alternative interpretation of this result in the context of our learning model below. We note that the CG slope at the one year horizon is too small in absolute value. We will show below that the one-year CG slope is the first moment that can be matched much better once behavioral frictions are introduced.

Panel B shows that forecast sensitivity to lagged realized HPA in the model is positive, declining in the forecast horizon, and generally declining in the look-back window over which the regressor is computed. All of these are also features of the data. The long-run forecasts become essentially insensitive to past

[27]This can be interpreted as a dogmatic prior where the parameters of the beta distribution C and D are chosen to deliver a constant γ. We have explored a model where we allow the prior γ_0 to follow a Beta distribution and where we use a Metropolis-Hastings procedure to draw its posterior distribution. Such a specification leads to similar results but requires much longer computation time.

house price growth in both model and data. Hence the model produces under-extrapolation, the fact that forecasters impute “too much" mean-reversion into HPA.

Table 4: Comparison of Individual-Level Model Fit

Moments	Horizon	Data	Benchmark	Overconfidence ($\phi = 2.5$)	Diagnostic Exp. ($\theta = 0.8$)	No LR.	No Disp.
Panel A: CG Regression Moments							
		(1)	(2)	(3)	(4)	(5)	(6)
CG *a*	**1-Yr**	3.151	2.553	2.552	2.553	1.858	2.015
		(0.068)	(0.059)	(0.060)	(0.060)	(0.064)	(0.058)
	2-Yr	4.099	3.879	3.863	3.864	3.268	3.364
		(0.063)	(0.057)	(0.058)	(0.057)	(0.069)	(0.052)
	3-Yr	4.304	4.274	4.277	4.273	3.874	3.665
		(0.062)	(0.059)	(0.059)	(0.059)	(0.072)	(0.052)
	4-Yr	4.010	4.256	4.262	4.252	3.892	3.590
		(0.075)	(0.073)	(0.074)	(0.073)	(0.090)	(0.064)
CG *b*	**1-Yr**	-0.323	0.029	-0.399	-0.302	-0.150	0.105
		(0.035)	(0.038)	(0.021)	(0.026)	(0.027)	(0.045)
	2-Yr	-0.570	-0.447	-0.460	-0.463	-0.193	-2.291
		(0.044)	(0.088)	(0.061)	(0.072)	(0.066)	(0.180)
	3-Yr	-0.576	-0.416	-0.527	-0.509	-0.248	-2.876
		(0.049)	(0.107)	(0.098)	(0.104)	(0.144)	(0.526)
	4-Yr	-0.411	-0.273	-0.284	-0.289	1.581	0.583
		(0.061)	(0.134)	(0.129)	(0.133)	(0.340)	(1.101)
Weighted Cost Function			**1165.5**	**282.1**	**249.8**	**13837.6**	**52197.7**

Moments	Horizon	Data	Benchmark	Overconfidence ($\phi = 2.5$)	Diagnostic Exp. ($\theta = 0.8$)	No LR.	No Disp.
Panel B: Sensitivity to Past House Price Growth							
Sensitivity to Past 1Q Growth	**1-Yr**	0.192	0.247	0.253	0.251	0.371	0.300
		(0.006)	(0.004)	(0.006)	(0.005)	(0.006)	(0.001)
	2-Yr	0.051	0.091	0.097	0.095	0.186	0.108
		(0.005)	(0.005)	(0.005)	(0.005)	(0.008)	(0.000)
	3-Yr	0.012	0.041	0.044	0.044	0.077	0.039
		(0.004)	(0.007)	(0.007)	(0.007)	(0.011)	(0.000)
	4-Yr	-0.003	0.025	0.026	0.026	0.034	0.014
		(0.004)	(0.007)	(0.007)	(0.007)	(0.013)	(0.000)
Sensitivity to Past 2Y Growth	**1-Yr**	0.108	0.198	0.193	0.195	0.279	0.240
		(0.006)	(0.005)	(0.007)	(0.006)	(0.007)	(0.003)
	2-Yr	0.029	0.088	0.093	0.091	0.160	0.095
		(0.005)	(0.005)	(0.006)	(0.005)	(0.008)	(0.001)
	3-Yr	0.015	0.039	0.041	0.040	0.058	0.027
		(0.004)	(0.006)	(0.006)	(0.006)	(0.010)	(0.000)
	4-Yr	0.008	0.027	0.028	0.027	0.026	0.010
		(0.004)	(0.006)	(0.006)	(0.006)	(0.011)	(0.000)
Sensitivity to Past 4Y Growth	**1-Yr**	0.102	0.218	0.221	0.221	0.312	0.250
		(0.007)	(0.005)	(0.008)	(0.007)	(0.008)	(0.003)
	2-Yr	0.022	0.092	0.098	0.096	0.161	0.089
		(0.005)	(0.005)	(0.006)	(0.006)	(0.009)	(0.001)
	3-Yr	0.010	0.046	0.048	0.047	0.060	0.025
		(0.004)	(0.006)	(0.006)	(0.006)	(0.010)	(0.000)
	4-Yr	0.007	0.034	0.034	0.034	0.028	0.009
		(0.004)	(0.007)	(0.007)	(0.006)	(0.011)	(0.000)
Weighted Cost Function			**896.7**	**994.7**	**964.9**	**3798.4**	**1177.4**

Table 4: Comparison of Individual-Level Model Fit (cont.)

Panel C: Biases and Inverse MZ Regression Moments							
Moments	**Horizon**	**Data**	**Benchmark**	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Forecast Biases	**1-Yr**	2.952	2.371	2.309	2.326	1.669	1.846
		(0.065)	(0.061)	(0.064)	(0.062)	(0.065)	(0.059)
	2-Yr	4.025	3.781	3.759	3.762	3.186	3.190
		(0.059)	(0.057)	(0.057)	(0.057)	(0.068)	(0.052)
	3-Yr	4.299	4.288	4.287	4.283	3.887	3.665
		(0.058)	(0.058)	(0.059)	(0.058)	(0.071)	(0.051)
	4-Yr	4.021	4.259	4.264	4.257	3.910	3.593
		(0.062)	(0.064)	(0.064)	(0.063)	(0.078)	(0.056)
Inverse MZ *a*	**1-Yr**	2.226	2.627	2.367	2.475	2.740	3.261
		(0.065)	(0.052)	(0.068)	(0.061)	(0.076)	(0.041)
	2-Yr	3.403	3.127	3.115	3.146	3.992	3.927
		(0.068)	(0.063)	(0.066)	(0.064)	(0.101)	(0.020)
	3-Yr	3.424	2.855	2.880	2.853	3.346	3.571
		(0.063)	(0.070)	(0.072)	(0.070)	(0.118)	(0.006)
	4-Yr	3.600	2.832	2.823	2.798	3.164	3.568
		(0.059)	(0.072)	(0.074)	(0.072)	(0.126)	(0.003)
Inverse MZ *b*	**1-Yr**	0.130	0.170	0.224	0.203	0.268	0.152
		(0.009)	(0.007)	(0.009)	(0.008)	(0.010)	(0.005)
	2-Yr	-0.064	0.004	0.008	0.004	-0.035	-0.026
		(0.009)	(0.008)	(0.009)	(0.008)	(0.013)	(0.003)
	3-Yr	-0.051	0.018	0.015	0.019	0.006	0.006
		(0.008)	(0.009)	(0.009)	(0.009)	(0.015)	(0.001)
	4-Yr	-0.057	0.014	0.015	0.019	0.017	0.005
		(0.007)	(0.009)	(0.009)	(0.009)	(0.016)	(0.000)
Weighted Cost Function			**623.7**	**694.6**	**686.4**	**1152.8**	**1039.8**

Panel D: Forecaster Dispersion Moments							
Moments	**Horizon**	**Data**	**Benchmark**	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Forecaster Dispersion	**1-Yr**	6.476	3.033	6.325	4.732	5.818	−−
	2-Yr	4.893	3.435	3.773	3.517	8.339	−−
	3-Yr	3.490	3.805	3.954	3.793	10.729	−−
	4-Yr	2.946	4.034	4.171	4.009	12.258	−−
Weighted Cost Function			**5865.1**	**2358.2**	**3047.8**	**150401.0**	−−

Panel E: Total Weighted Cost Function					
	Benchmark	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Total Weighted Cost Function	8551.1	4329.7	4948.9	169190.1	−−
Coefficient (weighted)	2686.0	1971.4	1901.0	18788.8	54414.8
Dispersion (weighted)	5865.1	2358.2	3047.8	150401.2	−−

Notes: This table shows the comparison of individual-level fit across different models. Column 1 shows the results calculated from the actual data. Columns 2 to 6 show the results calculated from $N = 100$ "forecasters" simulated from different models. The coefficients shown in Panels A to C are estimated using individual forecast data (actual or simulated) in panel regressions based on Equation (1) to Equation (4). Standard errors (not clustered) are reported in parentheses. Panel D shows the dispersion moments. Panel E shows the values of the total weighted cost function for different models. All forecast data are in percentages.

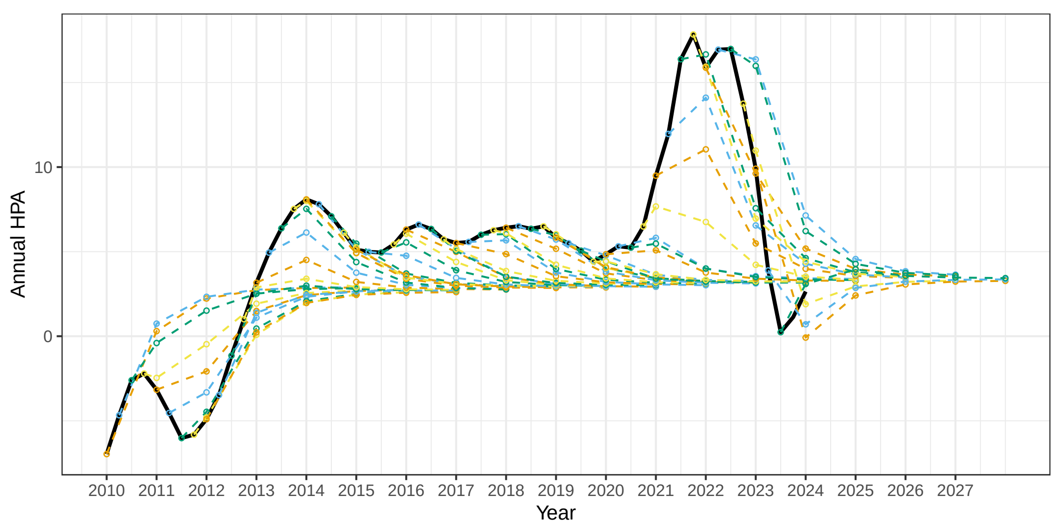
Panel C shows that our benchmark model produces large average in-sample forecast biases that are larger at the 2–4-year horizons, consistent with the data. It also produces forecasts whose sensitivity to contemporaneous realized HPA (inverse MZ b) is close to zero and far way from one, consistent with the data.

Panel D shows that the model does a reasonably good job matching the level of forecast dispersion, especially at longer horizons of 2–4 years. However, the benchmark model understates the forecast dispersion at short horizons. The one-year forecast dispersion is the second moment that can be matched much better with behavioral frictions.

Panel E shows that the distance between model and data, as captured by the cost function, attains a value of 8,551, a reference point for other model versions discussed below.

Figure 2 plots the estimated model-implied consensus forecast for HPA in the time series. It looks similar to Figure 1 which was based on empirical forecasts. The individual forecasts shown in Appendix Figure A6 are naturally more volatile than the consensus, which illustrates that the model is capable of generating substantial dispersion in forecasts.

Figure 2: Model-Implied Consensus Forecasts



Notes: This graph shows the consensus forecasts implied by the Benchmark rational learning model. The black solid line represents the realized annual HPA. Each short, colored dashed line with six small circles indicates the actual annual HPA (the first circle), the consensus forecast for the current calendar year (the second circle), and the consensus forecasts for the following four calendar years (the subsequent four circles).
Alt text: A graph illustrating the model-implied consensus forecasts for house price growth (colored dashed lines) compared to realized annual house price growth (solid black line), with each colored line segment representing forecasts for the current year and the next four years.

Table A9 reports 24 residual standard error moments that are not included in the estimation. They measure the volatility of the anomaly regression residuals, resulting both from time-series and cross-sectional variation. Since these moments are untargeted, they provide additional opportunities for model validation. The model fits these residual standard error moments quite well, confirming its ability to produce not only realistic average consensus forecasts but also the variation in these forecasts over time and across forecasters.

4.3 Individual versus Consensus CG Regression Slopes

The benchmark model also fits well the same moments obtained from consensus forecasts instead of individual forecasts (in both model and data). Outside of the 4 forecast dispersion moments, 24 of the remaining 32 moments included in the estimation are identical when estimated on consensus moments. Table 5 shows the fit for the remaining 8 moments which differ between individual and consensus: the CG regression intercept and slope moments. The moments in this table were not used in estimation, yet the fit is good. The CG regression slopes are, if anything, positive when estimated on consensus data. We note the large standard errors on the consensus CG slope estimates in both data and model.

The sign switch between b^{cg} at the individual level (negative) and at the consensus level (positive) is found for many other macro-economic and financial survey data and is the focus of Bordalo et al. (2020). They argue that the overreaction to news at the individual level and the underreaction to news at the consensus level can be explained in a model where forecasters receive heterogeneous signals and have diagnostic expectations. The positive sign at the consensus level comes from aggregating forecast revisions among forecasters with heterogeneous signals, which shrinks the average revisions towards zero. The lack of reaction to other forecasters' signals creates rigidity in the consensus belief and a positive consensus coefficient. They argue that the negative sign at the individual level arises from diagnostic expectations, which leads forecasters to react more strongly to news.

The underreaction at the consensus level arises naturally in Bayesian settings with signal dispersion, including ours. Our paper shows that learning can also generate the negative b^{cg} at the individual level. In our model, the negative slope in the individual CG regressions is not an actual overreaction to news but arises from learning about the long-run mean of HPA, a channel that is absent in Bordalo et al. (2020) where the long-run mean is known (and set to zero). We provide intuition for this result in the next subsection. We incorporate diagnostic expectations in our learning setting in Section 6.

Table 5: Comparison of Consensus-Level Model Fit of CG Moments

Moments	Horizon	Data	Benchmark	Overconfidence ($\phi = 2.5$)	Diagnostic Exp. ($\theta = 0.8$)	No LR.	No Disp.
		(1)	(2)	(3)	(4)	(5)	(6)
CG *a*	**1-Yr**	3.191	2.442	2.386	2.387	1.808	2.015
		(0.579)	(0.542)	(0.515)	(0.523)	(0.550)	(0.587)
	2-Yr	4.030	3.969	3.851	3.894	3.313	3.367
		(0.546)	(0.537)	(0.537)	(0.539)	(0.569)	(0.532)
	3-Yr	4.100	4.343	4.435	4.422	4.120	3.673
		(0.574)	(0.536)	(0.534)	(0.533)	(0.525)	(0.531)
	4-Yr	4.330	4.128	4.144	4.090	3.911	3.587
		(0.768)	(0.668)	(0.663)	(0.666)	(0.662)	(0.664)
CG *b*	**1-Yr**	0.668	0.952	0.772	0.933	0.303	0.106
		(0.629)	(0.580)	(0.457)	(0.509)	(0.379)	(0.462)
	2-Yr	-1.629	-1.959	-0.290	-0.919	-0.599	-2.352
		(2.045)	(2.657)	(2.167)	(2.311)	(1.159)	(1.872)
	3-Yr	-2.593	-4.093	-7.227	-6.766	-4.901	-3.801
		(3.198)	(6.448)	(5.696)	(5.740)	(2.449)	(6.342)
	4-Yr	4.514	7.834	7.734	11.347	-0.747	3.608
		(4.976)	(9.625)	(9.161)	(11.963)	(10.472)	(22.202)
Weighted Cost Function			**22.8**	**95.4**	**110.4**	**62.5**	**23.8**

Notes: This table shows the comparison of the consensus-level fit of CG moments across different models. Consensus forecasts are calculated as the average forecasts across all forecasters. Column 1 shows the results calculated from the actual data. Columns 2 to 6 show the results calculated from $N = 100$ "forecasters" simulated from different models. Coefficients are estimated using consensus forecasts based on time series regressions. Standard errors are reported in parentheses. The "Weighted Cost Function" for the coefficients is given by $Q_{cs}(\delta) = gcs(\delta)'\hat{W}csg_{cs}(\delta)$, as defined in the notes of Table 3. All forecast data are in percentages.

Figure 3 shows scatter plots of forecast revisions (x-axis) against forecast errors (y-axis) in the actual data (left panel) and in the benchmark model (right panel). The colored dots represent individual forecasters (organized by period) and the black dots represent the consensus (one dot per period). The dashed line shows the negative relationship based on the individual data points, while the solid black line shows the positive (and much more imprecisely estimated) relationship based on the consensus forecast. Model and data accord well.

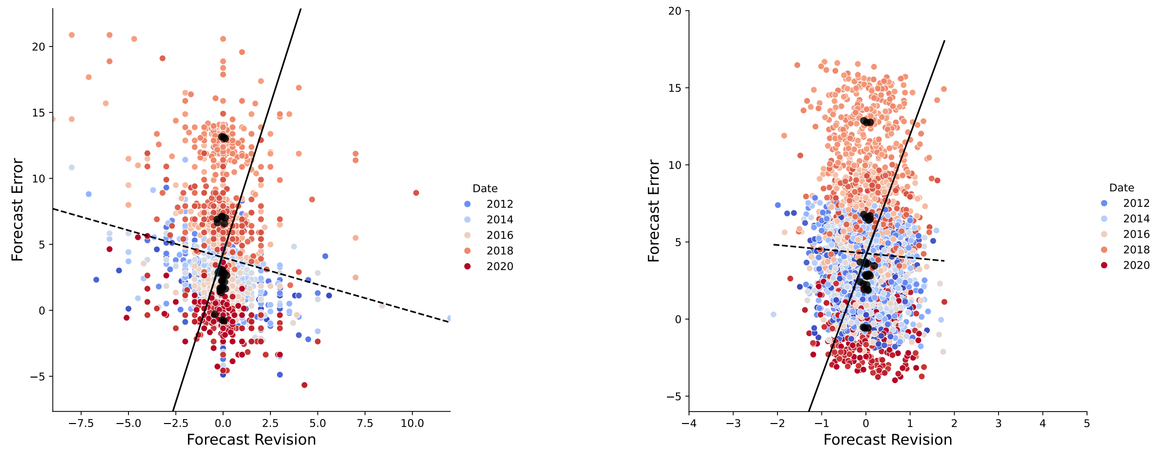
4.4 Understanding the Main Economic Forces

Two key economic forces drive our model's forecasts at the individual level and across forecasters. We discuss these economic forces and further demonstrate their workings through simulation.

4.4.1 Learning About the Long-Run Mean

Individual forecasters in our model do not know the true value of long-run house price growth ($\overline{\mu}$), and must learn about it by combining prior information and signals using Bayes' Law. This learning about long-run mean growth generates patterns that appear anomalous from the perspective of a rational expectations

Figure 3: Individual versus Consensus CG Regression Slopes of 4-Year Ahead Forecasts



(a) Actual (b) Benchmark Model

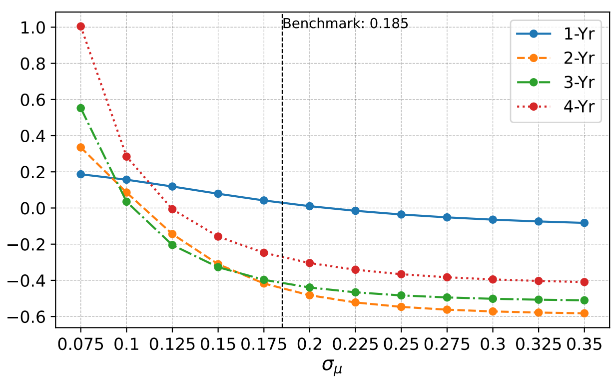
Notes: Black dots show the consensus forecasts at different periods, dashed lines show the individual CG regressions, and solid lines show the consensus CG regressions. The left panel shows the actual data, and the right panel shows the model-simulated data from the benchmark model.
Alt text: Scatter plot comparing individual and consensus forecast revisions against forecast errors for the actual (left panel) and the model-simulated data (right panel). Dashed lines illustrate negative relationships for individual forecasts, while solid lines indicate positive relationships for consensus forecasts.

model. The magnitude of σ_μ, which governs the precision of forecasters' prior beliefs, determines the volatility of their forecasts. This volatility in turn prominently affects the CG coefficients and, to a lesser extent, the sensitivity of forecasts to past realized growth.

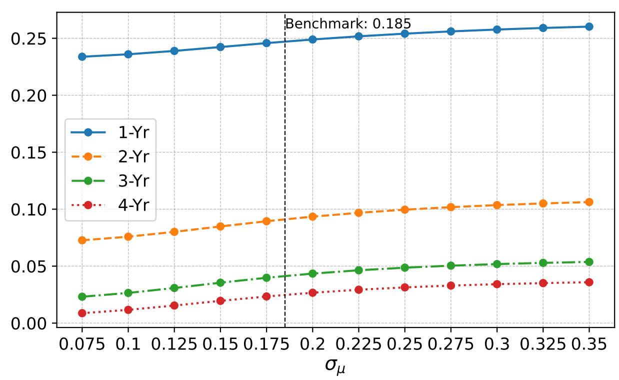
If forecasters have more diffuse priors (larger σ_μ), they assign more weight to their signals when making forecasts. Such "modest" forecasters' HPA predictions appear excessively volatile or to be overreacting to news about HPA. The econometrician estimates more negative individual CG coefficients and a stronger response of the forecast to past growth. Conversely, more "dogmatic" forecasters with tighter prior beliefs (smaller σ_μ) assign less weight to signals when making forecasts. Consequently, their forecast revisions do not predict future forecast errors, and their response to past realized growth appears too sluggish compared to the observed auto-correlation in realized house price growth. In the extreme case, where forecasters think they know the exact underlying value of the long-run house price growth ($\sigma_\mu = 0$), they do not revise their long-run forecasts at all, and the sensitivity of long-run forecasts to past HPA is zero. If their beliefs are consistent with the true parameter values, we are back in the FIRE case.

Figure 4 demonstrates this mechanism based on model simulations for different values of σ_μ, reported on the x-axis. On the y-axes, it plots the values of the individual CG coefficients (Panel a), the sensitivity of forecast to past 1Q growth (Panel b) and to past 4Y growth (Panel c). As the value of σ_μ becomes smaller than the benchmark value, the CG coefficients gradually increase, and become positive around $\sigma_\mu = 0.10$, mimicking the forecasts of a "dogmatic" forecaster. Conversely, the individual CG regression slope becomes more negative as σ_μ becomes larger, i.e., as the forecaster becomes more "modest". The effects on forecasters' sensitivity to past realized HPA are the opposite. As σ_μ increases, forecast sensitivity to past realized growth increases. Notably, the effect of σ_μ is the largest on the long-run forecasts and long-run CG slopes, confirming our "learning about the *long-run* mean" mechanism.

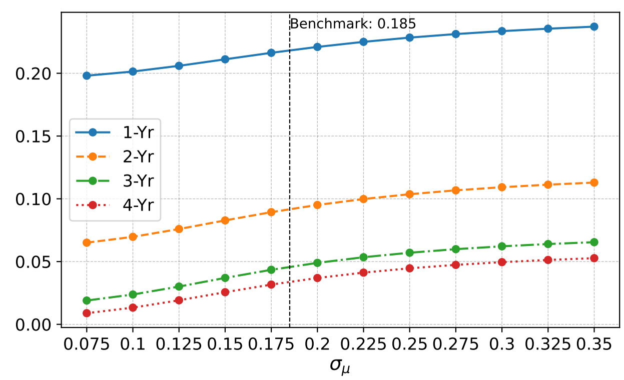
Figure 4: Impact of Forecaster Conviction (σ_μ) on Individual Forecasts



(a) Impact of σ_μ on CG-b coefficients



(b) Impact of σ_μ on sensitivity to 1Q past growth



(c) Impact of σ_μ on sensitivity to 4Y past growth

Notes: The graph varies the parameter σ_μ, holding all other parameters fixed at the benchmark model's values.
Alt text: Line graph showing how forecasters' conviction about the long-run mean house price growth (σ_μ) affects key individual forecasting metrics, with the benchmark model's estimated value of conviction parameter indicated by dotted black lines.

Figure 4 reveals how the two sets of empirical moments, individual CG coefficients and the sensitivity to past growth, help identify the parameter σ_μ in our model. Forecasters in the data exhibit negative CG coefficients, and a substantial positive short-run but weak long-run response to past realized growth. These moments point towards an intermediate value for σ_μ. A higher value than the estimated 0.18 would help to bring the model closer to the data in terms of the individual CG slopes at horizons 2-4, and could even turn the 1-year CG slope negative, but would hurt the model's fit for the sensitivity of 2-4-year-ahead forecasts to lagged HPA. The resulting intermediate value of 0.18 implies a high degree of prior uncertainty about long-run HPA. This is a world in which learning about long-run growth matters (far from the dogmatic case), and new data can sway forecasts.

To demonstrate cleanly that learning about the long-run mean is the key mechanism responsible for creating overreaction at the individual level, we simulate a model where we shut down this learning. We do so by setting σ_μ at increasingly small values, keeping the prior mean unbiased: $\mu_\mu = \overline{\mu}$. All other parameters are kept at their benchmark values. The simulation results in Figure A7 confirm that the models with little or no learning about the long-run mean (σ_μ below 0.1) generate positive rather than negative individual CG slopes at all forecast horizons. The model without uncertainty and learning about the long-run mean cannot generate overreaction at the individual level.

Our learning environment is similar to Collin-Dufresne, Johannes and Lochstoer (2016), where learning about the long-run mean in a multi-dimensional setting leads to "excessive" volatility in the subjective expectations of forecasters. Their model does not consider the implications for the CG regression nor the implications for the term structure of sensitivities to past realized house price growth. But like in our model, forecasters in their model face a difficult multi-dimensional learning problem.

In contrast, in Nagel and Xu (2022) there is (fading-memory) learning about the long-run mean, but inference is simpler because there is no unknown mean-reverting component (the $\rho_{i,t}\tilde{x}_t^i$ component of our forecast). To compare this mechanism to ours, we solve a model where HPA forecasts are made using fading-memory learning. The resulting model fit is substantially worse than in for the benchmark model (Table A10).

It is because of, not despite, slow learning about the long-run mean, that we are able to generate overreaction.[28]

4.4.2 Uncertain vs. Biased Initial Beliefs

The economic mechanism behind our model is different from the one proposed in Farmer, Nakamura and Steinsson (2023), where forecasters have biased prior beliefs. Farmer et al. (2023) show that, when prior beliefs about the persistence parameter $\overline{\rho}$ are downward-biased, their model generates underreaction of consensus forecasts for interest rates. Conversely, upward-biased beliefs about persistence generate over-

[28] Appendix I uses simulations to demonstrate that it takes about 100 years for beliefs about the long-run mean to converge to their true value if they start off either too high or too low. Forecasters' uncertainty about the long-run mean also leads them to mis-attribute part of the transitory shocks to the state of the housing market to long-run mean growth, similar to the mechanism suggested by Afrouzi et al. (2023). The same mechanism results in slow learning about $\overline{\rho}$. Our mechanism highlights the role of uncertainty in shaping experts' beliefs, consistent with evidence found among households (Coibion et al., 2024) and firms (Kumar, Gorodnichenko and Coibion, 2023).

reaction of consensus forecasts. Since our model nests theirs and extends it for forecaster heterogeneity, we show here that upward-biased beliefs about the persistence parameter can generate overreaction of individual beliefs. However, those beliefs also give rise to over-extrapolation of model forecasts, rather than the under-extrapolation we see in the HPA forecast data.

To highlight the role of biased beliefs about persistence, we vary the prior mean μ_ρ from 0.35 to 0.95, holding all other parameters at their benchmark estimates. Based on different priors, we let the forecasters learn from the HPA data and form their forecasts for future HPA. Panel (a) of Figure 5 shows the individual CG slope while panels (b) and (c) show the sensitivity of the forecast to past realized HPA. Higher values of μ_ρ result in more negative individual CG slopes.[29] The effect on the individual CG slope is strongest at the 1-year horizon. In fact, when μ_ρ is around 0.85, the CG slopes are closer to the data than at the benchmark estimate of $\mu_\rho = 0.675$. However, the overall model fit deteriorates significantly; see panel (d).

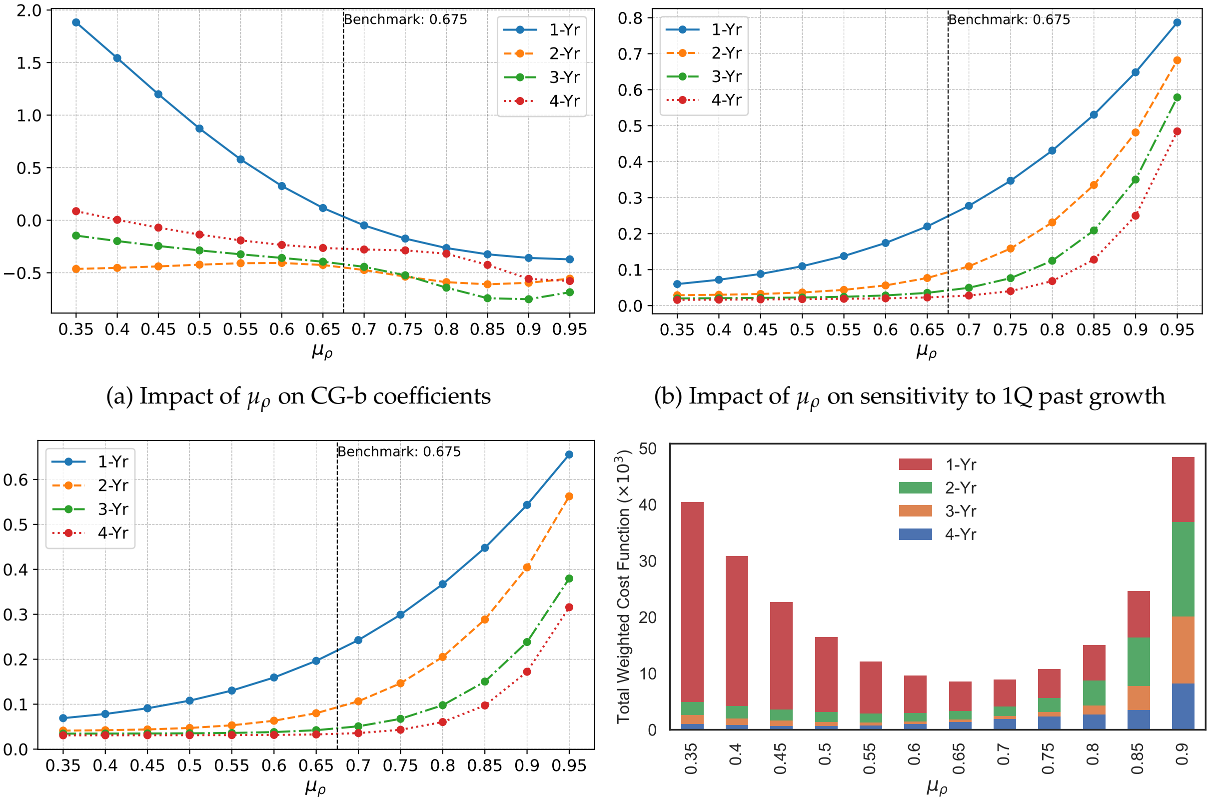
The main reason for the worse fit arises from the model's prediction for the forecast sensitivity to past realized HPA. Panels (b) and (c) show that these sensitivities increase strongly in the value of μ_ρ. As μ_ρ increases, forecasts become much more sensitive to past HPA, not only relative to the benchmark model but also relative to the data. If forecasters' prior beliefs about μ_ρ are high, their posterior beliefs about $\overline{\rho}$ are also higher, and forecasters believe house price growth to be persistent. That makes their forecasts overly sensitive to past realized HPA.

To demonstrate more cleanly that an upwardly-biased prior belief about persistence ($\mu_\rho > \overline{\rho}$) generates over-extrapolative forecasts, we simulate a HPA process from (5) and (6), setting the true persistence $\overline{\rho} = 0.675$. The correlations between current realized price growth and the past 1-quarter, 2-year and 4-year price growth, together with their two-standard error bands, are shown as the blue bars in Figure 6. Next, we generate forecasts based on prior μ_ρ equal to 0.92 and 0.42, representing upward- and downward-biased priors, respectively. We keep $\sigma_\rho = 0.020$ as estimated in the benchmark model, so that the prior is tight. We endow forecasters with an unbiased prior about the long-run mean $\mu_\mu = \overline{\mu} = 0.66$, and set $\sigma_\mu = 0$ in order to shut down the "uncertainty about the long-run mean" channel. The resulting forecast sensitivities to realized price growth at different horizons are plotted as the red bars (upward-biased, $\mu_\rho = 0.92$) and green bars (downward-biased, $\mu_\rho = 0.42$) in Figure 6, respectively. The upward-biased prior about persistence generates forecasts that are overly sensitive to past house price growth, i.e., over-extrapolation.[30] Conversely, if forecasters have a downward-biased prior about the persistence parameter, their forecasts under-extrapolate (and they also underreact to news, leading to positive individual CG slopes).

[29] High values for the prior mean of μ_ρ, such as values above 0.9, also result in high values for the posterior mean, and hence an upwardly-biased estimate for $\overline{\rho}$ if we assume that the true coefficient $\overline{\rho}$ takes on an intermediate value around 0.7. The simulation exercise below will confirm these results for a known $\overline{\rho}$.

[30] While it is clear why the estimation does not choose a higher estimate of μ_ρ, it leaves open the deeper question of why professional forecasters impute more mean reversion in HPA than could be estimated from OLS regressions.

Figure 5: Impact of Prior Belief about the Persistence Parameter (μ_ρ) on Individual Forecasts



(a) Impact of μ_ρ on CG-b coefficients

(b) Impact of μ_ρ on sensitivity to 1Q past growth

(c) Impact of μ_ρ on sensitivity to 4Y past growth

(d) Impact of μ_ρ on total weighted cost function

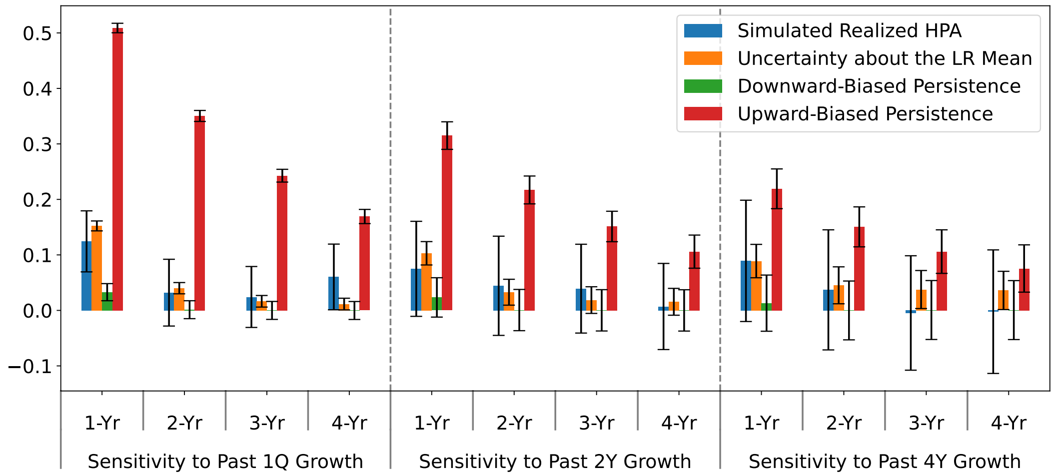
Notes: The graph varies the parameter μ_ρ, holding all other parameters fixed at the benchmark model's values.
Alt text: Graph showing how varying the prior belief about the persistence of house price growth (μ_ρ) affects key individual forecasting metrics, with the benchmark model's estimated value of persistence parameter indicated by dotted black lines.

For comparison, we also generate forecasts based on uncertainty about the long-run mean channel, in which forecasters have unbiased but uncertain priors about the long-run mean with $\mu_\mu = \overline{\mu} = 0.66$ and $\sigma_\mu = 0.185$, but know the true persistence parameter without uncertainty ($\sigma_\rho = 0$, $\mu_\rho = \overline{\rho} = 0.675$). The orange bars in Figure 6 show that the uncertainty about the long-run mean channel generates sensitivity of forecasts to lagged HPA that is in line with the simulated data.

4.4.3 Prior and Signal Heterogeneity and Forecast Dispersion

The second set of key targets for our model are the cross-sectional forecast dispersion moments. Both heterogeneity in forecasters' prior beliefs (governed by λ) and heterogeneity in their signals (governed by Γ) generate forecast dispersion. Our model reveals how these two different forces affect both the level and the term structure of the forecast dispersion, adding new insight to the previous literature (e.g., Patton and

Figure 6: Uncertain Priors vs. Biased Priors — Differential Impacts on Sensitivity to Past House Price Growth



Notes: This graph compares the impacts of uncertainty about the long-run mean ($\overline{\mu}$) and biased priors about persistence ($\overline{\rho}$) on the correlation of forecasts with past realized HPA. The HPA process is simulated according to equations (5) and (6), with parameters $\overline{\mu} = 0.660$, $\overline{\rho} = 0.675$, $\overline{\sigma} = 0.822$, and $\overline{\gamma} = 0.002$. The blue bars show the correlations between the (simulated) realized HPA and the (annualized) past 1-quarter, 2-year, and 4-year HPA. The orange bars show the correlations of forecasts based on uncertainty about the long-run mean mechanism, without incorporating learning about persistence ($\mu_\mu = \overline{\mu} = 0.660$, $\sigma_\mu = 0.185$, $\mu_\rho = \overline{\rho} = 0.675$, and $\sigma_\rho = 0$). The green and red bars show the correlations of forecasts with downward-biased ($\mu_\rho = 0.42$) and upward-biased ($\mu_\rho = 0.92$) priors about persistence, respectively, without learning about the long-run mean. In cases with biased priors, the other parameters are set to $\sigma_\rho = 0.020$, $\mu_\mu = \overline{\mu} = 0.660$, and $\sigma_\mu = 0$. Black whiskers represent one standard error above and below the mean estimates from panel regressions.
Alt text: Bar chart comparing the effects of uncertain priors about the long-run mean house price growth (HPA) and biased priors about persistence of HPA on forecast sensitivity to past realized HPA. The x-axis represents different horizons of forecasts (upper part) and the realized HPA (lower part), while the y-axis shows their respective correlations; black whiskers indicate one standard error above and below the mean estimates.

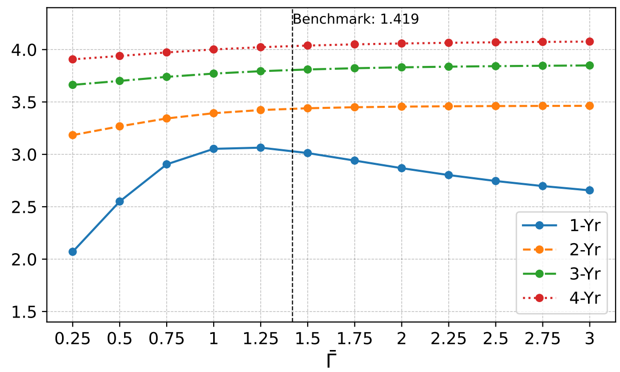
Timmermann, 2010).

First, the impact of signal heterogeneity ($\overline{\Gamma}$) on the level of forecast dispersion is limited in a rational learning model. When forecasters are rational and aware that their signals contain a lot of noise (high $\Gamma = \overline{\Gamma}$), they under-weight the signal when forming expectations following Bayes' Rule. An increase in $\overline{\Gamma}$ initially increases forecast dispersion due to learning from data. However, its impact diminishes as $\overline{\Gamma}$ increases further as forecasters discount their own signal when they realize that $\overline{\Gamma}$ is high.

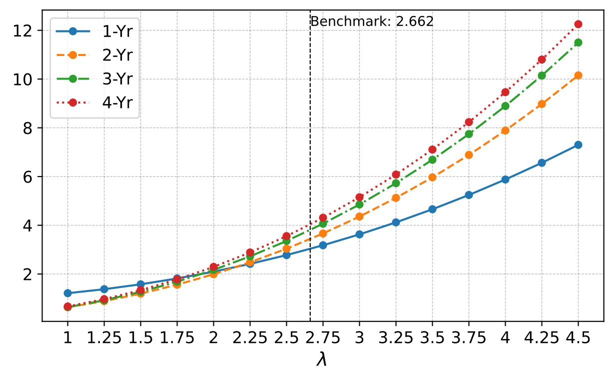
Panel (a) of Figure 7 confirms this mechanism. It plots forecast dispersion, the cross-sectional variance of the forecast, for different values of the parameter $\overline{\Gamma}$. The declining marginal impact of higher signal dispersion is visible at all horizons, but is especially strong for the 1-year forecast dispersion which displays a hump shape. At low levels of $\overline{\Gamma}$, increases in $\overline{\Gamma}$ lower the slope of the term structure of forecast dispersion, while beyond the benchmark value of $\overline{\Gamma} = 1.42$, increases in $\overline{\Gamma}$ result in a steeper term structure of forecast dispersion.

Second, the heterogeneity in forecasters' prior beliefs about the long-run mean (λ) has a first-order

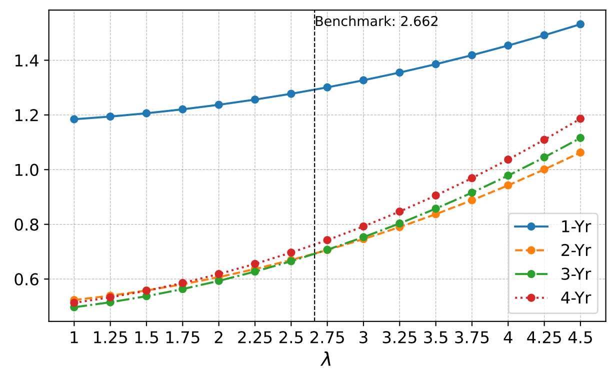
Figure 7: Impact of Signal Dispersion ($\overline{\Gamma}$) and Prior Heterogeneity (λ) on Forecast Dispersion



(a) Impact of $\overline{\Gamma}$ on forecast dispersion



(b) Impact of λ on forecast dispersion when $\sigma_\mu = 0.185$



(c) Impact of λ on forecast dispersion when $\sigma_\mu = 0.5$

Notes: Panel (a) varies the signal variance parameter $\Gamma = \overline{\Gamma}$, holding all other parameters fixed at the benchmark model's values. Panel (b) varies the parameter λ, which governs the dispersion in the prior about the long-run mean house price growth rate, holding all other parameters fixed at the benchmark model's values. Panel (c) redoes panel (b) for a value of $\sigma_\mu = 0.5$, higher than the benchmark value of 0.185.
Alt text (decorative): A set of graphs illustrating how forecasters' signal dispersion ($\overline{\Gamma}$) and prior belief heterogeneity (λ) affect forecast dispersion (y-axis) across different horizons, with the benchmark model's estimated values of parameters indicated by dotted black lines

impact on both the level and the slope of the term structure of forecaster dispersion. A higher value of λ increases the level of forecast dispersion at all horizons. It does so more for the long-run forecasts, leading to an upward-sloping term structure of forecast dispersion. Panel (b) of Figure 7 shows the quantitatively large impact of heterogeneity about the prior long-run mean, λ, on forecast dispersion at our benchmark estimate of σ_μ.

Finally, the effect of λ on forecast dispersion should be jointly considered with the level of forecaster conviction of their prior beliefs (σ_μ). For more "modest" forecasters who place a lower weight on their prior beliefs in their expectation formation process (high σ_μ), the impact of λ on the level of dispersion is limited. Modest forecasters place greater weight on the signals which erodes the dispersion in forecasts coming from prior heterogeneity. Panel (c) of Figure 7 shows that, when we increase σ_μ to 0.5, the increase in dispersion is much smaller when λ increases. We also notice a much lower level of forecast dispersion at all horizons. These results help explain the intermediate value for σ_μ found by the estimation routine.

In sum, generating enough forecast dispersion requires enough dispersion in the prior (a high value of λ), enough signal dispersion $\overline{\Gamma}$, as well as not too much uncertainty about the long-run mean (a low enough value of σ_μ). There are limits to how much forecast dispersion can be generated by cranking up signal dispersion when investors are rational (know the true signal noise). The parameter σ_μ, which modulates the strength of the learning about the long-run mean channel, is well identified. The individual CG slope moments tell the model that σ_μ cannot be too low, while the sensitivity of forecasts to lagged HPA and especially the forecast dispersion tell the model that σ_μ cannot be too high.

5 Supporting Evidence from Micro Data

The learning model makes four additional predictions for individual forecasters. Given the unique panel nature of our forecast data, we can test these predictions in the data.

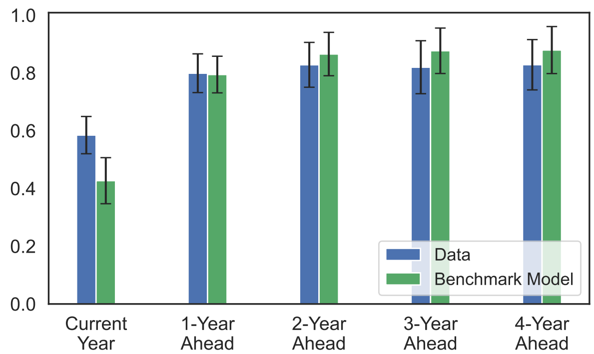
First, since learning is slow and investors have heterogeneous priors, forecast errors at the individual forecaster level are persistent over time in the model. Figure 8 shows that this is true in the Pulsenomics data as well, for both forecast errors (panel a) and absolute forecast errors (panel b). The annual persistence is around 0.8 at forecast horizons of 1–4 years. The persistence in model simulations is similar.

Second, the model predicts substantial heterogeneity in long-run forecasts, in no small part due to the high degree of heterogeneity in the prior beliefs about the long-run mean HPA governed by the parameter λ. Figure 9 shows that the data also feature forecasts that display enormous dispersion. For each forecaster and each survey quarter, we compute the implied compound quarterly growth rate of predicted house prices over the full forecast horizon. The dashed lines show the 5^{th} and 95^{th} percentiles of forecasts. The left tail of forecasters predicts negative HPA over the next four years while the right tail predicts HPA of 6% per year over the next four years.

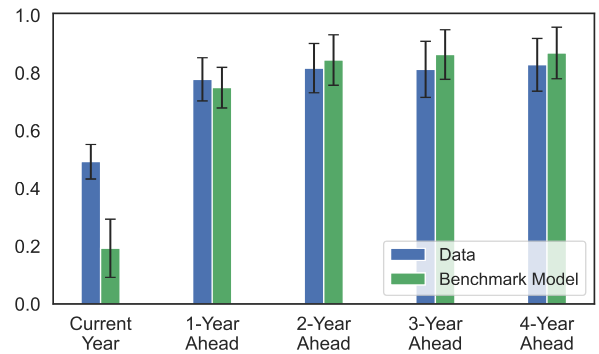
Third, we study in more depth the heterogeneity across forecasters. We match individual forecasters to their Linked-IN profiles to collect extensive information on their education and professional experience. The first two columns of Table 6 show that these demographic variables help account for some of the heterogeneity in forecasts. Compared to the specification in columns 1 that contains time fixed effects, the specification in column 2 that adds forecaster characteristics boosts the R^2 from 14% to 18.2%. Adding regional fixed effects increases it further to 20.1%.[31] We find similar improvements in fit when the depen-

[31] The long form of this table is reported as Appendix Table A5.

Figure 8: Persistence of Individual-Level Forecast Errors



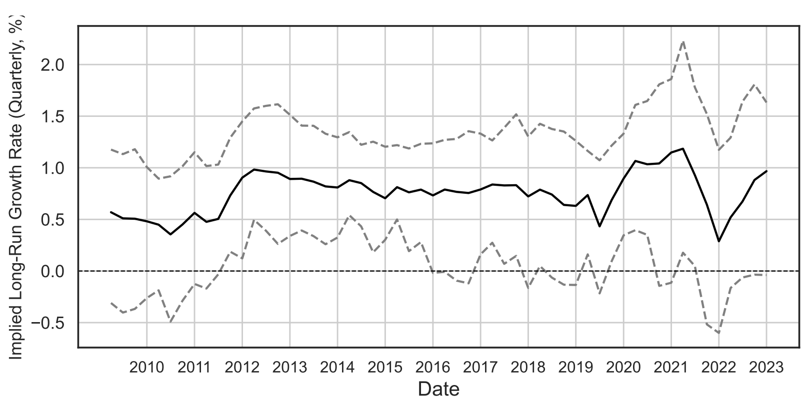
(a) Persistence of Forecast Errors



(b) Persistence of Absolute Forecast Errors

Notes: Panels (a) and (b) show the persistence of individual-level forecast errors and absolute forecast errors. Blue bars show the estimates calculated from the survey data. Green bars show the results calculated from the benchmark model. The coefficients are estimated from individual-level panel regressions: $e_{i,t+h|t} = a^p + b^p e_{i,t+h|t-1} + \varepsilon_{i,t+h}$, and $|e_{i,t+h|t}| = \tilde{a}^p + \tilde{b}^p |e_{i,t+h|t-1}| + \tilde{\varepsilon}_{i,t+h}$. Black whiskers represent one standard error above and below the point estimates from panel regressions. Standard errors are clustered at the quarter and forecaster levels.
Alt text (decorative): Bar charts showing the persistence of individual forecast errors (left panel) and absolute forecast errors (right panel) over time, comparing survey data (blue bars) with model simulations (green bars).

Figure 9: Implied Long-Run Growth Rate Derived from Forecasts



Notes: This graph shows the time series of forecasters' implied long-run house price growth rates, derived by combining their forecasts over the full forecast horizon. The grey dashed lines represent the 5th and 95th percentiles of the cross-sectional distribution of implied long-run house price growth rates (quarterly), computed from forecasts for the current calendar year (accounting for the portion already realized year-to-date) and forecasts for each of the next four calendar years: $\left(\frac{\prod_{h=0}^{4}(1+E_{it}[y_h])}{(1+y_{YTD})}\right)^{1/(20-q_s+1)} - 1$.
Alt text: A graph illustrating the time series of model-implied forecasters' long-run house price growth rates, with dashed lines representing the 5th and 95th percentiles of the cross-sectional distribution.

dent variable is the forecast error (column 7 versus column 5) or the absolute forecast error (column 11 versus column 9). The only forecaster characteristic that reliably correlates with smaller forecast errors is the length of experience. Relative to the reference group of forecasters whose career is fewer than 15 years long, forecasters with more than 35 years experience have 0.573% (0.335%) lower (absolute) forecast errors for quarterly HPA. They achieve these lower forecast errors by making higher average house price forecasts. These improvements in forecast errors represent 18% (8%) of the mean and 13% (9%) of the standard deviation of the forecast error (absolute forecast error).

Table 6: Demographic Characteristics and HPA Forecasts

	Forecasts				**Forecast Errors**				**Absolute Forecast Errors**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Forecast Horizon (qtrs.)		-0.040	-0.040	-0.048*		0.183***	0.183***	0.182***		0.121**	0.121**	0.123**
		(0.028)	(0.028)	(0.027)		(0.050)	(0.050)	(0.051)		(0.048)	(0.048)	(0.049)
Length of Exp. (yrs.)												
<15		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)
15-25		-0.113	-0.039	-0.115		0.079	0.025	0.074		0.149	0.141	0.139
		(0.206)	(0.196)	(0.221)		(0.207)	(0.196)	(0.223)		(0.154)	(0.145)	(0.165)
25-35		0.138	0.160	0.223		-0.132	-0.135	-0.232		-0.031	-0.021	-0.114
		(0.203)	(0.209)	(0.215)		(0.199)	(0.204)	(0.212)		(0.141)	(0.142)	(0.148)
>35		0.621***	0.572**	0.613**		-0.621**	-0.573**	-0.618**		-0.343**	-0.335**	-0.339*
		(0.228)	(0.239)	(0.245)		(0.233)	(0.235)	(0.252)		(0.171)	(0.159)	(0.182)
Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Other Demographics												
Highest Degree, Major, Skill Num., Num. of Followers		Y	Y	Y		Y	Y	Y		Y	Y	Y
Region FE (CBSA)			Y				Y				Y	
Lagged 1Y Local HPA				Y				Y				Y
Mean of Dep. Var.	3.266	3.266	3.266	3.266	3.217	3.217	3.217	3.217	4.000	4.000	4.000	4.000
Std. of Dep. Var.	2.829	2.829	2.829	2.829	4.320	4.320	4.320	4.320	3.608	3.608	3.608	3.608
R-Squared	0.140	0.182	0.201	0.183	0.209	0.276	0.283	0.273	0.170	0.216	0.222	0.217
Num. of Obs.	29,340	16,835	16,835	15,360	25,577	14,667	14,667	13,291	25,577	14,667	14,667	13,291

Notes: The region control is an indicator variable for being located in one of the largest Core-Based Statistical Areas (CBSAs), in one of the remaining CBSAs or a non-metropolitan area, or abroad. A separate CBSA indicator variable is included if there are at least five different forecasters from this region. There are 11 separate CBSA categorical variables, including: 1) New York-Newark-Jersey City, 2) Los Angeles-Long Beach-Anaheim, 3) Chicago-Naperville-Elgin, 4) Washington-Arlington-Alexandria, 5) Philadelphia-Camden-Wilmington, 6) Miami-Fort Lauderdale-West Palm Beach, 7) Boston-Cambridge-Newton, 8) San Francisco-Oakland-Fremont, 9) Seattle-Tacoma-Bellevue, 10) San Diego-Chula Vista-Carlsbad, and 11) Charlotte-Concord-Gastonia. Standard errors are clustered at the quarter and forecaster level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

We compare this experience effect to the one implied by the model. To that end, We conduct the following simulation. A forecaster who is born on date $t_{i,b}$, starts her career on date $t_{i,c}$, enters in the survey panel in $t_{i,en}$ and exits in $t_{i,ex}$ has an experience (quarters of learning) equal to $n_i = (t_{i,en} - t_{i,c}) + 1$ and makes HPA predictions for $(t_{i,ex} - t_{i,en}) + 1$ quarters. We assume that this forecaster has prior beliefs at $t_{i,c}$ that are determined by her lived experience of HPA as in Malmendier and Nagel (2011): $\mu_{\mu,i} = \sum_{j=0}^{t_{i,c}-t_{i,b}} (1-\nu)^j / (\sum_{j=0}^{t_{i,c}-t_{i,b}} (1-\nu)^j) y_{t_{i,c}-j}$. We set the value for $1 - \nu = 0.917$, corresponding to an annual decay parameter of around 0.7. By construction, the simulated panel matches the observed distribution of forecaster experience and age, and exactly replicates the entry and exit dates of each individual forecaster in the survey data. Heterogeneity in forecasts arises from heterogeneity in experience and heterogeneity in priors. The latter is affected most by (recent) HPA realizations prior to the start date of the forecaster's career. We then estimate the same panel regression on simulated data as we estimated in the actual data. The results from the model simulations in Table 7 are close to the results from the actual data in Table 6. The simulations also generate (i) a modest bump in R^2 from adding forecaster experience as a determinant of forecasts and forecast errors, and (ii) a significant reduction in mean (absolute) forecast errors for forecasters with more experience.[32]

Table 7: Effect of Experience Length on Forecasts and Forecast Errors in Simulation

	Forecasts		**Forecast Errors**		**Absolute Forecast Errors**	
	(1)	(2)	(3)	(4)	(5)	(6)
Forecast Horizon (qtrs.)		-0.007		0.153***		0.113***
		(0.030)		(0.043)		(0.035)
Length of Experience (yrs.)						
<15		(omitted)		(omitted)		(omitted)
15-25		-0.623**		0.623**		0.288
		(0.293)		(0.293)		(0.200)
25-35		0.593*		-0.593*		-0.270
		(0.324)		(0.324)		(0.187)
>35		1.026***		-1.026***		-0.436**
		(0.345)		(0.345)		(0.204)
Survey Quarter FE	Y	Y	Y	Y	Y	Y
R-Squared	0.340	0.388	0.262	0.330	0.219	0.268
Number of Observations	24,097	24,097	24,097	24,097	24,097	24,097

Notes: This table presents the effects of length of experience on forecasts, forecast errors, and absolute forecast errors in the simulation based on an extension of the benchmark model. A long-term HPA time series is assembled by combining (i) annual HPA data from 1891 to 1974 from the JST Macrohistory database (Jordà et al., 2019), (ii) quarterly HPA data from 1975 to 1995 from the Freddie Mac House Price Index (FMHPI), and (iii) quarterly HPA data from 1996 to 2023 from the Zillow Home Value Index (ZHVI). A panel of forecasters is simulated to closely match the key characteristics of real forecasters in the survey data. For each real forecaster (with LinkedIN data available, 239 in total), a simulated forecaster is generated to exactly match her/his a) date of birth $t_{i,b}$, b) length of experience n_i, and c) time of entering and exiting the survey ($t_{i,en}$ and $t_{i,ex}$). The date of birth is inferred from the forecasters' educational information, and the length of experience is defined as the number of quarters between the start of their career ($t_{i,c}$) and the time they make forecasts. Each simulated forecaster i has a prior belief at $t_{i,c}$, starts learning about HPA, and begins making forecasts from $t_{i,en}$ to $t_{i,ex}$. At $t_{i,c}$, the prior $\mu_{\mu,i}$ for forecaster i is a weighted average of the experienced HPA during their lifetime, calculated as $\sum_{j=0}^{k}\left((1-\nu)^j / \sum_{j=0}^{k}(1-\nu)^j\right) y_{t_{i,c}-j}$, where $k = t_{i,c} - t_{i,b}$ and $\nu = 0.08$. Other prior parameters are set as follows: $\sigma_\mu = 0.185$, $\mu_\rho = 0.675$, and $\sigma_\rho = 0.020$, consistent with the benchmark model.

Fourth, the model predicts that the forecast dispersion should be smaller among forecasters with more experience than among forecasters with less experience, a direct consequence of longer periods of learning

[32] The results are similar for values of $1-\nu$ that imply annual decay parameters between 0.5 and 0.9.

for more experienced forecasters. Table A6 shows that this prediction is borne out in the data. Forecast dispersion measured within experience groups decreases in experience. The decline is measured precisely.

These four pieces of empirical evidence based on individual forecaster panel data significantly bolster our learning explanation.

6 The Impact of Behavioral Biases

We now enrich the learning model with one of two different behavioral frictions and show that these frictions help to further improve the fit along two important dimensions: the short-run CG regressions slope and the short-run cross-sectional forecast dispersion.

6.1 Overconfidence

We allow for a difference in the true and the perceived signal variance, modulated by the parameter ϕ. The rational learning benchmark, where the two are the same, is $\phi = 1$. When $\phi > 1$, the true signal variance is higher than the perceived one. Forecasters are over-confident about the informational content of the signal. The opposite is true for $\phi < 1$.

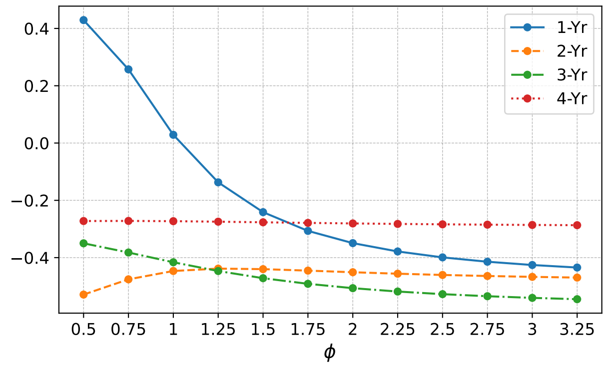
Panel (a) of Figure 10 shows the model-implied CG slope as we change the value of ϕ. As ϕ increases from 0.75 (under-confidence) to 3.25 (over-confidence), the CG slope based on 1-year forecasts changes from +0.4 to -0.4. Overconfidence can generate short-run overreaction in the housing market, similar to what Broer and Kohlhas (2022) finds for a broader set of macroeconomic forecasts.

On the other hand, the CG slopes based on forecasts of longer horizons barely change. Forecasts of 3-4 year still exhibit negative CG slopes, due solely to forecasters learning about the long-run mean. This result is intuitive. A large ϕ should lead to a more negative CG regression slope as over-confident forecasters revise their forecasts *too much* compared to the true information content of their signals. Furthermore, the effect of ϕ should be most visible in the short-term forecasts, as private signals mostly concern the current period. After all, everybody agrees on the true, past house price growth, limiting the potency of over-confidence to affect longer-run forecasts.

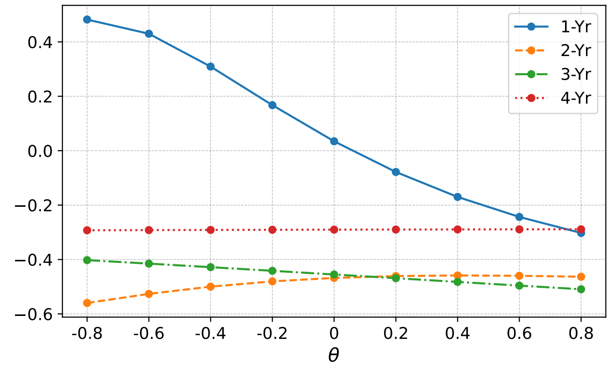
Panel (c) of Figure 10 demonstrates how the cross-sectional forecast dispersion changes with ϕ. As ϕ increases, forecast dispersion of all horizon increases. However, the 1-year forecast dispersion displays by far the strongest impact. Intuitively, overconfidence should increase the short-term forecast dispersion. As we discussed in Section 4.4.3, signal heterogeneity has limited impact on forecast dispersion because rational forecasters discount their signals when forming forecasts if their signals contain a lot of noise.

However, when forecasters over-estimate the precision of their signals, they discount the signal less than they should. In this setting, higher signal heterogeneity ($\overline{\Gamma}$) *does* translate into higher forecast dispersion.

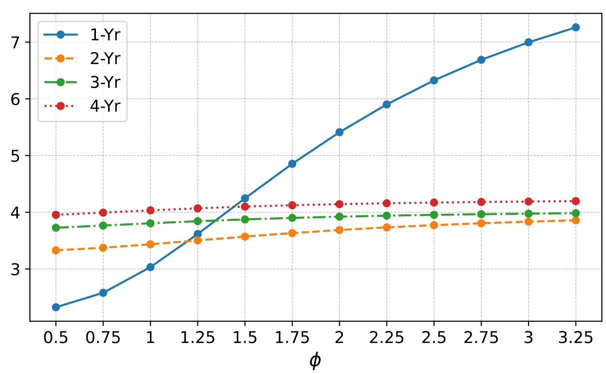
Figure 10: The Impact of Overconfidence (ϕ) and Diagnostic Expectations (θ)



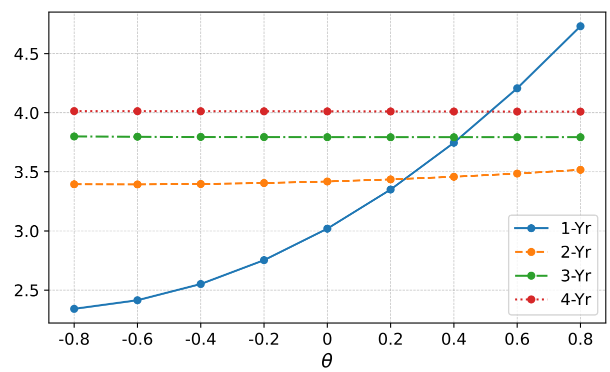
(a) Impact of ϕ on CG-b coefficients



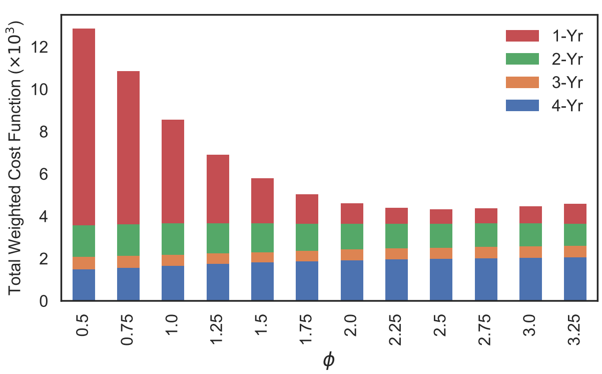
(b) Impact of θ on CG coefficients



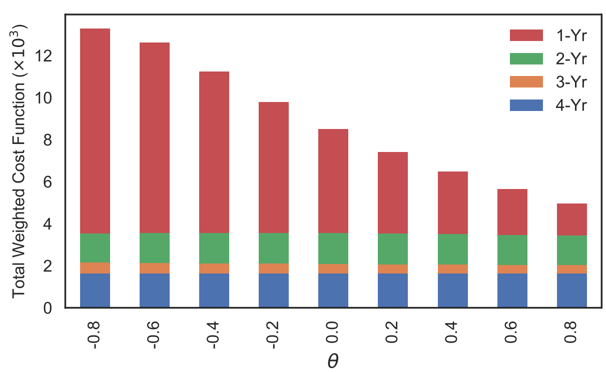
(c) Impact of ϕ on forecast dispersion



(d) Impact of θ on Dispersion



(e) Impact of ϕ on total weighted cost function



(f) Impact of θ on total weighted cost function

Notes: This figure shows the impact of behavioral frictions on the model's fit. In the model with overconfidence, the parameter ϕ controls the ratio of the objective signal variance, $\overline{\Gamma}$, to the subjective variance, Γ. In the model that incorporates diagnostic expectations, the parameter θ controls the degree of distortion in forecasters' estimates of the state of the housing market, $\hat{x}_{i,t}$, when applying the Kalman filter, given their posterior estimates for the parameters. The value of θ is varied in increments of 0.2. Panels (e) and (f) show the impact of behavioral frictions on the value of the overall cost function. The height of the bars represents the value of the total weighted cost function, which reflects the goodness of fit for the 36 moments presented in Table 4. The cost function is decomposed into contributions from each of the four different forecast horizons. In each panel, all parameters are set to those in the benchmark model, except for the parameter of interest.

Alt text: A figure showing the impact of behavioral frictions, including overconfidence and diagnostic expectations (right column), on the model's implied forecasts in terms of the CG coefficients, forecaster dispersion and the overall fit.

6.2 Diagnostic Expectations

We allow forecasters to display diagnostic expectations, modulated by the behavioral filter parameter θ as proposed in the diagnostic expectations literature (Bordalo et al., 2019, 2020, among others). The higher θ the more weight forecasters put on recent data points when applying the Kalman filter. Figure 10 Panels (b) and (d) show how the individual CG regression slope (Panel b) and the cross-sectional forecast variance (Panel d) are affected by changing θ up or down by 0.2 increments. The rational learning benchmark model sets $\theta = 0$.

The diagnostic filter affects short forecasting horizons (1-to-2 years) more than long horizons (3-4 years). This is intuitive given that the behavioral filter affects the dynamics of the latent state and the latent state has a modest half-life.

While the rational learning model can generate a negative slope coefficient in the 1-year CG regression, at least for some combinations of parameters (different from the benchmark model), the distance between the one-year CG b coefficient in the benchmark model and the data is substantial. Consistent with the intuition in Bordalo et al. (2020), diagnostic expectations can help generate a *strongly* negative CG short-horizon slope coefficient. In fact, the best fit for the one-year CG slope and cross-sectional forecast variance are obtained for a value around $\theta = 0.8$, close to the average of the range of values estimated by Bordalo et al. (2020).

Diagnostic expectations also help resolve the tension between short-run and long-run forecast dispersion. Learning in the presence of uncertainty about the long-run mean growth rate $\overline{\mu}$ generates high long-run forecast dispersion, while diagnostic expectations helps generate the high short-run forecast dispersion.

6.3 The Impact of Behavioral Biases on Model Fit

Figure 10 Panels (e) and (f) compare the cost function values for different levels of overconfidence (Panel e) or diagnostic expectations (Panel f). Incorporating either overconfidence ($\phi > 1$) or a diagnostic filter ($\theta > 0$) improves the model fit, i.e., reduces the total cost function. An overconfidence level of $\phi = 2.5$ and a diagnostic expectation parameter of $\theta = 0.8$ can reduce the cost function by 49% and 42%, respectively, bringing the model significantly closer to the data. Columns 3 and 4 of Table 4 show that the reduction in total costs comes from both the reduction in regression coefficient estimates (first 32 moments), and the forecast dispersion (4 moments). The improvement comes mainly from the short-term CG slopes and short-term dispersion. Figure 10 Panels (e) and (f), which decompose the total cost functions into the contributions from each forecast horizon, show that nearly the entire improvement comes from the one-year-horizon forecasts. The long-run moments, on the other hand, are not much impacted by either behavioral frictions.

Finally, we ask whether the behavioral frictions can account for the data *without* learning. To that end, we shut down parameter

uncertainty and optimize over the remaining parameters, including ϕ and θ, to obtain the best possible fit. The results are reported in Table A11 and show that the models with only overconfidence or only diagnostic expectations have significantly worse model fits.

These results highlight the importance of having a rich enough rational benchmark model when making inference on the rationality of subjective forecasts. With the rational learning benchmark, we can make quantitative inference on *how much* and *where* behavioral biases play a role in shaping expectations. When added to a full-fledged learning model with uncertainty about the long-run mean, either overconfidence or diagnostic expectations are very helpful to generate realistic short-run CG coefficients and short-run forecast dispersion. These biases have little impact on the longer-run moments, which are successfully accounted for by learning dynamics.

7 Assessing Importance of Term-Structure and Cross-Sectional Data

Most survey data sources contain limited information on the term structure of forecasts and/or limited information on the cross-section of forecasters. Indeed, the typical survey only contains data on the one-period-ahead consensus forecast. This section discusses the importance of term structure and cross-sectional information for parameter inference and hence for the interpretation of forecast data.

7.1 The Impact of Forecaster Heterogeneity

We consider a first special case of the benchmark model estimation where the econometrician only has access to consensus forecast data, i.e., the average across the forecasters. This case serves as a tool for explaining the role of forecaster heterogeneity for inference. It allows us to speak to the literature that studies individual versus consensus forecasts (e.g., Bordalo et al., 2020).

We use 32 moments in the estimation of this model. Compared to the benchmark model, we drop the four moments that measure forecast dispersion since one cannot measure dispersion from consensus forecasts. These 32 moments are the same as in the benchmark model, except that they are based on consensus forecasts rather than individual forecasts.[33] This only makes a difference for the CG slope, which changes sign when going from individual to aggregate forecasts.

[33] The parameters are estimated by minimizing the distance between the model-implied and the empirical consensus moments: $\hat{\delta}_{NoDisp} = \text{argmin}_{\delta}\, Q_{cs}(\delta) = \text{argmin}_{\delta}\, g_{cs}(\delta)'\hat{W}_{cs} g_{cs}(\delta)$, where $\hat{W}_{cs} = \text{diag}(\hat{\Omega}_{cs})^{-1}$. Similarly, $\hat{\Omega}_{cs}$ is the covariance matrix of consensus moments, calculated using the influence function approach.

Table 3 shows the parameter estimates from the simpler model estimated on consensus data in the column labeled "No Disp". The main difference with the benchmark model is the much lower estimate for prior uncertainty about the long-run mean house price growth rate, σ_μ. The latter is 0.009 in the No Disp model, twenty times smaller than the 0.185 in the benchmark model.

Section 4 extensively discussed the crucial role of σ_μ. Without a high enough value for σ_μ, beliefs about mean house price growth become too "dogmatic." The $\sigma_\mu = 0.009$ parameter value effectively shut down the "learning about the long-run mean" channel. This channel is the key driver of the negative CG regression slopes at horizons $h = 2 - 4$.

Table 4 shows the resulting model fit for the model without investor heterogeneity in column "No Disp.". Not surprisingly, the model produces no cross-sectional forecast dispersion. The fit for the remaining 32 moments deteriorates dramatically from 2,686 in the benchmark to 54,415.

As anticipated, the model without forecaster dispersion cannot consistently produce negative individual CG slopes because of its low value for σ_μ. Rather the model implies similar CG slopes at the individual and the consensus level. Since the latter tend to be positive in the data, and well-matched (see Table 5), the No Disp model generally implies the wrong sign for the individual CG slopes. The remaining moments, including the sensitivity of forecasts to lagged house price growth, are matched reasonably well.

In sum, having access to panel survey data is crucial when it comes to identifying the important role of learning about the long-run mean, as revealed by the negative sign of the long-run CG regression slopes.

7.2 Benchmark Model Estimation without Term Structure

We consider a second special case where the econometrician does not have access to data on the term structure of forecasts, only to the one-year ahead forecast ($h = 1$). The estimation of this model involves 9 rather than 36 moments. Its purpose is to illustrate the role of term structure of forecaster data. The parameter estimates are in the column "No LR." of Table 3. The main differences with the benchmark estimates are (i) a much lower σ_μ, (ii) a higher λ, and (iii) a higher σ_ρ. This shows that the lack of long-run forecasts can lead to significantly different parameter inference.

Table 4 shows the resulting model fit for the model without long-run forecast moments in the column "No LR.". It reports all model-implied moments even though only the 1-year moments are used in estimation. The overall fit for all 36 moments is much worse than in the benchmark model; the distance between the model-implied and observed moments rises from 8,551 to 169,190. This shows that there is important information in long-horizon forecasts that does not get captured when estimating the model on one-year data and studying its implications for longer forecast horizons.

Given the parameter estimates, the “No LR." model misses badly on the CG slopes, especially at the 4-year horizon, for the same reason that the “No Disp." model did. It also misses badly on the forecast dispersion moments, significantly overstating the dispersion at longer horizons. Finally, the “No LR." model generates excess sensitivity to lagged HPA which arises from the higher value for the posterior persistence parameter in the no LR model. In turn, this high posterior estimate results from the higher prior mean but especially from the higher prior uncertainty about the persistence parameter. Intuitively, when agents are more uncertain about the persistence parameter, they are likely to revise it more strongly. In the benchmark model, the prior on persistence is tight, limiting the role for learning about $\overline{\rho}$.

8 Under-Extrapolation and Overreaction

The previous sections have shown the co-existence of under-extrapolation, the positive correlation between HPA forecast errors and lagged HPA, and overreaction, the negative relationship between forecast errors and lagged forecast revisions. In this final section, we explore how common the combination of overreaction and under-extrapolation is among a broad set of macroeconomic variables using Survey of Professional Forecasts data.

Table 8 shows that the combination exhibited by professional HPA forecasts is the most common, with 6 instances of macro series displaying overreaction and under-extrapolation, followed by 4 instances of overreaction and over-extrapolation, 3 instances of underreaction and under-extrapolation, and 2 instances of underreaction and over-extrapolation. This shows that the empirical pattern in the HPA data is of broader interest to the belief formation literature.

Moreover, the learning model is capable of explaining three of the four pattern combinations, exhibited by 13 out of the 15 macro series. Indeed, Figure 11 illustrates how our model can deliver different combinations of over/under-extrapolation and over/under-reaction for different combinations of parameters (σ_μ, μ_ρ). A small ridge of parameter combinations generates the observed combination of overreaction and under-extrapolation, providing tight parameter identification. The flexibility of our learning model to generate various combinations of extrapolation and reaction patterns sets it apart from other candidate explanations, including simpler learning models.[34]

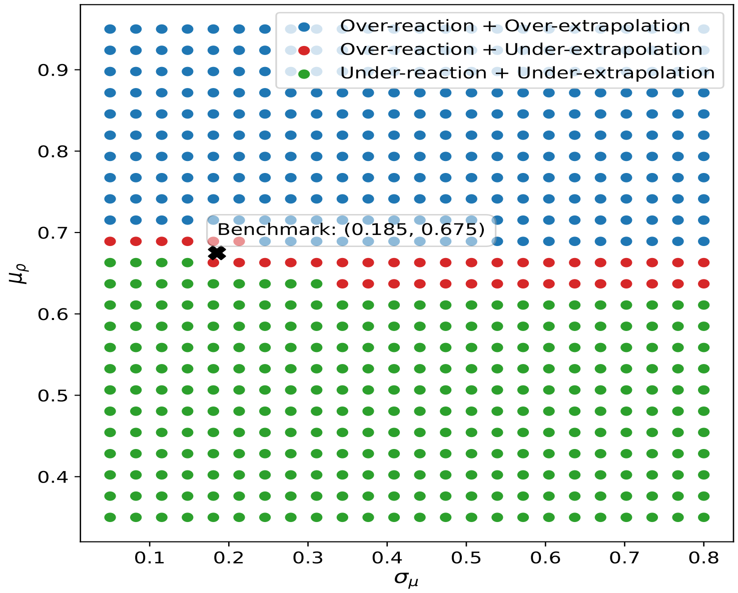
[34] For example, in the experimental setting of Afrouzi et al. (2023), forecasters receive news about the lagged variable. When the variable to be predicted follow an AR(1), overreaction and over-extrapolation necessarily coincide in their setting. This is not true in our richer learning model, where the HPA process contains a latent persistent component and where forecasters have other, heterogeneous sources of private information besides lagged HPA.

Table 8: Extrapolation Regression and CG Regression Results for Other Macroeconomic Forecasts

	Regressions of Forecast Errors on Past Growth (b^{ex})	**CG Regressions (b^{cg})**
Nominal GDP	0.029	-0.218
Real GDP	-0.041	-0.152
GDP Price Index Inflation	0.141	0.178
CPI	-0.308	-0.190
Real Consumption	0.246	-0.241
Industrial Production	-0.081	-0.161
Real Non-Residential Investment	-0.062	0.077
Real Residential Investment	0.292	0.009
Real Federal Government Consumption	0.216	-0.588
Real State & Local Govt Consumption	0.221	-0.429
Housing Start	0.386	-0.236
Unemployment	0.052	0.337
3M Treasury Rate	-0.109	0.274
10Y Treasury Rate	0.041	-0.185
AAA Corporate bond Rate	-0.031	-0.222

Notes: This table shows the coefficient estimates of the extrapolation regression (forecast error on past growth: $e_{i,t+h|t} = a^{ex} + b^{ex} y_{t,t-j} + \varepsilon_{i,t+h}$) and of the CG regression (4) for forecasts of different macroeconomic variables from the Survey of Professional Forecasters (SPF). Variables are constructed following the methods in Bordalo et al. (2020) and for the same sample period of 1968-2016. The forecasts are for three-quarter-ahead growth rates. The independent variable in the extrapolation regression is realized three-quarter growth.

Figure 11: Extrapolation and Reaction



Notes: This graph shows how different values of forecaster conviction (σ_μ) and prior belief about the persistence parameter (μ_ρ) affect the level of extrapolation and reaction. Other parameters are set to the values in the benchmark model. Extrapolation and CG regressions are estimated following the same procedure as in Table (8). Over- and under-extrapolation are defined as $b^{ex} < 0$ and $b^{ex} > 0$. Over- and under-reaction are defined as $b^{cg} < 0$ and $b^{cg} > 0$.
Alt text: A graph showing how different values of forecaster conviction (σ_μ) and prior belief about the persistence parameter (μ_ρ) affect the level of extrapolation and reaction implied by the model. The black cross indicates the estimated parameter values of the benchmark model.

9 Conclusion

We argue that a rational learning model with parameter uncertainty about the long-run mean and forecaster heterogeneity goes a long way towards accounting for the features of forecast data. We do so in the context of U.S. house price forecasts, exploiting a unique panel data set of individual-level forecasts that contains a rich term structure of forecasts. Theories of behavioral biases such as overconfidence and diagnostic expectations are useful complements that help the learning model account for the high degree of overreaction and dispersion in short-horizon forecasts, while leaving the longer-run forecasts largely unaffected. Estimating our model on macro-economic and financial survey data is another exciting avenue for future work.

References

Adam, Klaus, Pei Kuang, and Shihan Xie. 2024. "Overconfidence in Private Information Explains Biases in Professional Forecasts." Working Paper University College London.

Afrouzi, Hassan, Spencer Y Kwon, Augustin Landier, Yueran Ma, and David Thesmar. 2023. "Overreaction in Expectations: Evidence and Theory*." *The Quarterly Journal of Economics*, 138(3): 1713–1764.

Aladangady, Aditya. 2017. "Housing Wealth and Consumption: Evidence from Geographically-Linked Microdata." *American Economic Review*, 107(11): 3415–46.

Andersen, Steffen, Cristian Badarinza, Lu Liu, Julie Marx, and Tarun Ramadorai. 2022. "Reference dependence in the housing market." *American Economic Review*, 112(10): 3398–3440.

Andrade, Philippe, Richard K Crump, Stefano Eusepi, and Emanuel Moench. 2016. "Fundamental disagreement." *Journal of Monetary Economics*, 83: 106–128.

Angeletos, George-Marios, Zhen Huo, and Karthik A Sastry. 2021. "Imperfect macroeconomic expectations: Evidence and theory." *NBER Macroeconomics Annual*, 35(1): 1–86.

Armona, Luis, Andreas Fuster, and Basit Zafar. 2019. "Home price expectations and behaviour: Evidence from a randomized information experiment." *The Review of Economic Studies*, 86(4): 1371–1410.

Bailey, Michael, Eduardo Dávila, Theresa Kuchler, and Johannes Stroebel. 2019. "House price beliefs and mortgage leverage choice." *The Review of Economic Studies*, 86(6): 2403–2452.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel. 2018. "The economic effects of social networks: Evidence from the housing market." *Journal of Political Economy*, 126(6): 2224–2276.

Barberis, Nicholas C, and Lawrence J Jin. 2023. "Model-free and model-based learning as joint drivers of investor behavior." National Bureau of Economic Research.

Barberis, Nicholas, Robin Greenwood, Lawrence Jin, and Andrei Shleifer. 2015. "X-CAPM: An extrapolative capital asset pricing model." *Journal of Financial Economics*, 115(1): 1–24.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer. 2019. "Diagnostic expectations and stock returns." *The Journal of Finance*, 74(6): 2839–2874.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2020. "Overreaction in Macroeconomic Expectations." *American Economic Review*.

Broer, Tobias, and Alexandre N. Kohlhas. 2022. "Forecaster (Mis-)Behavior." *The Review of Economics and Statistics*, 1–45.

Case, Karl E, and Robert J Shiller. 1988. "The Behavior of Home Buyers in Boom and Post-Boom Markets." National Bureau of Economic Research.

Case, Karl E, Robert J Shiller, and Anne K Thompson. 2012. "What Have They Been Thinking?: Homebuyer Behavior in Hot and Cold Markets." *Brookings Papers on Economic Activity*, 2012(2): 265–315.

Chopra, Felix, Christopher Roth, and Johannes Wohlfart. 2023. "Home price expectations and spending: Evidence from a field experiment." *Available at SSRN 4452588*.

Cogley, Timothy, and Thomas J Sargent. 2008. "Anticipated utility and rational expectations as approximations of Bayesian decision making." *International Economic Review*, 49(1): 185–221.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *American Economic Review*, 105(8): 2644–78.

Coibion, Olivier, Dimitris Georgarakos, Yuriy Gorodnichenko, Geoff Kenny, and Michael Weber. 2024. "The Effect of Macroeconomic Uncertainty on Household Spending." *The American economic review*, 114(3): 645–677.

Collin-Dufresne, Pierre, Michael Johannes, and Lars A. Lochstoer. 2016. "Parameter Learning in General Equilibrium: The Asset Pricing Implications." *American Economic Review*, 106(3): 664–98.

Croce, Mariano M, Martin Lettau, and Sydney C Ludvigson. 2015. "Investor information, long-run risk, and the term structure of equity." *The Review of Financial Studies*, 28(3): 706–742.

D'Acunto, Francesco, and Michael Weber. 2024. "Why Survey-Based Subjective Expectations are Meaningful and Important." *Annual Review of Economics*, forthcoming.

Daniel, Kent, and David Hirshleifer. 2015. "Overconfident Investors, Predictable Returns, and Excessive Trading." *Journal of Economic Perspectives*, 29(4): 61–88.

Erickson, Timothy, and Toni M Whited. 2002. "Two-step GMM estimation of the errors-in-variables model using high-order moments." *Econometric Theory*, 18(3): 776–799.

Farmer, Leland, Emi Nakamura, and Jón Steinsson. 2023. "Learning About the Long Run." *Journal of Political Economy*.

Favilukis, Jack, Sydney C Ludvigson, and Stijn Van Nieuwerburgh. 2017. "The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium." *Journal of Political Economy*, 125(1): 140–223.

Friedman, Benjamin M. 1979. "Optimal expectations and the extreme information assumptions of ârational expectationsâ macromodels." *Journal of Monetary Economics*, 5(1): 23–41.

Fuster, Andreas, David Laibson, and Brock Mendel. 2010. "Natural expectations and macroeconomic fluctuations." *Journal of Economic Perspectives*, 24(4): 67–84.

Fuster, Andreas, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar. 2022. "Expectations with endogenous information acquisition: An experimental investigation." *Review of Economics and Statistics*, 104(5): 1059–1078.

Gabaix, Xavier. 2019. "Behavioral Inattention." In *Handbook of Behavioral Economics: Applications and Foundations 1*. Vol. 2, , ed. B. Douglas Bernheim, Stefano DellaVigna and David Laibson, 261–343. Elsevier.

Giacoletti, Marco, Kristoffer T Laursen, and Kenneth J Singleton. 2021. "Learning from disagreement in the us treasury bond market." *The Journal of Finance*, 76(1): 395–441.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2021*a*. "Five facts about beliefs and portfolios." *American Economic Review*, 111(5): 1481–1522.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. 2021*b*. "The joint dynamics of investor beliefs and trading during the COVID-19 crash." *Proceedings of the National Academy of Sciences*, 118(4): e2010316118.

Glaeser, Edward L, and Charles G Nathanson. 2017. "An extrapolative model of house price dynamics." *Journal of Financial Economics*, 126(1): 147–170.

Goldstein, Nathan, and Yuriy Gorodnichenko. 2023. "Expectations Formation and Forward Information." National Bureau of Economic Research.

Greenwald, Daniel L, and Adam Guren. 2021. "Do credit conditions move house prices?" National Bureau of Economic Research.

Greenwood, Robin, and Andrei Shleifer. 2014. "Expectations of returns and expected returns." *The Review of Financial Studies*, 27(3): 714–746.

Guren, Adam M, Alisdair McKay, Emi Nakamura, and Jón Steinsson. 2021. "Housing wealth effects: The long view." *The Review of Economic Studies*, 88(2): 669–707.

Jacobson, Margaret. 2022. “Beliefs, Aggregate Risk, and the US Housing Boom.”

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor. 2019. “The Rate of Return on Everything, 1870â2015*.” *The Quarterly Journal of Economics*, 134(3): 1225–1298.

Kacperczyk, Marcin, Stijn Van Nieuwerburgh, and Laura Veldkamp. 2016. “A Rational Theory of Mutual Funds’ Attention Allocation.” *Econometrica*, 84(2): 571–626.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante. 2020. “The housing boom and bust: Model meets evidence.” *Journal of Political Economy*, 128(9): 3285–3345.

Kindermann, Fabian, Julia Le Blanc, Monika Piazzesi, and Martin Schneider. 2021. “Learning about housing cost: Survey evidence from the german house price boom.” National Bureau of Economic Research.

Kohlhas, Alexandre N., and Ansgar Walther. 2021. “Asymmetric Attention.” *American Economic Review*, 111(9): 2879–2925.

Kozlowski, Julian, Laura Veldkamp, and Venky Venkateswaran. 2020. “The tail that wags the economy: Beliefs and persistent stagnation.” *Journal of Political Economy*, 128(8): 2839–2879.

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

Kuchler, Theresa, Monika Piazzesi, and Johannes Stroebel. 2023. “Housing market expectations.” In *Handbook of Economic Expectations*. 163–191. Elsevier.

Kumar, Saten, Yuriy Gorodnichenko, and Olivier Coibion. 2023. “The effect of macroeconomic uncertainty on firm decisions.” *Econometrica*, 91(4): 1297–1332.

Landvoigt, Tim. 2017. “Housing demand during the boom: The role of expectations and credit constraints.” *The Review of Financial Studies*, 30(6): 1865–1902.

Lewellen, Jonathan, and Jay Shanken. 2002. “Learning, asset-pricing tests, and market efficiency.” *The Journal of Finance*, 57(3): 1113–1145.

Liu, Haoyang, and Christopher Palmer. 2021. “Are stated expectations actual beliefs? New evidence for the beliefs channel of investment demand.” National Bureau of Economic Research.

Lorenzoni, Guido. 2009. “A Theory of Demand Shocks.” *American Economic Review*, 99(5): 2050–84.

Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3): 769–803.

Malmendier, Ulrike, and Laura Veldkamp. 2022. "Information Resonance."

Malmendier, Ulrike, and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?*." *The Quarterly Journal of Economics*, 126(1): 373–416.

Malmendier, Ulrike, and Stefan Nagel. 2016. "Learning from inflation experiences." *The Quarterly Journal of Economics*, 131(1): 53–87.

Mankiw, N Gregory, Ricardo Reis, and Justin Wolfers. 2003. "Disagreement about inflation expectations." *NBER Macroeconomics Annual*, 18: 209–248.

Mincer, Jacob A, and Victor Zarnowitz. 1969. "The evaluation of economic forecasts." In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. 3–46. NBER.

Nagel, Stefan, and Zhengyang Xu. 2022. "Asset pricing with fading memory." *The Review of Financial Studies*, 35(5): 2190–2245.

Patton, Andrew J., and Allan Timmermann. 2010. "Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion." *Journal of Monetary Economics*, 57(7): 803–820.

Renxuan, Wang. 2020. "Subjective return expectations." *Available at SSRN 3740792*.

Singleton, Kenneth J. 2021. "Presidential address: How much "rationality" is there in bondâmarket risk premiums?" *The Journal of Finance*, 76(4): 1611–1654.

Sodini, Paolo, Stijn Van Nieuwerburgh, Roine Vestman, and Ulf von Lilienfeld-Toal. 2023. "Identifying the benefits from homeownership: A Swedish experiment." *American Economic Review*.

Timmermann, Allan G. 1993. "How learning in financial markets generates excess volatility and predictability in stock prices." *The Quarterly Journal of Economics*, 108(4): 1135–1145.

Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2009. "Information immobility and the home bias puzzle." *The Journal of Finance*, 64(3): 1187–1215.

Van Nieuwerburgh, Stijn, and Laura Veldkamp. 2010. "Information Acquisition and Under-Diversification." *The Review of Economic Studies*, 77(2): 779–805.

Vestman, Roine, Jesper Bojeryd, Björn Tyrefors, and Dany Kessel. 2023. "The Housing Wealth Effect: Quasi-experimental Evidence."

Weber, Michael, Francesco DâAcunto, Yuriy Gorodnichenko, and Olivier Coibion. 2022. "The subjective inflation expectations of households and firms: Measurement, determinants, and implications." *Journal of Economic Perspectives*, 36(3): 157–184.

Online Appendix

A The Timing of Pulsenomics Survey

The Pulsenomics Survey asks its participants to make forecasts for the end-of-year HPA of different forecasting horizons at quarterly frequency. We therefore design the model and notation to fully take advantage of the survey structure.

First, define the quarterly HPA as the log change in the house price index:

$$y_t = log\left(\frac{P_t}{P_{t-1}}\right).$$

In each quarter, the Pulsenomics Survey is conducted at the beginning of the second month of each quarter. For example, the survey for 2020-Q1 is conducted in the middle of February. Zillow Home Value Index (ZHVI) for a given month is published on the third Thursday of the following month. Thus, the latest available ZHVI we have for 2020-Q1 Survey is ZHVI of Dec 2019. Therefore, in the 2020-Q1 survey, the latest information that experts actually use is from 2019-Q4. Suppose the survey is conducted at t_s (in year y_s, quarter q_s), e.g., 2020Q1. Then experts actually form expectation using the data of 2019Q4:

- t is the end-of-quarter date of the experts' information set, e.g. $t = 2019Q4$;
- $h = 0, 1, 2, 3, 4$ represent forecasting horizons, e.g. $h = 0$ corresponds to a forecast for the current year, $h = 1$ represents a forecast for the next year, etc.
- $q_s = 1, 2, 3, 4$ are the calendar quarters in which the surveys are conducted in real time;
- $q = q_s - 1$ is the effective quarter end, based on the information from which the experts make their forecasts,

The target for h-year ahead forecast in the survey is defined as y_{t+h}, which denotes the h-year ahead end-of-year annual HPA (from Dec to Dec). Formally,

$$y_{t+h} = log\left(\frac{P_{t+4(h+1)-q}}{P_{t+4(h+1)-q-4}}\right).$$

Note that the forecast horizons are different for surveys conducted in different quarters.

As an example for why the quarter in which the survey is conducted matters, consider the forecast for the next year ($h = 1$), conducted in 2020-Q2. In this case, forecasters' information set is up until 2020-Q1.

The forecast target would be the price growth from 2020-Q4 (2 quarters from now) to 2021-Q4 (6 quarters from now). Indeed, we have $t+4(h+1)-q$ = 2020Q1+8$Q-1Q$ = 2021Q4 and $t+4(h+1)-q-4$ = 2020Q1+8$Q-1Q-4Q$ = 2020Q4. Notice that for the current year forecast ($h=0$), at the time that forecasts are made, there had already been realized price growth observed within the forecasting horizon, which will be taken into account in the framework.

Define forecaster i's (subjective) expectation of h-year ahead HPA as

$$E_{i,t}(y_{t+h}) = E_{i,t}\left[log\left(\frac{P_{t+4(h+1)-q}}{P_{t+4(h+1)-q-4}}\right)\right].$$

Note that $E_{i,t-1}(y_{t+h})$ represents the expectation of the same target y_{t+h} formed in the survey conducted in the previous quarter.

Finally, we define y_{t-j} as lagged annualized house price growth over the past j years. Formally, y_{t-j} is calculated as

$$y_{t-j} = \frac{1}{j}log\left(\frac{P_t}{P_{t-4j}}\right)$$

B Posterior Distributions of Parameters and States

B.1 Prior Beliefs of Parameters and States

We estimate the six parameters that govern the prior distributions for the following parameters

$$\mu_0 \sim N(\mu_\mu, \sigma_\mu^2)$$
$$\rho_0 \sim N(\mu_\rho, \sigma_\rho^2)$$
$$\sigma_0^2 \sim IG(A, B)$$

Furthermore, we fix the initial beliefs of the states to be

$$x_1 \sim N(0, 1)$$

B.2 Gibbs Sampling

We start with an initial guess of the parameter $\theta^{(0)} = (\mu^{(0)}, \rho^{(0)}, \sigma^{(0)}, x_{1:T}^{(0)})'$. Subsequently, we conduct Bayesian updates given these initial draws as follows:

1. Draw $\rho^{(b+1)}|\mu^{(b)}, \sigma^{(b)}, x_{1:T}^{(b)}, y_{1:T}$ from

   $$\rho^{(b+1)} \sim N(\hat{\mu}_\rho, \hat{\sigma}_\rho^2)$$

   where

   $$\hat{\sigma}_\rho^2 \equiv [\sigma_\rho^{-2} + \frac{\Sigma_{s=2}^T (x_{s-1}^{(b)})^2}{(\sigma^{(b)})^2(1-\gamma^{(b)})}]^{-1}$$
   $$\hat{\mu}_\rho \equiv \hat{\sigma}_\rho^2 [\frac{\mu_\rho}{\sigma_\rho^2} + \frac{\Sigma_{s=2}^T x_{s-1}^{(b)} x_s^{(b)}}{(\sigma^{(b)})^2(1-\gamma^{(b)})}],$$

   as beliefs about ρ should satisfy:

   $$\frac{1}{\sqrt{1-\gamma^{(b)}}} x_t^{(b)} = \rho \frac{1}{\sqrt{1-\gamma^{(b)}}} x_{t-1}^{(b)} + \sigma^{(b)} \eta_t.$$

2. Draw $\mu^{(b+1)}|\rho^{(b+1)}, \sigma^{(b)}, x_{1:T}^{(b)}, y_{1:T,}$ from $N(\hat{\mu}_\mu, \hat{\sigma}_\mu^2)$, where

$$\hat{\sigma}_\mu^2 \equiv [\sigma_\mu^{-2} + \frac{T-1}{\gamma^{(b)}(\sigma^{(b)})^2} + \frac{1}{(\gamma^{(b)} + \Gamma^2)(\sigma^{(b)})^2}]^{-1}$$
$$\hat{\mu}_\mu \equiv \hat{\sigma}_\mu^2 [\frac{\mu_\mu}{\sigma_\mu^2} + \frac{\Sigma_{s=1}^{T-1}(y_s - x_s^{(b)})}{\gamma^{(b)}(\sigma^{(b)})^2} + \frac{y_T - x_T^{(b)}}{(\gamma^{(b)} + \Gamma^2)(\sigma^{(b)})^2}],$$

as belief updates about μ follows the regression:

$$y_t - x_t = \mu + e_t.$$

or

$$\begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_T - x_T \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \mu + \begin{pmatrix} e_1 \\ \vdots \\ e_T \end{pmatrix}, \begin{pmatrix} e_1 \\ \vdots \\ e_T \end{pmatrix} \sim N(0, \sigma^2 \underbrace{\begin{bmatrix} \gamma & & & 0 \\ & \ddots & & \\ & & \gamma & \\ 0 & & & \gamma + \Gamma^2 \end{bmatrix}}_{\Lambda})$$

3. Draw $\sigma^{(b+1)}|\mu^{(b+1)}, \rho^{(b+1)}, x_{1:T}^{(b)}, y_{1:T,}$ from $(\sigma^{(b+1)})^2 \sim IG(\hat{A}, \hat{B})$, where

$$\hat{A} \equiv A + \frac{2T-1}{2}$$
$$\hat{B} \equiv B + \frac{\Sigma_{s=1}^{T-1}(y_s - x_s^{(b)} - \mu^{(b+1)})^2}{2\gamma^{(b)}} + \frac{(y_T - x_T^{(b)} - \mu^{(b+1)})^2}{2(\gamma^{(b)} + \Gamma^2)} + \frac{\Sigma_{s=2}^{T}(x_s^{(b)} - \rho^{(b+1)} x_{s-1}^{(b)})^2}{2(1-\gamma^{(b)})},$$

To understand the posterior distribution, consider that beliefs about μ can be updated from the following regressions:

$$y_t - x_t = \mu + e_t$$
$$\frac{1}{\sqrt{1-\gamma^{(b)}}} x_t^{(b)} = \rho^{(b+1)} \frac{1}{\sqrt{1-\gamma^{(b)}}} x_{t-1}^{(b)} + \sigma^{(b)} \eta_t$$

where the matrix form of the first equation is

$$\begin{pmatrix} y_1 - x_1 \\ \vdots \\ y_T - x_T \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \mu + \begin{pmatrix} e_1 \\ \vdots \\ e_T \end{pmatrix}, \begin{pmatrix} e_1 \\ \vdots \\ e_T \end{pmatrix} \sim N(0, \sigma^2 \underbrace{\begin{bmatrix} \gamma & & & 0 \\ & \ddots & & \\ & & \gamma & \\ 0 & & & \gamma + \Gamma^2 \end{bmatrix}}_{\Lambda})$$

4. Obtain $x_{1:T}^{(b+1)} | \mu^{(b+1)}, \rho^{(b+1)}, \sigma^{(b+1)}, \gamma, y_{1:T}$, following the standard Kalman filter and simulation smoother.

C Forecasting with Diagnostic Filters

C.1 General Model Setup

Define linear Gaussian state-space model

$$y_t = Z\alpha_t + \epsilon_t, \epsilon_t \sim N(0, H)$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t, \eta_t \sim N(0, Q)$$

$$\alpha_1 \sim N(a_1, P_1), t = 1, ..., n.$$

where Y_t denote the set of past observations $y_1, ..., y_t$, and the other variables defined as

Expression	
$a_{t\|t}$	$E(\alpha_t\|Y_t)$
a_{t+1}	$E(\alpha_{t+1}\|Y_t)$
$P_{t\|t}$	$Var(\alpha_t\|Y_t)$
P_{t+1}	$Var(\alpha_{t+1}\|Y_t)$
v_t	$y_t - Z_t a_t$
F_t	$Var(v_t\|Y_{t-1})$

C.2 Representative Function and Distorted Posterior

Define the representativeness of a state α at t as the likelihood ratio:

$$R_t(\alpha) = \frac{f(\alpha|Y_t)}{f(\alpha|Y_{t-1} \cup \{\alpha_{t|t-1}\})}$$

State α is more representative at t if the signal y_t received in this period raises the probability of that state relative to the case where the news equals the ex-ante forecast, as described in the denominator.

The forecaster then over-weighs representative states by using the distorted posterior,

$$f^{\theta}(\alpha_t|Y_t) = f(\alpha_t|Y_t)R_t(\alpha_t)^{\theta}\frac{1}{J_t},$$

where J_t is a normalization factor and θ is a constant.

C.3 Distorted Posterior Distribution in the Steady State

The distorted posterior $f^{\theta}(\alpha_t|Y_t)$ is normal in the steady state characterized by a time-varying mean $a^{\theta}_{t|t}$ and constant covariance Σ, where

$$\begin{aligned} a^{\theta}_{t|t} &= a_{t|t} + \theta(a_{t|t} - a_t) \\ \Sigma &= \bar{P} - \bar{P}Z'\bar{F}^{-1}Z\bar{P} \end{aligned}$$

and $\bar{P}$ and $\bar{F}$ are the steady state solutions of P_{t+1} and F_t.

C.4 Proof of the Distorted Posterior Distribution in the Steady State

For $y_t = Za_t$, we have $a_{t|t} = a_t$ as we receive no news this case, i.e., $v_t = 0$. So $f(\alpha_t|Y_{t-1} \cup \{\alpha_{t|t-1}\}) \sim N(a_t, \Sigma)$.

As we now know the distributions of $f(\alpha_t|Y_{t-1} \cup \{\alpha_{t|t-1}\})$ and $f(\alpha_t|Y_t)$, the log likelihood of $f^{\theta}(\alpha_t|Y_t)$ is the following

$$\begin{aligned} lnf^{\theta}(\alpha_t|Y_t) &\propto -\frac{1}{2}(\alpha_t - a_{t|t})'\Sigma^{-1}(\alpha_t - a_{t|t}) - \frac{1}{2}\theta(\alpha_t - a_{t|t})'\Sigma^{-1}(\alpha_t - a_{t|t}) + \frac{1}{2}\theta(\alpha_t - a_t)'\Sigma^{-1}(\alpha_t - a_t) \\ &\propto -\frac{1}{2}\alpha_t'\Sigma^{-1}\alpha_t + \alpha_t'\Sigma^{-1}a_{t|t} - \frac{1}{2}\theta\alpha_t'\Sigma^{-1}\alpha_t + \theta\alpha_t'\Sigma^{-1}a_{t|t} + \frac{1}{2}\theta\alpha_t'\Sigma^{-1}\alpha_t - \theta\alpha_t'\Sigma^{-1}a_t \\ &\propto -\frac{1}{2}\alpha_t'\Sigma^{-1}\alpha_t + \alpha_t'\Sigma^{-1}(a_{t|t} + \theta(a_{t|t} - a_t)) \end{aligned}$$

By completing the square, we see that $f^{\theta}(\alpha_t|Y_t)$ is normal with mean $a_{t|t} + \theta(a_{t|t} - a_t)$ and covariance Σ.

C.5 Implementation of the Diagnostic Filter

To forecast with diagnostic filters, we include an additional step in the previous Gibbs sampling procedure. We additionally draw $\alpha_{diag,t|t} \sim N(a_{t|t} + \theta(a_{t|t} - a_t), \Sigma)$, which is later used in the forecasting step and θ

controls the degree of diagnosticity.

D Estimating Prior Beliefs of Forecasters

D.1 Forecasting with the Posterior Parameters and States

The algorithm described in the above section yields B samples of the posterior of the states and parameters of our model at each point in time t. We remove the first J samples as burn-in and re-index these samples by b as follows $\{\mu^{(b)}, \rho^{(b)}, \sigma^{(b)}, x_{1:t}^{(b)}\}_{b=1}^{B}$. We then use the following algorithm to produce a real-time forecast distribution for the housing growth rate at time t:

1. For each $b = 1, ..., B$, simulate a path of shocks $\{\eta_{t+k}^{(b)}\}_{k=1}^{K}$ from the standard normal distribution.

2. Starting from $k = 1$, construct a simulated path of the states over K subsequent periods using the following equations:
$$x_{t+k}^{(b)} = \rho^{(b)} x_{t+k-1|t}^{(b)} + \sqrt{1-\gamma^{(b)}}\sigma^{(b)}\omega_{t+k}^{(b)}$$

3. Use the simulated states to construct the forecast distribution of the quarterly rate $\{y_{t+k|t}^{(b)}\}$ where
$$y_{t+k|t}^{(b)} = \mu_{t|t}^{(b)} + x_{t+k|t}^{(b)}$$

4. The quarterly forecast of y_{t+k} given time t information is computed as
$$E_t y_{t+k} = \frac{1}{B}\Sigma_{b=1}^{B} y_{t+k|t}^{(b)}$$

5. The quarterly forecasts are then aggregated to annual forecasts by considering both the simulated forecasts and realized current year's quarterly rate.

D.2 Searching over Prior Beliefs

Let $\delta = (\mu_\mu, \sigma_\mu, \mu_\rho, \sigma_\mu, \lambda, \Gamma)'$ be the set of parameters that we wish to optimize. The moment function is defined in (13). The parameters are then estimated by minimizing the sum of squared moments as in (12). Every evaluation of the moment function $m(\delta)$ requires us to add noise in the last observation and sample from the posterior of the model 100 times which corresponds to a simulated panel of 100 individuals. We use a burn-in sample of 2000 draws and keep the subsequent 8000 draws. The global minimum is found using the Bayesian optimization routine from the Python's scikit-optimize package.

E Comparison with Household Surveys

In this section, we compare the Pulsenomics Home Price Expectations Survey of professionals with various household expectation surveys.

E.1 Household Survey Data

We study three sets of household survey data: the University of Michigan Survey of Consumers, the New York Fed's Survey of Consumer Expectations (SCE) Regular Module, and the SCE Housing Module. The Michigan survey measures (local) HPA forecasts over the next year and over the next 5 years (annualized), and is monthly starting in 2007. The SCE Regular Module survey data is also monthly and measures (national) HPA forecasts over the next year and over the one year two years hence. The SCE Housing Module data is available only at annual frequency and measures (local) HPA forecasts over the next year and over the next 5 years (annualized). The latter is used by Armona, Fuster and Zafar (2019).

We align the sample as much as possible with our Pulsenomics sample frame from 2010 until 2023. We are able to align the Michigan Survey perfectly. The SCE regular module runs from 2013 until 2023, while the SCE housing module runs from 2014 until 2022.

Table A1 shows the correlations between the time series of average expectations from the various surveys that are available at monthly or quarterly frequency. Given that the specific questions in the surveys and the timing of these surveys differ, we have selected and constructed the most comparable series of forecasts for calculating the correlations, details of which are described in the table notes. The results indicate that short-term household expectations are highly correlated with professional forecasts, with correlations in Panel A of 0.83 and 0.82 for the Michigan and NY Fed SCE surveys, respectively. The long-term forecasts from both series also show positive, albeit lower, correlations (Panel B).

Table A1: Correlations Between Pulsenomics and Household Surveys

	Pulsenomics Survey	**Michigan Survey**	**NY Fed - SCE Core**
Pulsenomics Survey	1.00	0.82	0.83
Michigan Survey	0.82	1.00	0.84
NY Fed - SCE Core	0.83	0.84	1.00
Panel B: Correlations of Long-term Forecasts			
	Pulsenomics Survey	**Michigan Survey**	**NY Fed - SCE Core**
Pulsenomics Survey	1.00	0.45	0.50
Michigan Survey	0.45	1.00	-0.12
NY Fed - SCE Core	0.50	-0.12	1.00

Notes: This table shows the correlations between the Pulsenomics Home Price Expectations Survey, the University of Michigan Surveys of Consumers, and the New York Fed's Survey of Consumer Expectations Core Module. Panel A shows the correlations between short-term consensus forecasts, defined as follows: (1) Pulsenomics: forecasts of nationwide HPA for the next calendar year, (2) Michigan: forecasts of local HPA over the next year, and (3) SCE Core: forecasts of nationwide HPA over the next year. Panel B shows the correlations between long-term consensus forecasts, defined as follows: (1) Pulsenomics: forecasts of average annual nationwide HPA over the next four calendar years, (2) Michigan: forecasts of average annual local HPA over the next five years, and (3) SCE Core: forecasts of 3-year-ahead annual nationwide HPA. Forecasts from the Pulsenomics Survey (quarterly) are aligned with the Michigan and SCE Core household surveys (monthly) for the months of February, May, August, and November, based on the timing of when the surveys are conducted. The sample period is 2010-2023 for the Pulsenomics and Michigan surveys and 2013-2023 for the SCE Core survey.

E.2 Extrapolation and Anomaly Moments

We begin by estimating the extrapolation regressions in household survey data and comparing them to the professional survey data evidence. Table A2 shows regressions of forecast errors on lagged HPA. One set of regressions has the 1-year lagged HPA on the right-hand side, while a second set has the annualized long-run lagged HPA, measured over the past 5 year on the right-hand side.

Overall, the results for households and professionals are quite similar. First, on the long-run mean reversion, the results are unanimous. In all three household surveys and in the professional survey, forecasters under-extrapolate. The slope of a regression of 2-5 year ahead or 3-year ahead HPA forecast errors on the past 5-year annual HPA results in large, positive, and significant coefficient estimates.

Second, for the regressions of short-run forecast errors on one-year lagged HPA, we find mild positive coefficients for the Michigan, SCE Housing Module, and Pulsenomics surveys, and a negative but insignificant coefficient for the SCE Regular Module. Hence, these results are also broadly consistent with short-run under-extrapolation (underestimating short-run momentum).

Third, the dependence of long-run forecasts on short-lagged HPA differs across household surveys. The Michigan survey, which matches the time period of the Pulsenomics survey, shows under-extrapolation vis-a-vis recent HPA (a positive coefficient), just like the Pulsenomics survey. The SCE modules show over-extrapolation of long-run forecasts to recent HPA (a negative coefficient). However, the SCE sample only starts in 2013 (Regular Module) or 2014 (Housing Module), which misses the important 2010-2012 period when HPA was weak coming out of the Great Financial Crisis. The housing market experienced a long expansion from 2013 until 2020. Ignoring the years of weak lagged HPA could lead to a higher estimate of the households' belief extrapolation coefficient in the SCE sample and hence a negative coefficient in the regression of forecast errors on lagged HPA. Consistent with this hypothesis, restricting the sample for the Michigan survey to 2013-2023 lowers the coefficient of the 2-5-year ahead forecast on the 1-year lagged HPA. The same is true in the Pulsenomics survey.

Table A2: Extrapolation in Household and Professional Surveys

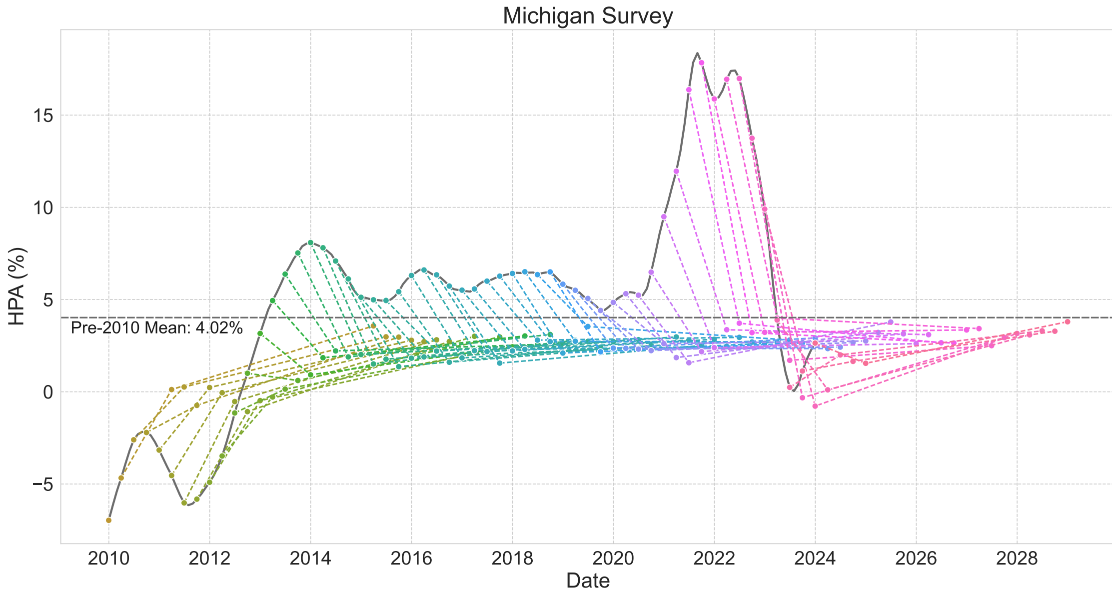
	Michigan Survey (2010-2023)				**Pulsenomics Survey (2010-2023)**					
	Next 1 Yr		**Next 2-5 Yrs**		**1-Yr Ahead**		**3-Yrs Ahead**		**Next 2-4 Yrs**	
Past 1Y HPA	0.317***		0.222***		0.081		0.143*		0.187***	
	(0.071)		(0.032)		(0.106)		(0.085)		(0.057)	
Past 5Y Ann. HPA		0.283***		0.349***		0.165		0.349***		0.406***
		(0.078)		(0.024)		(0.111)		(0.115)		(0.053)
Num. of Obs.	61,784	61,784	40,070	40,070	5,503	5,503	4,737	4,737	4,352	4,352
R-Squared	0.076	0.044	0.049	0.102	0.010	0.029	0.026	0.152	0.095	0.411

	SCE Core (2013-2023)				**SCE Housing Module (2014-2022)**			
	Next 1 Yr		**3-Years Ahead**		**Next 1 Yr**		**Next 2-5 Yrs**	
Past 1Y HPA	-0.093		-1.099***		0.256		-0.809	
	(0.123)		(0.387)		(0.849)		(0.728)	
Past 5Y Ann. HPA		0.175*		0.654***		0.502		0.559***
		(0.097)		(0.119)		(0.699)		(0.069)
Num. of Obs.	146,265	146,265	101,637	101,637	7,808	7,808	4,501	4,501
R-Squared	0.001	0.005	0.014	0.055	0.004	0.014	0.052	0.227

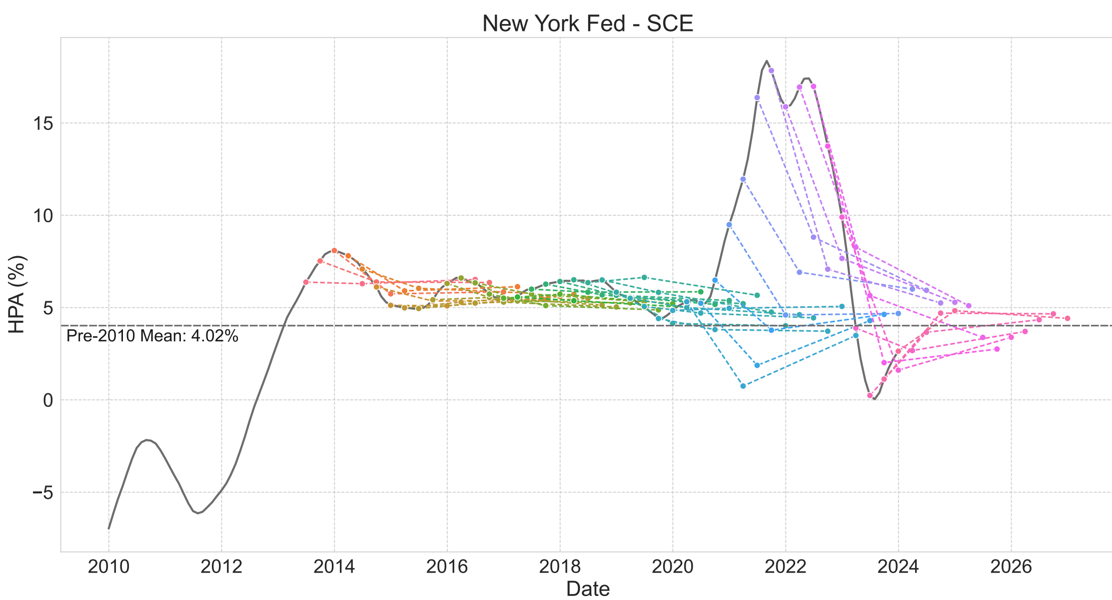
Notes: This table shows the relationship between forecast errors and past HPA across different surveys. The coefficients are estimated from individual-level extrapolation regressions: $e_{i,t+h|t} = a^{ex} + b^{ex} y_{t,t-j} + \varepsilon_{i,t+h}$. All past HPA variables are annualized. Standard errors are clustered at the date and forecaster level. Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Similar to Figure 1, Figure A1 plots the consensus HPA forecasts from the two household surveys alongside the realized annual nationwide HPA, which is based on the Zillow Home Value Index (ZHVI). Both household surveys show a similar pattern of "pulling back towards the long-run mean": long-term forecasts are close to the unconditional mean.

Figure A1: Consensus HPA Forecast by Households



(a) University of Michigan – Surveys of Consumers



(b) New York Fed – Survey of Consumer Expectations

Notes: This figure shows the time series of consensus forecasts for the University of Michigan Surveys of Consumers and the New York Fed's Survey of Consumer Expectations Core Module. The black solid lines represent the realized annual nationwide HPA based on ZHVI. In Panel A, each short colored dashed line with three dots plots the actual realized annual HPA (the first circle), the average forecast of annual HPA over the next year (the second circle), and the average forecast of annual HPA 2-5 year ahead (the third circle). In Panel B, each short colored dashed line with three dots plots the actual annual HPA (the first circle), the average forecast of annual HPA over the next year (the second circle), and the average forecast of annual HPA 3-year ahead (the third circle). The survey forecasts from the last month of each quarter are plotted in the figure.
Alt text: Graphs showing the realized annual house price growth (solid black line) alongside the average survey forecasts for house price growth over different horizons (colored dashed lines) from different data sources. Each colored line segment represents forecasts for the current, the next and the next 2-5 year ahead in Panel (a) and the 3-year ahead in Panel (b).

We also study the other anomaly regressions that we are able to estimate on household data. Those include the individual inverse-MZ regressions and the individual average forecast bias. We do not have individual household data on forecast revisions, so we cannot estimate the CG regressions on household survey data. As table A3 shows, these moments also line up nicely between household and professional survey data.

Table A3: Forecast Anomalies in Household and Professional Surveys

Panel A. Individual Inverse MZ Regressions									
	Michigan Survey (2010-2023)		**SCE Core (2013-2023)**		**SCE Housing Module (2014-2022)**		**Pulsenomics Survey (2010-2023)**		
	Next 1 Yr	**Next 2-5 Yrs**	**Next 1 Yr**	**3-Years Ahead**	**Next 1 Yr**	**Next 2-5 Yrs**	**1-Yr Ahead**	**3-Yrs Ahead**	**Next 2-4 Yrs**
b^{MZ}	0.140***	-0.024***	-0.001***	-0.030***	-1.338***	-0.078***	0.130***	-0.051***	-0.084***
	(0.016)	(0.018)	(0.061)	(0.018)	(0.207)	(0.042)	(0.050)	(0.013)	(0.029)
Num. of Obs.	61,784	40,070	146,265	101,637	7,808	4,501	5,503	4,737	4,352
R-squared	0.023	0.000	0.000	0.000	0.155	0.002	0.039	0.009	0.012
Panel B. Biases									
	Michigan Survey (2010-2023)		**SCE Core (2013-2023)**		**SCE Housing Module (2014-2022)**		**Pulsenomics Survey (2010-2023)**		
	Next 1 Yr	**Next 2-5 Yrs**	**Next 1 Yr**	**3-Years Ahead**	**Next 1 Yr**	**Next 2-5 Yrs**	**1-Yr Ahead**	**3-Yrs Ahead**	**Next 2-4 Yrs**
a^{fb}	4.110***	4.304***	2.038***	2.743***	4.911	5.380***	2.952***	4.299***	4.292***
	(0.395)	(0.196)	(0.425)	(0.525)	(3.468)	(0.989)	(0.584)	(0.556)	(0.385)
Num. of Obs.	61,784	40,070	146,265	101,637	7,808	4,501	5,503	4,737	4,352

Notes: This table shows the regression estimates of other forecast anomalies at different horizons across different surveys. Panel A shows the estimates of inverse MZ regression specified in Equation 3. Stars indicate the level of statistical significance in testing the hypothesis: $b = 1$. Panel B shows the estimates of bias in Equation 2. All coefficients are estimated from individual-level panel regressions. Standard errors are in parentheses. Standard errors are clustered at the date and forecaster level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

In summary, the evidence on extrapolation and anomaly moments is broadly similar between household and professional forecasters over our sample period.

E.3 Comparison With AFZ (2019)

In addition to their different focus on household forecasts, Armona, Fuster and Zafar (2019, AFZ for short) emphasize the perception gap in shaping expectations. The perception gap measures the difference between actual past HPA and the perceived past HPA. Our forecasters are all shown the true past HPA on the survey data entry screen, so that there is no room for a perception gap in our setting. In their Table 5, AFZ find that households adjust their short-term and long-term forecasts of future HPA positively to *their own perceptions of* short- and long-lagged HPA, with coefficients that are stronger for short-run than for long-run forecasts. These regressions are based on one cross-section of data from the April 2015 SCE Housing Module. Hence, these results are not directly comparable to our *panel* regressions which have *actual* lagged HPA on the right-hand side. That said, the dependence of forecasts on lagged house prices (perceived or real) is a common finding across papers.

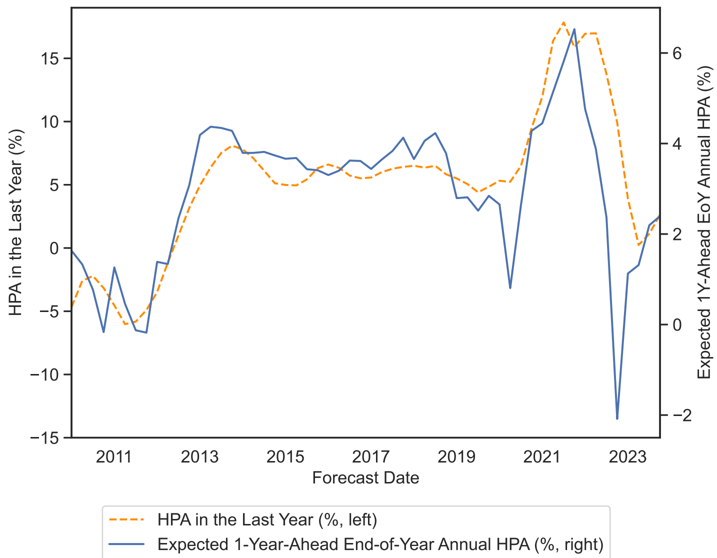
In addition, AFZ find that respondents perceive substantial uncertainty in their estimates of future home price growth (their section 4.2), consistent with our findings and the premise of our Bayesian learning model.

Finally, AFZ ascribe an important role to forecaster heterogeneity and argue for “type heterogeneity", rather than merely information heterogeneity. Similarly, we find that prior heterogeneity plays a more important role in generating dispersion in professional forecasters’ expectations than signal (information) heterogeneity.

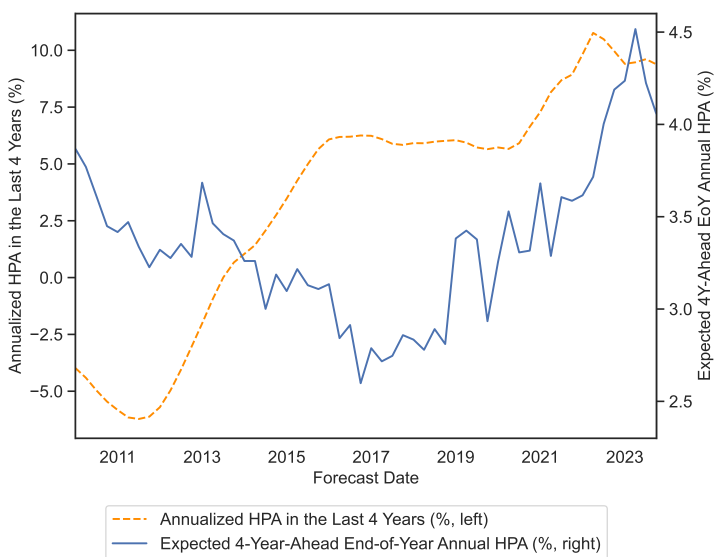
The consistency between our findings and AFZ strengthens the case for our Bayesian learning model.

F Additional Empirical Results

Figure A2: Time-Series of HPA Forecasts



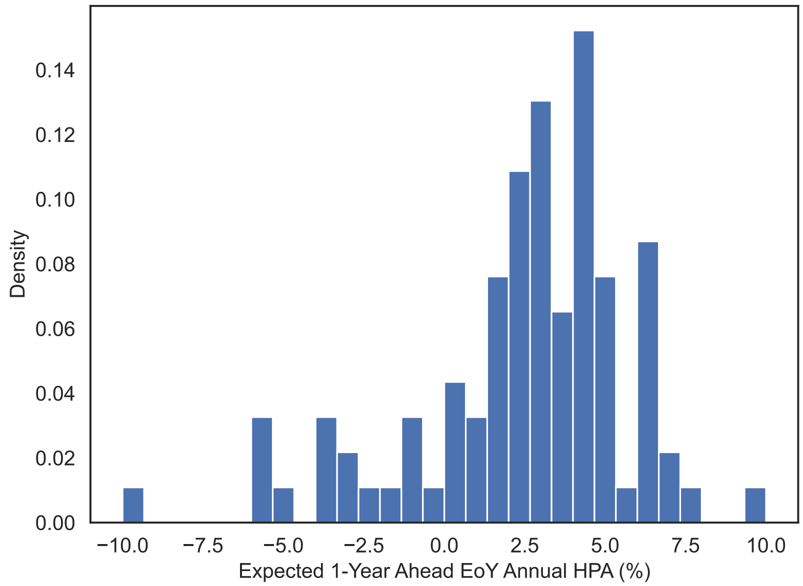
(a) 1-Year Ahead



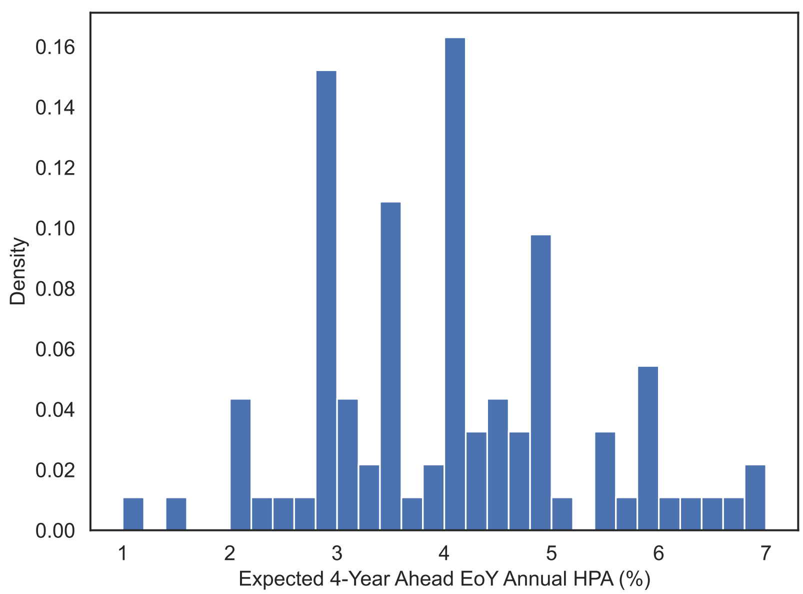
(b) 4-Year Ahead

Notes: This figure shows the time series of average professional forecasts of future annual HPA and realized annual HPA. Panel A shows the forecasts of 1-year-ahead end-of-year annual HPA (e.g., from Dec 31, 2010, to Dec 31, 2011, for the 2010Q1 survey; solid line, right y-axis) and the realized HPA in the last year (e.g., from Mar 31, 2009, to Mar 31, 2010, for the 2010Q1 survey; dashed line, left y-axis). Panel B shows the forecasts of 4-year-ahead end-of-year annual HPA (e.g., from Dec 31, 2013, to Dec 31, 2014, for the 2010Q1 survey; solid line, right y-axis) and the annualized realized HPA in the past 4 years (e.g., annualized HPA from Mar 31, 2006, to Mar 31, 2010, for the 2010Q1 survey; dashed line, left y-axis).
Alt text (decorative): Two-panel line graph comparing realized house price appreciation (HPA) with forecasts.

Figure A3: Cross-Sectional Dispersion of HPA Forecasts



(a) 1-Year Ahead



(b) 4-Year Ahead

Notes: This figure shows the cross-sectional dispersion of HPA forecasts in the 2023Q4 survey. Panel A shows the histogram of forecasts for 1-year-ahead end-of-year annual HPA (from Dec 31, 2023, to Dec 31, 2024). Panel B shows the histogram of forecasts for 4-year-ahead end-of-year annual HPA (from Dec 31, 2026, to Dec 31, 2027).
Alt text: Figure showing the cross-sectional dispersion of HPA forecasts in the 2023Q4 survey for 1-year-ahead forecasts (Panel A) and 4-year-ahead forecasts (Panel B).

Table A4: Forecast Anomalies with Individual Fixed Effects

	Forecast Horizon				
	Current Year	**1-Year Ahead**	**2-Years Ahead**	**3-Years Ahead**	**4-Years Ahead**
Panel A: Sensitivity to Past 1Q Growth					
b	0.492***	0.187***	0.047***	0.009	-0.007
	(0.050)	(0.032)	(0.015)	(0.011)	(0.011)
Num. of Obs.	5,847	5,847	5,847	5,847	5,847
R-squared	0.599	0.483	0.480	0.398	0.350
Panel B: Sensitivity to Past 2Y Growth					
b	0.444***	0.106**	0.029	0.018	0.009
	(0.067)	(0.053)	(0.018)	(0.013)	(0.014)
Num. of Obs.	5,847	5,847	5,847	5,847	5,847
R-squared	0.450	0.397	0.471	0.400	0.350
Panel C: Sensitivity to Past 4Y Growth					
b	0.412***	0.101**	0.021	0.011	0.007
	(0.079)	(0.047)	(0.020)	(0.017)	(0.016)
Num. of Obs.	5,847	5,847	5,847	5,847	5,847
R-squared	0.370	0.389	0.469	0.398	0.350
Panel D: Inverse MZ Regressions					
b	0.617***	0.097***	-0.063***	-0.046***	-0.056***
	(0.067)	(0.049)	(0.016)	(0.013)	(0.012)
Num. of Obs.	5,847	5,486	5,101	4,717	4,337
R-squared	0.588	0.393	0.517	0.460	0.446
Panel E: CG Regressions					
b	-0.346***	-0.342**	-0.549***	-0.540***	-0.419***
	(0.105)	(0.157)	(0.045)	(0.023)	(0.058)
Num. of Obs.	5,023	4,701	4,384	4,040	2,853
R-squared	0.236	0.239	0.258	0.213	0.198

Notes: This table shows the regression estimates of forecast anomalies at different horizons with individual fixed effects. Panel A to Panel C show the relationship between house price forecasts and past house price growth, estimated using equation 1. Panel D shows the estimates of inverse MZ regression specified in Equation 3. Stars indicate the level of statistical significance in testing the hypothesis: $b = 1$. Panel E shows the estimates of CG regression in equation 4. All coefficients are estimated from individual-level panel regressions. Standard errors are in parentheses. Standard errors are clustered at the quarter and forecaster level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A5: Demographic Characteristics and HPA Forecasts

	Forecasts				Forecast Errors				Absolute Forecast Errors			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Forecast Horizon (qtrs.)		-0.040	-0.040	-0.048*		0.183***	0.183***	0.182***		0.121**	0.121**	0.123**
		(0.028)	(0.028)	(0.027)		(0.050)	(0.050)	(0.051)		(0.048)	(0.048)	(0.049)
Length of Exp. (yrs.)												
<15		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)
15-25		-0.113	-0.039	-0.115		0.079	0.025	0.074		0.149	0.141	0.139
		(0.206)	(0.196)	(0.221)		(0.207)	(0.196)	(0.223)		(0.154)	(0.145)	(0.165)
25-35		0.138	0.160	0.223		-0.132	-0.135	-0.232		-0.031	-0.021	-0.114
		(0.203)	(0.209)	(0.215)		(0.199)	(0.204)	(0.212)		(0.141)	(0.142)	(0.148)
>35		0.621***	0.572**	0.613**		-0.621**	-0.573**	-0.618**		-0.343**	-0.335**	-0.339*
		(0.228)	(0.239)	(0.245)		(0.233)	(0.235)	(0.252)		(0.171)	(0.159)	(0.182)
Highest Degree												
Bachelor		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)
Master		0.502	0.342	0.579		-0.435	-0.309	-0.502		-0.325	-0.264	-0.405
		(0.317)	(0.262)	(0.359)		(0.289)	(0.252)	(0.329)		(0.228)	(0.197)	(0.258)
Doctor		0.103	-0.157	0.136		-0.053	0.164	-0.069		-0.097	0.044	-0.119
		(0.311)	(0.261)	(0.358)		(0.288)	(0.253)	(0.333)		(0.233)	(0.209)	(0.266)
Major												
Econ&Finance		0.179	0.163	0.221		-0.208	-0.204	-0.257		-0.146	-0.121	-0.170
		(0.226)	(0.242)	(0.233)		(0.212)	(0.223)	(0.221)		(0.180)	(0.166)	(0.184)
Skill Number												
<25% Percentile		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)
25% - 50% Percentile		0.091	-0.119	0.070		-0.054	0.162	-0.024		0.049	0.203	0.088
		(0.254)	(0.281)	(0.259)		(0.254)	(0.274)	(0.259)		(0.191)	(0.192)	(0.196)
50 - 75% Percentile		-0.090	-0.057	-0.128		0.164	0.172	0.220		0.108	0.168	0.169
		(0.193)	(0.211)	(0.213)		(0.198)	(0.208)	(0.219)		(0.152)	(0.146)	(0.160)
>75% Percentile		-0.227	-0.304	-0.183		0.343	0.415*	0.302		0.301*	0.365**	0.280*
		(0.207)	(0.208)	(0.215)		(0.208)	(0.211)	(0.216)		(0.162)	(0.163)	(0.166)
Number of Followers												
<25% Percentile		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)		(omitted)	(omitted)	(omitted)
25% - 50% Percentile		0.330	0.255	0.465*		-0.320	-0.247	-0.461*		-0.138	-0.058	-0.235
		(0.239)	(0.238)	(0.245)		(0.241)	(0.236)	(0.247)		(0.165)	(0.169)	(0.164)
50 - 75% Percentile		0.252	0.376*	0.298		-0.240	-0.370*	-0.272		-0.017	-0.098	-0.039
		(0.223)	(0.189)	(0.221)		(0.231)	(0.202)	(0.235)		(0.174)	(0.165)	(0.175)
>75% Percentile		0.292	0.248	0.354		-0.203	-0.167	-0.250		0.032	0.050	0.011
		(0.255)	(0.246)	(0.263)		(0.254)	(0.240)	(0.263)		(0.169)	(0.169)	(0.174)
Region (CBSA)			Y				Y				Y	
Lagged 1Y Local HPA				-0.000				0.003				0.007
				(0.019)				(0.020)				(0.014)
Mean of Dep. Var.	3.266	3.266	3.266	3.266	3.217	3.217	3.217	3.217	4.000	4.000	4.000	4.000
Std. of Dep. Var.	2.829	2.829	2.829	2.829	4.320	4.320	4.320	4.320	3.608	3.608	3.608	3.608
Quarter FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
R-Squared	0.140	0.182	0.201	0.183	0.209	0.276	0.283	0.273	0.170	0.216	0.222	0.217
Num. of Obs.	29,340	16,835	16,835	15,360	25,577	14,667	14,667	13,291	25,577	14,667	14,667	13,291

Notes: The region control is an indicator variable for being located in each of the largest Core-Based Statistical Areas (CBSA), for being in one of the remaining CBSAs or in a non-metropolitan area, and for being abroad. A separate CBSA indicator variable is included if there are at least 5 different forecasters in that region at some point. There are 11 separate CBSA categorical variables, including: 1) New York-Newark-Jersey City, 2) Los Angeles-Long Beach-Anaheim, 3) Chicago-Naperville-Elgin, 4) Washington-Arlington-Alexandria, 5) Philadelphia-Camden-Wilmington, 6) Miami-Fort Lauderdale-West Palm Beach, 7) Boston-Cambridge-Newton, 8) San Francisco-Oakland-Fremont, 9) Seattle-Tacoma-Bellevue, 10) San Diego-Chula Vista-Carlsbad, and 11) Charlotte-Concord-Gastonia. Standard errors are clustered at the quarter and forecaster level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

Table A6: Length of Experience and Forecast Dispersion

	Dispersion of Implied Long-Run Growth Rates
Length of Experience (yrs.)	
<15	(omitted)
15-25	-0.019*
	(0.007)
25-35	-0.047**
	(0.011)
>35	-0.052**
	(0.014)
Survey Quarter FE	Y
R-Squared	0.657
Number of Observations	220

Notes: This table shows the relationship between the dispersion of implied long-run house price growth rates and the length of forecasters' experience. For each group of forecasters with different lengths of experience (k), the cross-sectional variance of their implied long-run house price growth rates is calculated at each date t: $Disp_{k,t}^{LR}$. Groups with fewer than 5 forecasters are dropped. The coefficients in the table are estimated from panel regressions: $Disp_{k,t}^{LR} = a_t + \sum_{k=2}^{4} b_k LoE_k + \varepsilon_{k,t}$, where a_t represents the quarter fixed effects. Standard errors are in parentheses. Standard errors are clustered at the quarter and LoE category level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

F.1 MZ vs. Inverse MZ Regression Estimation

The standard Mincer and Zarnowitz (1969) regression takes the form: $y_{t+h} = \widetilde{a}^{mz} + \widetilde{b}^{mz} E_t^c[y_{t+h}] + \widetilde{\varepsilon}_{i,t+h}$. In this regression, full-information rational expectations (FIRE) imply that $\widetilde{b}^{mz} = 1$, which means that realized values should move one-for-one with forecasts. However, in an individual-level panel regression, if FIRE does not hold, the slope coefficient will be underestimated. The dependent variable y_{t+h} is identical for every individual in a single cross-section, causing the error terms to be correlated with the regressors. In our model, the magnitude of this downward bias is influenced by (i) the distribution of forecasters' prior mean about the long-run mean HPA, captured by the parameter λ, and (ii) the cross-sectional distribution of private signals, determined by $\bar{\Gamma}$ and $\bar{\sigma}^2$. Thus, the model parameters to be estimated affect the level of estimation bias of the model-implied moments, complicating the estimation of parameters. Adding individual fixed effects cannot eliminate the bias, given that private signals are i.i.d. across time and cannot be absorbed by individual fixed effects.

To address this issue, we instead include the intercept and slope coefficients of the *inverse* MZ regressions, as described in equation (3), as moments to be matched in the estimation of our model. The inverse MZ regression has the realized HPA as the independent variable and HPA forecasts as the dependent variable, thus avoiding the aforementioned bias in the panel regression setting. This regression may suffer from attenuation bias due to noise in the realized HPA. However, the magnitude of this bias depends solely on the true DGP of HPA, specifically on the proportion of the variance of the noise term e_t in the total variance of HPA y_t. Thus, the level of bias in the model-implied moments is not influenced by the model's parameters. Both the empirical moments and the model-implied moments are subject to the same degree of attenuation bias, making the estimation process cleaner. Furthermore, in a balanced panel, both individual and consensus *inverse* MZ regressions yield identical coefficients, a desirable feature not observed for the MZ regressions.

Finally, Table A7 reports the slope estimates from the MZ regression estimated using consensus forecasts. The consensus MZ regression is free from any of the aforementioned biases. FIRE implies $\widetilde{b}^{mz} = 1$. We find that all slope estimates are negative and of large magnitudes for $h = 2, 3, 4$, indicating significant deviations from FIRE. This evidence is consistent with the evidence from the inverse MZ regression results we presented in Table 2 Panel B.

Table A7: Consensus MZ Regression Results

	Forecast Horizon				
	Current Year	**1-Year Ahead**	**2-Years Ahead**	**3-Years Ahead**	**4-Years Ahead**
$\widetilde{b}^{mz}$	1.129	0.985	-2.773**	-6.235***	-7.759***
	(0.154)	(0.548)	(1.536)	(2.597)	(2.300)
Num. of Obs.	56	52	48	44	40
R-squared	0.705	0.123	0.184	0.311	0.445

Notes: This table shows the estimates of consensus-level MZ regressions: $y_{t+h} = \widetilde{a}^{mz} + \widetilde{b}^{mz} E_t^c[y_{t+h}] + \widetilde{\varepsilon}_{t+h}$, across different horizons. Stars indicate the level of statistical significance in testing the hypothesis: $\widetilde{b}^{mz} = 1$. Standard errors are shown in parentheses and are calculated using Newey-West with a lag length of $L = \lceil 0.75 \times T^{1/3} \rceil$. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

G Forecasts of Others

The model assumes that the forecasters do not use the lagged forecasts of other forecasters to predict future house price growth.

We investigate empirically whether this is a good assumption. We ask whether the lagged consensus (i.e., average) forecast adds explanatory power to the individual forecasts. We control for the own lagged forecast and lagged realized HPA. The results are in Table A8 below. In Panel A, we find that the lagged consensus forecast does not enter significantly, either for the short-run or for the long-run forecasts. We find the same result for the lagged forecast of the best ten forecasters (lowest RMSE) rather than the consensus forecast in Panel B. This result demonstrates the observed behavior that forecasters ignore the forecasts of others.

A more subtle question is how rational this behavior is. Forecasters in our model receive private signals about the current quarter T's HPA, y_T. All forecasters see the same lagged history of HPA, $y_{T-j},\ j = 1, \cdots, 5$, when they fill out the survey. They cannot observe the contemporaneous forecasts of other forecasters because the survey results have not been published at the time they make their forecasts. The lagged consensus forecast cannot add information about lagged HPA. It theoretically could add information about the current and future HPA rates if it helps inference about the parameters of the model, notably the prior mean of the long-run growth rate of house prices, μ_μ. However, the prior μ_μ is not necessarily close to the true long-run mean, $\bar{\mu}$. We have shown (in Appendix H) that learning about the long-run mean is slow, potentially taking one hundred years. This means that even if the consensus forecast fully revealed the μ_μ, that would still not be very helpful for learning $\bar{\mu}$.

Finally, our baseline model makes the simplifying assumption that all forecasters start making forecasts in the same year (1996). In reality—and in the model extension discussed in section 5 for the purposes of creating table 7—forecasters are born, start their careers, and begin forecasting in different years. This heterogeneity makes the consensus forecast an extremely complex object to unravel and to use for improved inference on the DGP, both in the data and in the extended model.

Table A8: Effect of Lagged Consensus Forecasts on Individual Forecasts

	Panel A: Lagged Consensus Forecasts				
	Current Year	**1-Year Ahead**	**2-Years Ahead**	**3-Years Ahead**	**4-Years Ahead**
Consensus Forecasts (t-1)	0.077	0.225	0.073	0.050	0.057
	(0.113)	(0.163)	(0.062)	(0.051)	(0.049)
Forecasts (t-1)	0.555***	0.606***	0.604***	0.593***	0.490***
	(0.055)	(0.042)	(0.048)	(0.047)	(0.037)
Forecast (t-2)	-0.013	0.228***	0.259***	0.243***	0.235***
	(0.051)	(0.050)	(0.039)	(0.032)	(0.021)
Past HPA Realizations	Y	Y	Y	Y	Y
Num. of Obs.	4,469	4,469	4,469	4,469	2,308
R-Squared	0.751	0.622	0.622	0.540	0.451

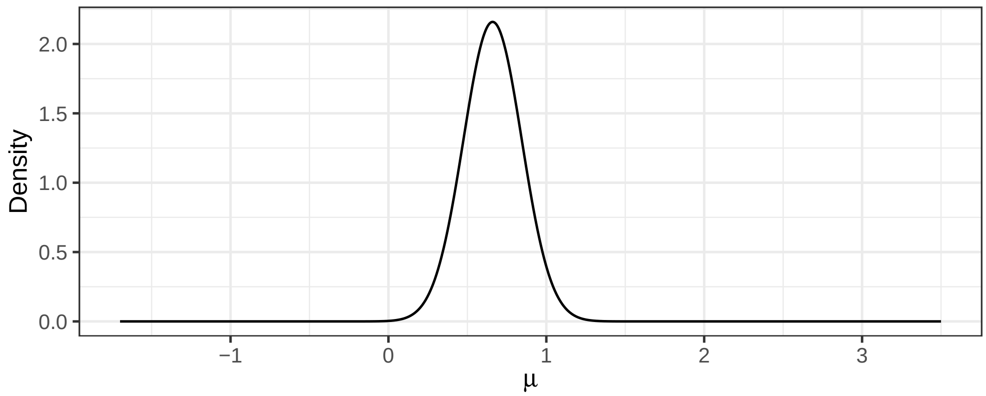
	Panel B: Average Forecasts from the Best Forecasters				
	Current Year	**1-Year Ahead**	**2-Years Ahead**	**3-Years Ahead**	**4-Years Ahead**
Average Forecasts from	-0.013	0.164	-0.082	-0.083	0.028
the Best Forecasters (t-1)	(0.093)	(0.102)	(0.058)	(0.052)	(0.054)
Forecasts (t-1)	0.571***	0.587***	0.582***	0.572***	0.472***
	(0.060)	(0.045)	(0.053)	(0.055)	(0.028)
Forecast (t-2)	-0.021	0.238***	0.288***	0.250***	0.257***
	(0.052)	(0.054)	(0.042)	(0.040)	(0.019)
Past HPA Realizations	Y	Y	Y	Y	Y
Num. of Obs.	4,126	3,779	3,426	3,101	1,462
R-Squared	0.725	0.580	0.598	0.507	0.448

Notes: Panel A shows the relationship between individual forecasts and lagged consensus forecasts. Panel B shows the relationship between individual forecasts and the lagged average forecasts from the 10 best forecasters. For each period, the 10 best forecasters are selected based on having the lowest root mean square error (RMSE) for the same-horizon forecasts made in all surveys from the most recent year with available RMSE data. For example, to identify the 10 best forecasters for 4-year-ahead forecasts in 2015Q1, performance is evaluated using all 4-year-ahead forecasts made in 2010. These forecasters must have participated in all four surveys conducted in that year. Past HPA realizations include lagged quarterly HPA from 1 to 6 quarters. The coefficients are estimated from individual-level panel regressions. All past HPA variables are annualized. Standard errors are clustered at the quarter and forecaster level. Standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

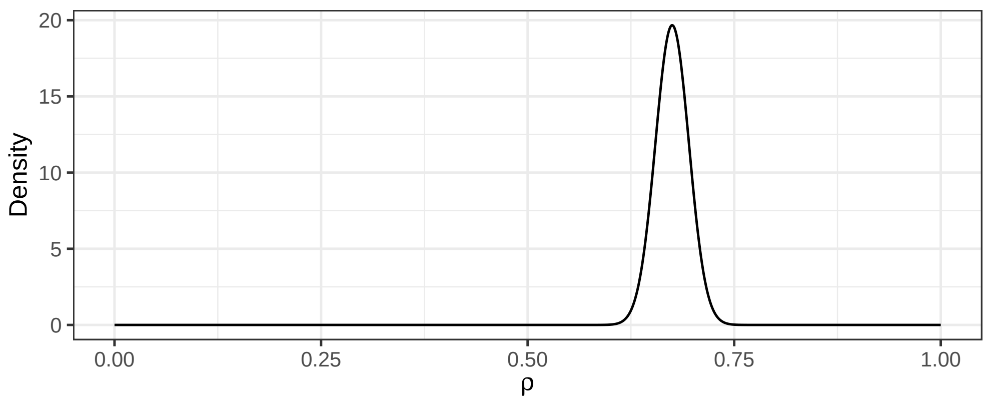
H Additional Results from the Model

Figure A4: Prior Distributions of Parameters in the Benchmark Model

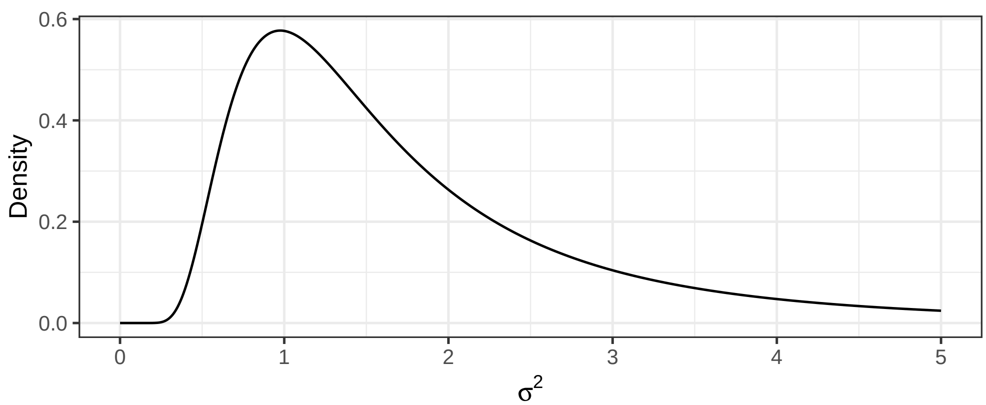
Alt text (decorative): Figure showing the estimated prior distributions of key parameters of the Benchmark Model proposed in the paper.



(a) Prior distribution of the long-run mean (μ)



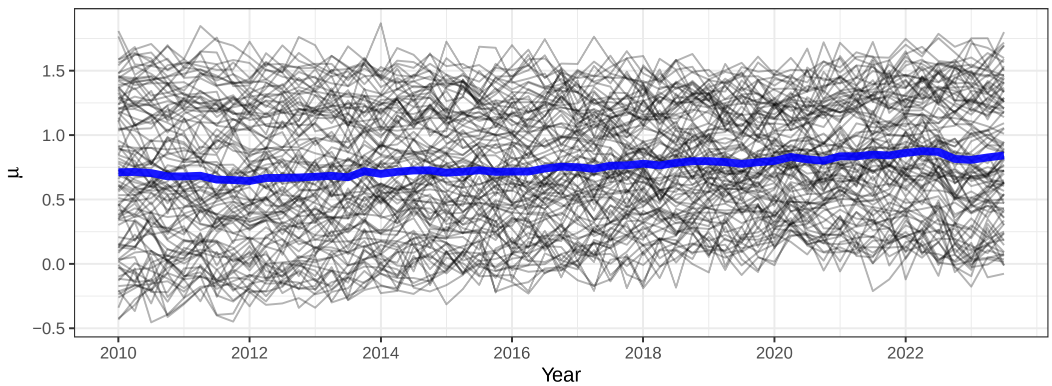
(b) Prior distribution of the persistent parameter (ρ)



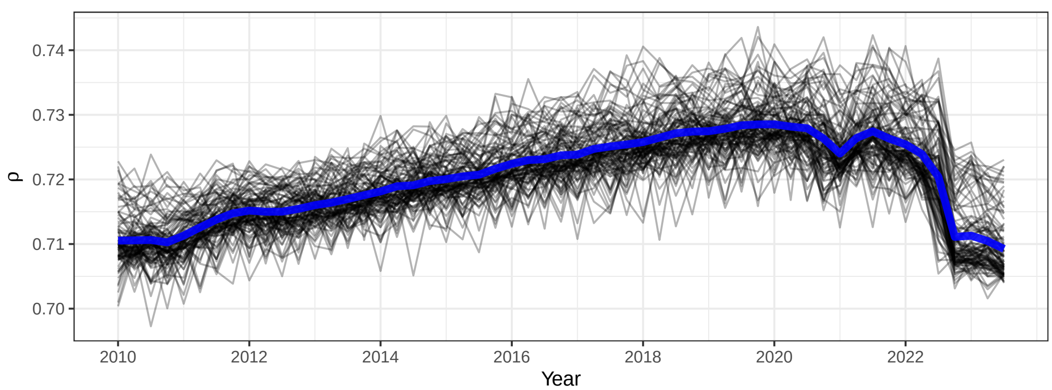
(c) Prior distribution of the volatility parameter (σ^2)

Figure A5: Posterior Distributions of Parameters in the Benchmark Model

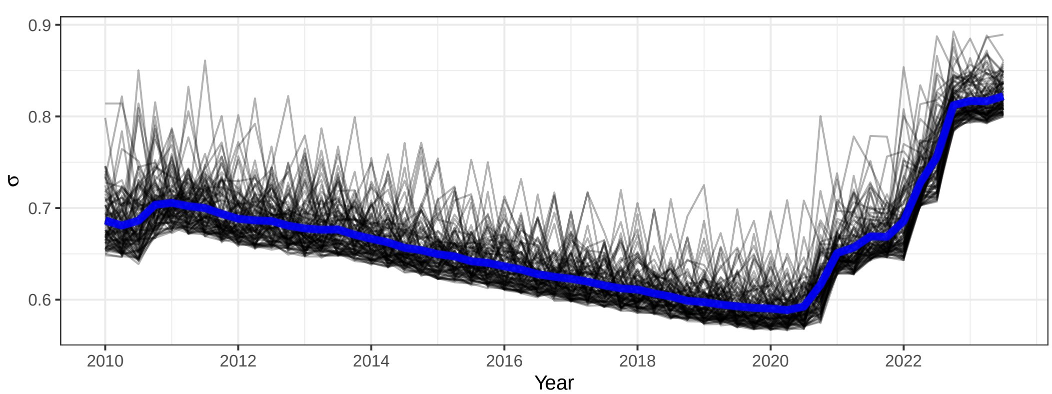
Alt text (decorative): Figure showing the estimated posterior distributions of key parameters of the Benchmark Model proposed in the paper.



(a) Posterior distribution of the long-run mean (μ)

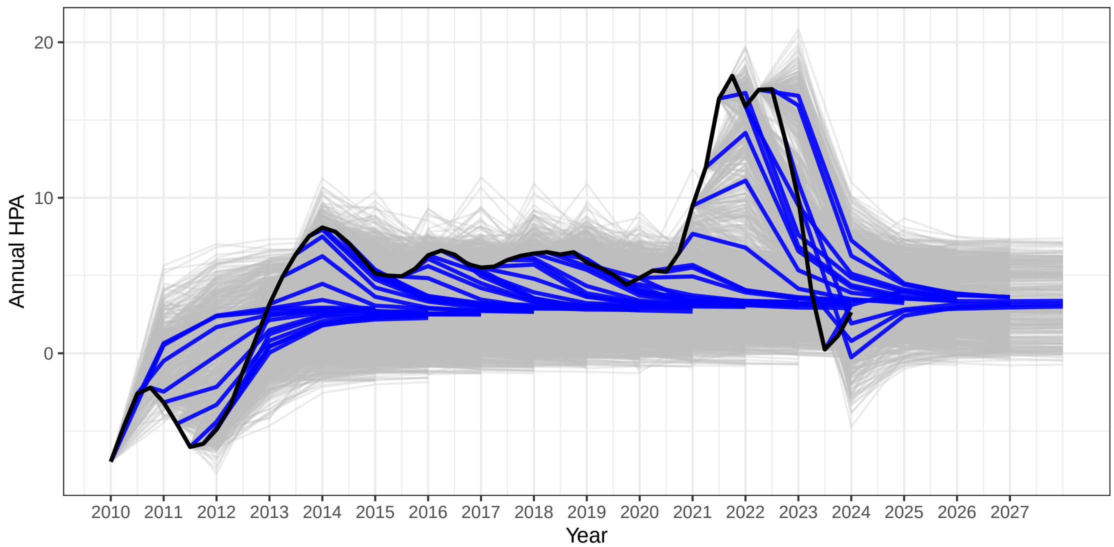


(b) Posterior distribution of the persistent parameter (ρ)



(c) Posterior distribution of the volatility parameter (σ^2)

Figure A6: Model Implied Forecasts: Individuals



Notes: This figure shows individual forecasts in the simulation of the benchmark rational learning model. The grey areas represent the range of forecasts across the 100 forecasters in the model simulation. The blue solid line represents the consensus forecasts.
Alt text (decorative): Figure showing individual forecasts in the simulation of the benchmark rational learning model.

Table A9: Comparison of Residual Standard Errors from Individual Forecaster Regressions

Panel A: CG Regressions								
Moments	**Horizon**	**Data**	**Benchmark**	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
		(1)	(2)	(3)	(4)	(5)	(6)
CG	**1-Yr**	4.649	4.235	4.313	4.276	4.534	4.102
	2-Yr	4.163	3.910	3.935	3.920	4.708	3.519
	3-Yr	3.960	3.872	3.895	3.864	4.723	3.386
	4-Yr	4.026	4.013	4.028	3.996	4.902	3.514
Cost Function			**0.24**	**0.17**	**0.21**	**1.66**	**1.31**

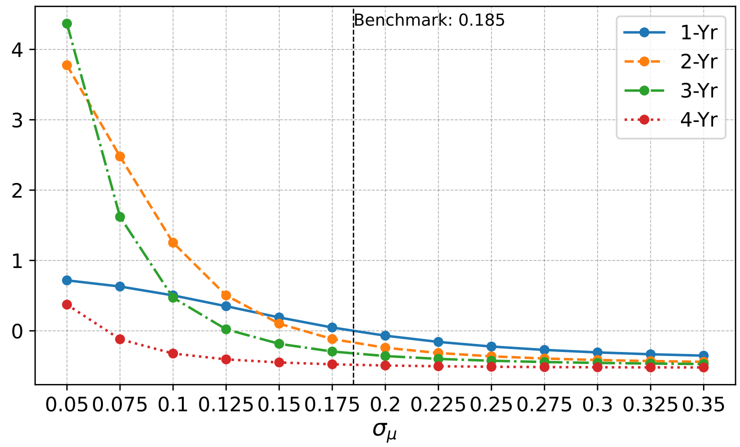
Panel B: Sensitivity to Past House Price Growth							
Moments	**Horizon**	**Data**	**Benchmark**	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Sensitivity to	**1-Yr**	2.765	1.851	2.732	2.351	2.663	0.490
Past 1Q	**2-Yr**	2.257	1.869	1.966	1.896	2.922	0.176
Growth	**3-Yr**	1.905	1.959	1.997	1.956	3.283	0.061
	4-Yr	1.764	2.014	2.048	2.008	3.502	0.040
Sensitivity to	**1-Yr**	2.927	2.079	2.923	2.558	3.082	1.251
Past 2Y	**2-Yr**	2.271	1.882	1.980	1.911	2.987	0.383
Growth	**3-Yr**	1.904	1.958	1.996	1.955	3.287	0.118
	4-Yr	1.764	2.012	2.046	2.006	3.502	0.054
Sensitivity to	**1-Yr**	2.943	2.089	2.913	2.553	3.082	1.354
Past 4Y	**2-Yr**	2.275	1.881	1.979	1.908	2.992	0.435
Growth	**3-Yr**	1.905	1.954	1.993	1.952	3.287	0.132
	4-Yr	1.764	2.010	2.044	2.004	3.502	0.059
Cost Function			**2.93**	**0.52**	**1.04**	**16.31**	**40.32**

Panel C: Biases and Inverse MZ Regressions							
Moments	**Horizon**	**Data**	**Benchmark**	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Forecast	**1-Yr**	4.841	4.411	4.624	4.503	4.688	4.256
Biases (%)	**2-Yr**	4.241	3.925	3.962	3.940	4.709	3.571
	3-Yr	3.999	3.854	3.883	3.850	4.706	3.362
	4-Yr	4.109	4.022	4.037	4.003	4.933	3.515
Inverse MZ	**1-Yr**	2.884	2.267	2.976	2.659	3.290	1.778
	2-Yr	2.159	1.936	2.037	1.967	3.096	0.615
	3-Yr	1.820	1.966	2.006	1.964	3.299	0.177
	4-Yr	1.704	2.016	2.050	2.010	3.504	0.072
Cost Function			**0.87**	**0.32**	**0.44**	**7.90**	**10.52**

Panel D: Total Weighted Cost Function					
	Benchmark	**Overconfidence ($\phi = 2.5$)**	**Diagnostic Exp. ($\theta = 0.8$)**	**No LR.**	**No Disp.**
Total Cost Function	4.04	1.01	1.69	25.87	52.15

Notes: This table compares residual standard errors (RSE) from individual forecaster regressions across different models. Column 1 shows the results calculated from the actual data. Columns 2 to 6 show the results calculated from $N = 100$ “forecasters" simulated from different models. The "Cost Function" for RSE of the regressions is defined as: (Model Implied RSE − Actual RSE)^T(Model Implied RSE − Actual RSE).

Figure A7: Impact of Forecaster Conviction (σ_μ) on CG-b Coefficients with Unbiased Prior Belief



Notes: This graph shows the impact of forecaster conviction (σ_μ) on the individual CG slope coefficients b^{cg}, in a simulation with unbiased prior beliefs. The HPA process is simulated according to equations (5) and (6) with parameters $\overline{\mu} = 0.660$, $\overline{\rho} = 0.675$, $\overline{\sigma} = 0.822$, and $\overline{\gamma} = 0.002$. The prior beliefs about the long-run mean and persistence are unbiased: $\mu_\mu = \overline{\mu} = 0.660$, and $\mu_\rho = \overline{\rho} = 0.675$. Other parameters are set to the values in the benchmark model, including $\sigma_\rho = 0.020$.

Alt text (decorative): A figure showing the impact of forecaster conviction (σ_μ) on CG-b coefficients when forecasters in the model has unbiased prior beliefs.

Table A10: Comparing Consensus-Level Model Fit: Benchmark vs. Learning with Fading Memory

Panel A: CG Regression Moments					
Moments	**Horizon**	**Data**	**Benchmark**	**NX (2022)**	
				$\nu = \mathbf{0.0180}$	$\nu = \mathbf{0.0196}$
		(1)	(2)	(3)	(4)
CG *a*	**1-Yr**	3.191	2.442	3.459	3.420
		(0.579)	(0.542)	(0.688)	(0.687)
	2-Yr	4.030	3.969	5.563	5.527
		(0.546)	(0.537)	(0.525)	(0.521)
	3-Yr	4.100	4.343	5.191	5.163
		(0.574)	(0.536)	(0.558)	(0.555)
	4-Yr	4.330	4.128	4.992	4.971
		(0.768)	(0.668)	(0.688)	(0.685)
CG *b*	**1-Yr**	0.668	0.952	7.030	6.137
		(0.629)	(0.580)	(5.944)	(5.495)
	2-Yr	-1.629	-1.959	-14.457	-13.790
		(2.045)	(2.657)	(5.073)	(4.655)
	3-Yr	-2.593	-4.093	1.195	0.751
		(3.198)	(6.448)	(7.053)	(6.476)
	4-Yr	4.514	7.834	8.546	7.571
		(4.976)	(9.625)	(8.723)	(8.012)
Weighted Cost Function			**22.8**	**2102.0**	**1725.8**

Panel B: Sensitivity to Past House Price Growth					
Moments	**Horizon**	**Data**	**Benchmark**	**NX (2022)**	
				$\nu = \mathbf{0.0180}$	$\nu = \mathbf{0.0196}$
Sensitivity to	**1-Yr**	0.188	0.247	0.113	0.122
Past 1Q		(0.025)	(0.014)	(0.017)	(0.018)
Growth	**2-Yr**	0.049	0.091	0.113	0.122
		(0.010)	(0.006)	(0.017)	(0.018)
	3-Yr	0.011	0.041	0.113	0.122
		(0.008)	(0.006)	(0.017)	(0.018)
	4-Yr	-0.003	0.025	0.113	0.122
		(0.009)	(0.006)	(0.017)	(0.018)
Sensitivity to	**1-Yr**	0.103	0.198	0.141	0.152
Past 2Y		(0.033)	(0.027)	(0.013)	(0.014)
Growth	**2-Yr**	0.029	0.088	0.141	0.152
		(0.012)	(0.009)	(0.013)	(0.014)
	3-Yr	0.017	0.039	0.141	0.152
		(0.008)	(0.005)	(0.013)	(0.014)
	4-Yr	0.010	0.027	0.141	0.152
		(0.009)	(0.004)	(0.013)	(0.014)
Sensitivity to	**1-Yr**	0.097	0.218	0.170	0.183
Past 4Y		(0.037)	(0.030)	(0.012)	(0.013)
Growth	**2-Yr**	0.022	0.092	0.170	0.183
		(0.013)	(0.009)	(0.012)	(0.013)
	3-Yr	0.013	0.046	0.170	0.183
		(0.009)	(0.004)	(0.012)	(0.013)
	4-Yr	0.011	0.034	0.170	0.183
		(0.010)	(0.003)	(0.012)	(0.013)
Weighted Cost Function			**116.2**	**1515.5**	**1777.7**

Table A10: Comparing Consensus-Level Model Fit: Benchmark vs. Learning with Fading Memory (cont.)

Panel C: Biases and Inverse MZ Regression Moments					
Moments	**Horizon**	**Data**	**Benchmark**	**NX (2022)**	
				$\nu = \mathbf{0.0180}$	$\nu = \mathbf{0.0196}$
Forecast Biases (%)	**1-Yr**	3.012	2.371	3.669	3.604
		(0.598)	(0.567)	(0.618)	(0.618)
	2-Yr	3.969	3.781	4.765	4.713
		(0.537)	(0.505)	(0.488)	(0.489)
	3-Yr	4.219	4.288	5.255	5.214
		(0.542)	(0.507)	(0.491)	(0.490)
	4-Yr	4.005	4.259	5.280	5.246
		(0.599)	(0.558)	(0.544)	(0.543)
Inverse MZ *a*	**1-Yr**	2.260	2.627	2.014	2.055
		(0.356)	(0.340)	(0.232)	(0.249)
	2-Yr	3.423	3.127	1.811	1.837
		(0.159)	(0.186)	(0.268)	(0.288)
	3-Yr	3.420	2.855	1.543	1.544
		(0.092)	(0.089)	(0.239)	(0.257)
	4-Yr	3.602	2.832	1.531	1.531
		(0.083)	(0.066)	(0.221)	(0.239)
Inverse MZ *b*	**1-Yr**	0.125	0.170	0.056	0.060
		(0.047)	(0.045)	(0.031)	(0.033)
	2-Yr	-0.066	0.004	0.052	0.055
		(0.021)	(0.024)	(0.035)	(0.037)
	3-Yr	-0.050	0.018	0.066	0.071
		(0.011)	(0.011)	(0.030)	(0.032)
	4-Yr	-0.057	0.014	0.053	0.058
		(0.010)	(0.008)	(0.028)	(0.030)
Weighted Cost Function			**349.3**	**1971.7**	**2009.1**

Panel D: Total Weighted Cost Function			
	Benchmark	**NX (2022)**	
		$\nu = \mathbf{0.0180}$	$\nu = \mathbf{0.0196}$
Total Weighted Cost Function	488.3	5589.2	5512.6

Notes: This table shows the comparison of the consensus-level fit between the benchmark model and the model of learning with fading memory of Nagel and Xu (2022). The coefficients shown in Panels A, B, and C are estimated using time-series regressions on consensus forecast data (actual or simulated). Column 1 shows the results calculated from the actual data. Column 2 shows the consensus moments of the benchmark model. Columns 3 and 4 present the results for the model of learning with fading memory: $E_{t+1}(y_{t+h}) = E_t(y_{t+h}) + \nu(y_{t+1} - E_t(y_{t+h}))$. In column 3, the constant gain parameter ν is set to 0.018, taken directly from Nagel and Xu (2022). Column 4 optimizes ν by minimizing the total weighted cost function for consensus moments: $\hat{\nu} = \text{argmin}_{\nu}\, g_{cs}(\nu)'\hat{W}_{cs} g_{cs}(\nu)$, as described in the notes of Table 3. Standard errors are reported in parentheses. Panel D shows the values of the total weighted cost function for different models. All forecast data are in percentages.

Table A11: Comparing Individual-Level Model Fit: Benchmark vs. Only Behavioral Frictions

Panel A: CG Regression Moments					
Moments	**Horizon**	**Data**	**Benchmark**	**Only Overconfidence ($\phi = 1.93$)**	**Only Diagnostic Exp. ($\theta = 1.52$)**
		(1)	(2)	(3)	(4)
CG a	**1-Yr**	3.151	2.553	1.643	1.078
		(0.068)	(0.059)	(0.065)	(0.069)
	2-Yr	4.099	3.879	2.781	2.273
		(0.063)	(0.057)	(0.055)	(0.064)
	3-Yr	4.304	4.274	3.131	3.136
		(0.062)	(0.059)	(0.052)	(0.053)
	4-Yr	4.010	4.256	3.046	3.030
		(0.075)	(0.073)	(0.063)	(0.064)
CG b	**1-Yr**	-0.323	0.029	-0.537	-0.483
		(0.035)	(0.038)	(0.017)	(0.011)
	2-Yr	-0.570	-0.447	-0.668	-0.430
		(0.044)	(0.088)	(0.025)	(0.026)
	3-Yr	-0.576	-0.416	-0.605	-1.648
		(0.049)	(0.107)	(0.040)	(0.041)
	4-Yr	-0.411	-0.273	-0.551	-0.430
		(0.061)	(0.134)	(0.078)	(0.113)
Weighted Cost Function			**1165.5**	**1873.3**	**8463.9**

Panel B: Sensitivity to Past House Price Growth					
Moments	**Horizon**	**Data**	**Benchmark**	**Only Overconfidence ($\phi = 1.93$)**	**Only Diagnostic Exp. ($\theta = 1.52$)**
Sensitivity to Past 1Q Growth	**1-Yr**	0.192	0.247	0.360	0.429
		(0.006)	(0.004)	(0.007)	(0.013)
	2-Yr	0.051	0.091	0.221	0.333
		(0.005)	(0.005)	(0.004)	(0.006)
	3-Yr	0.012	0.041	0.142	0.238
		(0.004)	(0.007)	(0.003)	(0.005)
	4-Yr	-0.003	0.025	0.085	0.123
		(0.004)	(0.007)	(0.002)	(0.002)
Sensitivity to Past 2Y Growth	**1-Yr**	0.108	0.198	0.377	0.263
		(0.006)	(0.005)	(0.007)	(0.015)
	2-Yr	0.029	0.088	0.235	0.296
		(0.005)	(0.005)	(0.005)	(0.008)
	3-Yr	0.015	0.039	0.124	0.169
		(0.004)	(0.006)	(0.003)	(0.005)
	4-Yr	0.008	0.027	0.073	0.087
		(0.004)	(0.006)	(0.002)	(0.002)
Sensitivity to Past 4Y Growth	**1-Yr**	0.102	0.218	0.392	0.348
		(0.007)	(0.005)	(0.008)	(0.016)
	2-Yr	0.022	0.092	0.220	0.300
		(0.005)	(0.005)	(0.005)	(0.008)
	3-Yr	0.010	0.046	0.112	0.163
		(0.004)	(0.006)	(0.003)	(0.005)
	4-Yr	0.007	0.034	0.067	0.077
		(0.004)	(0.007)	(0.002)	(0.002)
Weighted Cost Function			**896.7**	**8903.0**	**15695.6**

Table A11: Comparing Individual-Level Model Fit: Benchmark vs. Only Behavioral Frictions (cont.)

Panel C: Biases and Inverse MZ Regression Moments

Moments	Horizon	Data	Benchmark	Only Overconfidence ($\phi = 1.93$)	Only Diagnostic Exp. ($\theta = 1.52$)
Forecast Biases	**1-Yr**	2.952	2.371	1.419	0.913
		(0.065)	(0.061)	(0.071)	(0.081)
	2-Yr	4.025	3.781	2.679	2.155
		(0.059)	(0.057)	(0.058)	(0.065)
	3-Yr	4.299	4.288	3.159	2.888
		(0.058)	(0.058)	(0.052)	(0.061)
	4-Yr	4.021	4.259	3.048	3.003
		(0.062)	(0.064)	(0.055)	(0.055)
Inverse MZ *a*	**1-Yr**	2.226	2.627	3.391	0.964
		(0.065)	(0.052)	(0.082)	(0.131)
	2-Yr	3.403	3.127	4.316	4.396
		(0.068)	(0.063)	(0.068)	(0.101)
	3-Yr	3.424	2.855	3.907	4.955
		(0.063)	(0.070)	(0.042)	(0.062)
	4-Yr	3.600	2.832	3.895	3.903
		(0.059)	(0.072)	(0.025)	(0.028)
Inverse MZ *b*	**1-Yr**	0.130	0.170	0.201	0.688
		(0.009)	(0.007)	(0.011)	(0.017)
	2-Yr	-0.064	0.004	-0.009	0.055
		(0.009)	(0.008)	(0.009)	(0.013)
	3-Yr	-0.051	0.018	0.029	-0.078
		(0.008)	(0.009)	(0.005)	(0.008)
	4-Yr	-0.057	0.014	0.035	0.040
		(0.007)	(0.009)	(0.003)	(0.003)
Weighted Cost Function			**623.7**	**2505.5**	**7468.5**

Panel D: Forecaster Dispersion Moments

Moments	Horizon	Data	Benchmark	Only Overconfidence ($\phi = 1.93$)	Only Diagnostic Exp. ($\theta = 1.52$)
Forecaster Dispersion	**1-Yr**	6.476	3.033	7.695	−−
	2-Yr	4.893	3.435	2.689	−−
	3-Yr	3.490	3.805	0.942	−−
	4-Yr	2.946	4.034	0.332	−−
Weighted Cost Function			**5865.1**	**16609.8**	−−

Panel E: Total Weighted Cost Function

	Benchmark	Only Overconfidence ($\phi = 1.93$)	Only Diagnostic Exp. ($\theta = 1.52$)
Total Weighted Cost Function	8551.1	29891.6	−−
Coefficient (weighted)	2686.0	13281.8	31627.9
Dispersion (weighted)	5865.1	16609.8	−−

Notes: This table shows the comparison of the individual-level fit between the benchmark model and models with only overconfidence or diagnostic expectations, as described in Section (6). Forecasters are assumed to know the true parameters, $\bar{\mu}$ and $\bar{\rho}$, of the house price growth process. Column (3) shows the individual-level fit of the model with only overconfidence. The parameters of this model are $(\phi, \overline{\Gamma}, \overline{\sigma}, \overline{\gamma})$, and are obtained by optimizing the same objective function as in Equation (12). Column (4) shows the individual-level fit of the model with only diagnostic expectations. Forecasters are assumed to have no heterogeneous signals and can observe the current quarter's realized HPA without noise. The parameters of this model are $(\theta, \overline{\sigma}, \overline{\gamma})$, and are obtained by optimizing Equation (12). $\bar{\mu}$ and $\bar{\rho}$ are calibrated by fitting an AR(1) process to the whole sample. The procedure for coefficient estimation is the same as described in Table (4).

I The Speed of Learning

Our key mechanism, "uncertainty about the long-run mean" relies on the premise that forecasters' learning process is slow. We examine how slow this process is through simulation.

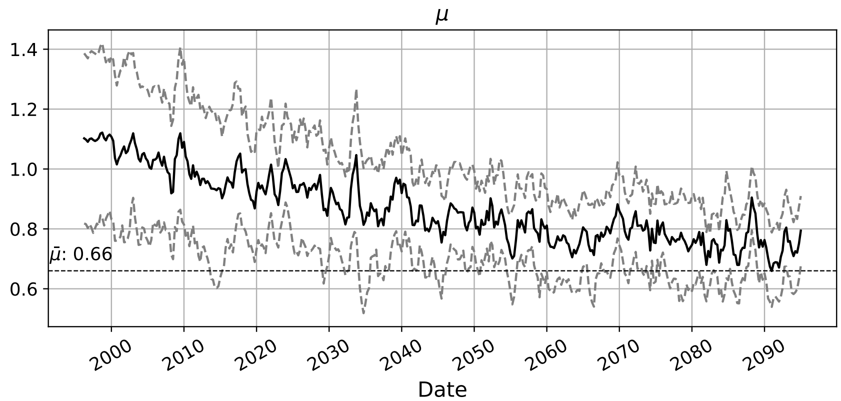
We simulate 100 years of house price growth data from (5) and (6) with $\overline{\mu} = 0.660$, $\overline{\rho} = 0.675$, $\overline{\sigma} = 0.822$, and $\overline{\gamma} = 0.002$. Subsequently, we generate HPA forecasts based on our rational learning model with upward- (Panel a), downward- (Panel b) and un-biased (Panel c) priors about $\overline{\mu}$, all with the same level of uncertainty of $\sigma_\mu = 0.185$. The resulting posterior parameter estimates are plotted in Figure A8. Across different values of the prior, parameter learning shrinks the uncertainty about the parameter $\overline{\mu}$, as indicated by the narrower bands (5% and 95% percentile) around the posterior estimates as time goes by. Furthermore, the learning is indeed slow; after 100 years of learning, forecasters still revise their posterior estimates visibly, although their estimates have become much closer to the true value.

We conduct a similar simulation exercise on the learning of $\overline{\rho}$, with a tight prior $\sigma_\rho = 0.020$, consistent with our benchmark estimate. As Figure A9 shows, the speed of learning about $\overline{\rho}$ is even slower than the learning about $\overline{\mu}$, due to the tighter priors. Indeed, after 100 years of learning, forecasters' posterior estimates have only moved about 0.05 closer to the true value.

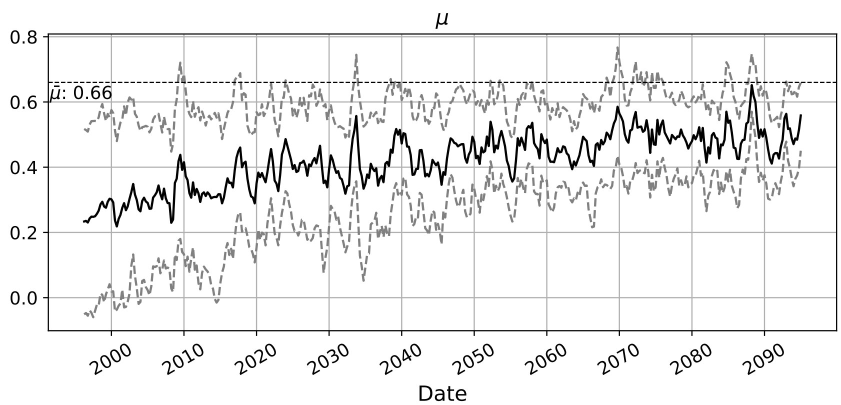
Our benchmark model estimation indicates that the forecasters in the Pulsenomics survey have downward-biased priors about $\overline{\rho}$ on average (corresponding to panel b in Figure A9). Furthermore, these forecasters are quite certain about their prior estimate. This is the key feature that leads to the salient pattern of "pulling forecasts back to the mean", highlighted in Figure 1. In other words, the forecasters believe that house price growth is more "transitory" than it actually is.

Overall, the results in Figure A8 and A9 show that learning about the parameters is indeed slow, and more so when forecasters have tight priors. Furthermore, forecasters in our sample exhibit more "modest" prior beliefs about the long-run mean than about the persistence. This observation motivates us to investigate the impact of behavioral biases on the short-term versus the long-term forecasts in Section 6.

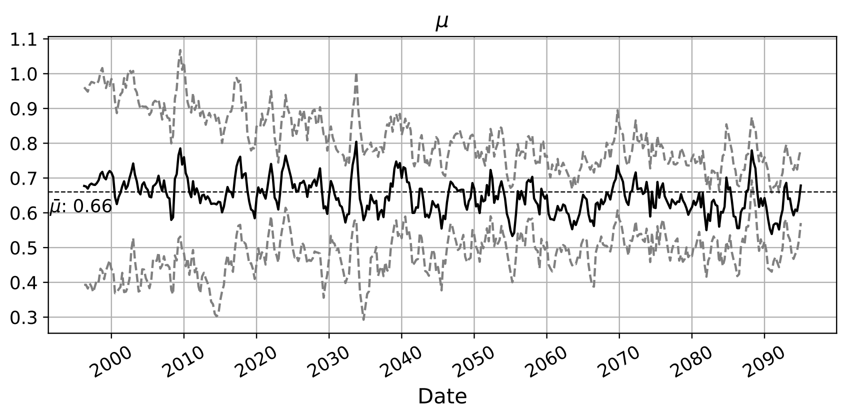
Figure A8: Posterior Beliefs about the Long-Run Mean ($\overline{\mu}$) with Different Priors Based on Simulated Data



(a) Upward-Biased Prior



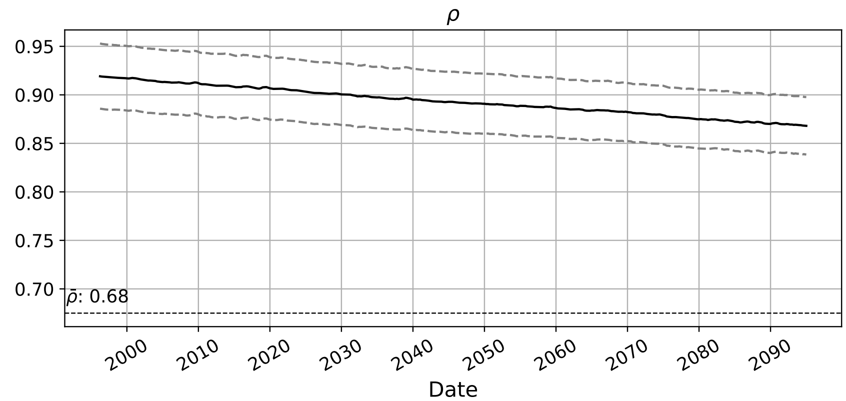
(b) Downward-Biased Prior



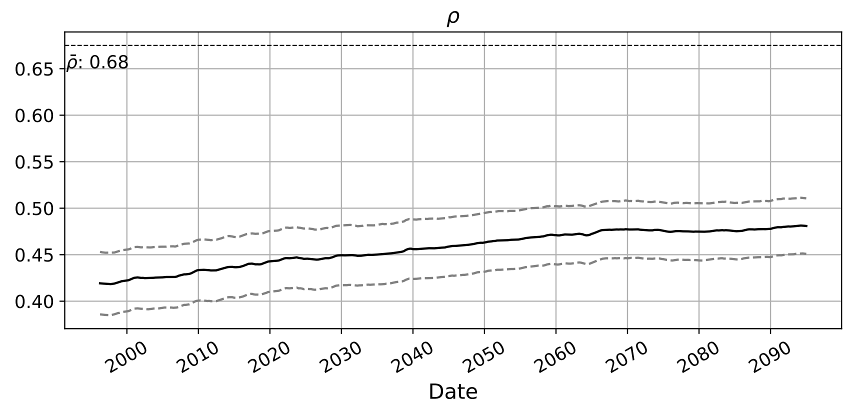
(c) Unbiased Prior

Notes: This figure shows the evolution of posterior beliefs about the long-run mean ($\overline{\mu}$) over time. The HPA process is simulated according to equations (5) and (6), with parameters $\overline{\mu} = 0.660$, $\overline{\rho} = 0.675$, $\overline{\sigma} = 0.822$, and $\overline{\gamma} = 0.002$. Panels (a), (b), and (c) show the evolution of posterior beliefs about the long-run mean $\overline{\mu}$ under different priors: upward-biased ($\mu_\mu = 1.1$), downward-biased ($\mu_\mu = 0.2$), and unbiased ($\mu_\mu = \overline{\mu} = 0.660$), all with a standard deviation of $\sigma_\mu = 0.185$. Prior beliefs about persistence are unbiased: $\mu_\rho = \overline{\rho} = 0.675$, with $\sigma_\rho = 0.020$. The black lines show the mean of the posterior distribution. The grey dashed lines show the 5% and 95% percentiles of the posterior distribution.
Alt text: A figure showing the long term (100 year) evolution of model-implied posterior beliefs about the long-run mean ($\overline{\mu}$) with upward-biased (Panel A), downward-biased (Panel B) and unbiased prior (Panel C) beliefs.

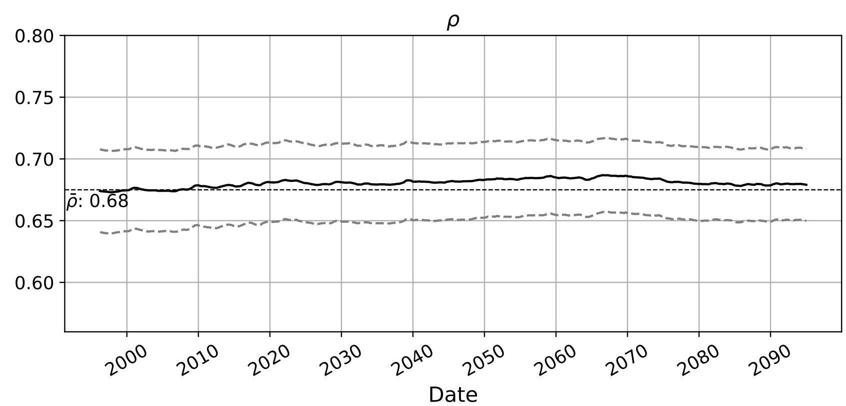
Figure A9: Posterior Beliefs about the Persistence $\overline{\rho}$ with Simulated Data



(a) Upward-Biased Prior



(b) Downward-Biased Prior



(c) Unbiased Prior

Notes: This figure shows the evolution of posterior beliefs about the persistence ($\overline{\rho}$) over time. The HPA process is simulated according to equations (5) and (6), with parameters $\overline{\mu} = 0.660$, $\overline{\rho} = 0.675$, $\overline{\sigma} = 0.822$, and $\overline{\gamma} = 0.002$. Panels (a), (b), and (c) show the evolution of posterior beliefs about the persistence $\overline{\rho}$ under different priors: upward-biased ($\mu_\rho = 0.92$), downward-biased ($\mu_\rho = 0.42$), and unbiased ($\mu_\rho = \overline{\rho} = 0.675$), all with a standard deviation of $\sigma_\rho = 0.020$. Prior beliefs about the long-run mean are unbiased: $\mu_\mu = \overline{\mu} = 0.660$, with $\sigma_\mu = 0.185$. The black lines show the mean of the posterior distribution. The grey dashed lines show the 5% and 95% percentiles of the posterior distribution.
Alt text: A figure showing the long term (100 year) evolution of model-implied posterior beliefs about the persistence ($\overline{\rho}$) with upward-biased (Panel A), downward-biased (Panel B) and unbiased prior (Panel C) beliefs.