DISCUSSION PAPER SERIES

DP16223

Learning about Housing Cost: Survey Evidence from the German House Price Boom

Fabian Kindermann, Julia Le Blanc, Monika Piazzesi and Martin Schneider

FINANCIAL ECONOMICS

MONETARY ECONOMICS AND FLUCTUATIONS

CEPR

ISSN 0265-8003

Learning about Housing Cost: Survey Evidence from the German House Price Boom

Fabian Kindermann, Julia Le Blanc, Monika Piazzesi and Martin Schneider

Discussion Paper DP16223
Published 03 June 2021
Submitted 01 June 2021

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Learning about Housing Cost: Survey Evidence from the German House Price Boom

Abstract

This paper uses new household survey data to study expectation formation during the recent housing boom in Germany. The cross section of forecasts depends on only two household characteristics: location and tenure. The average household in a region responds to local conditions but underpredicts local price growth. Renters make on average higher and hence more accurate forecasts than owners, although their forecasts are more dispersed and their mean squared forecast errors are higher. A quantitative model of learning about housing cost can match these facts. It emphasizes the unique information structure of housing among asset markets: renters who do not own the asset are relatively well informed about its cash flow, since they pay for housing services that owners simply consume. Renters then make more accurate forecasts in a boom driven by an increase in rents and recovery from a financial crisis.

JEL Classification: N/A

Keywords: N/A

Fabian Kindermann - fabian.kindermann@ur.de
University of Regensburg and CEPR

Julia Le Blanc - julia.le-blanc@ec.europa.eu
European Commission, Joint Research Centre

Monika Piazzesi - piazzesi@stanford.edu
Stanford University and CEPR

Martin Schneider - schneidr@stanford.edu
Stanford University and CEPR

Acknowledgements
We thank Rudi Bachmann, Theresa Kuchler, Ulrike Malmendier, Charles Nathanson, Tom Sargent, Johannes Stroebel, Stijn van Nieuwerburgh, Michael Weber, and seminar participants at Bonn, Chicago, Deutsche Bundesbank, ECB, Federal Reserve Bank at Cleveland, Harvard, HULM, Leuven, Manchester, Munich, "Micro and Macro Implications of Household Finance", NBER Summer Institute, NYU, Queens, UBC, and UC Santa Cruz for helpful comments. A large part of this paper was written while Julia Le Blanc was a senior research economist at the Deutsche Bundesbank in Frankfurt. We thank the Deutsche Bundesbank for kindly providing us with all their avail- able data on house prices and expectations as well as with an excellent research infrastructure.

Learning about Housing Cost: Survey Evidence from the German House Price Boom

Fabian Kindermann
University of Regensburg & CEPR

Julia Le Blanc
European Commission, Joint Research Center

Monika Piazzesi
Stanford & NBER

Martin Schneider
Stanford & NBER

May 31, 2021

Abstract

This paper uses new household survey data to study expectation formation during the recent housing boom in Germany. The cross section of forecasts depends on only two household characteristics: location and tenure. The average household in a region responds to local conditions but underpredicts local price growth. Renters make on average higher and hence more accurate forecasts than owners, although their forecasts are more dispersed and their mean squared forecast errors are higher. A quantitative model of learning about housing cost can match these facts. It emphasizes the unique information structure of housing among asset markets: renters who do not own the asset are relatively well informed about its cash flow, since they pay for housing services that owners simply consume. Renters then make more accurate forecasts in a boom driven by an increase in rents and recovery from a financial crisis.

Email addresses: fabian.kindermann@ur.de, julia.le-blanc@ec.europa.eu, piazzesi@stanford.edu, schneidr@stanford.edu. We thank Rudi Bachmann, Theresa Kuchler, Ulrike Malmendier, Charles Nathanson, Tom Sargent, Johannes Stroebel, Stijn van Nieuwerburgh, Michael Weber, and seminar participants at Bonn, Chicago, Deutsche Bundesbank, ECB, Federal Reserve Bank at Cleveland, Harvard, HULM, Leuven, Manchester, Munich, "Micro and Macro Implications of Household Finance", NBER Summer Institute, NYU, Queens, UBC, and UC Santa Cruz for helpful comments.

A large part of this paper was written while Julia Le Blanc was a senior research economist at the Deutsche Bundesbank in Frankfurt. We thank the Deutsche Bundesbank for kindly providing us with all their available data on house prices and expectations as well as with an excellent research infrastructure.

1 Introduction

Expectation formation is a key ingredient of any account of asset market fluctuations. Recent literature has turned to survey data to directly measure subjective expectations. For housing markets, heterogeneity of expectations is particularly important: market participants include a diverse group of individual households and booms often go along with strong regional differences. Moreover, there are only few transactions in housing markets, and these often involve buyers who are currently renting: households who do not own the asset before the transaction.[1] Unfortunately, these same features have made it more difficult to measure the expectations that matter in these transactions. During past booms, such as the US housing boom of the early 2000s, major representative household surveys did not yet incorporate detailed questions on expectation formation that would have allowed researchers to explore the distribution of forecasts.

This paper studies expectation formation during the recent German housing boom. We first use new Bundesbank survey data to characterize quantitative house price forecasts across regions with different local housing market conditions, controlling for a rich set of demographics. On average, households underpredict actual house price growth. In the cross section of households, location and housing tenure are key determinants of price growth forecasts, while other characteristics such as age, income, wealth, risk aversion, and financial literacy play only minor roles. On average, renters' price growth forecasts are 2 percentage points higher, and hence more accurate, than those of owners. At the same time, the average renter's root mean squared error is more than 1pp higher than that of the average owner. The differences are sizeable in light of actual growth between 3% and 9% per year, depending on the region.

We then develop a theory of learning about housing cost that can quantitatively account for our facts on expectation formation. The basic idea is that housing is a special asset when it comes to information about cash flow. Non-owners, who rent and tend to have renter neighbors, have easy access to information about cash flow. In contrast, owners directly consume services from their home (as do their owner neighbors) and thus need not pay much attention to rent, but get better information about house prices. During a boom that features strong rent growth as well as recovery from financial distress, knowledge of rents gives the average renter an edge in forecasting, even though noisier signals about house prices generate a higher dispersion of forecasts and larger mean squared forecast errors among renters.

[1] In Germany, 58 percent of homeowners who changed their main residence within the last decade were previously renting. In contrast, only 3 percent of renters who changed their main residence within the last decade were previously homeowners. These numbers are from the 2014 Panel of Household Finances, which is further explained below.

In our model, differences in information matter for forecasts because house prices reflect two persistent factors. First, as in many other asset markets, mean reverting fluctuations in the stochastic discount factor generate excess volatility in prices over the present values of future rents. Second, rent growth is itself quite persistent, unlike cash flow growth in other markets such as equity. In our model, forecasters at any point in time are not sure which force is driving prices: this is because they observe prices and rents with error, for example they might sample only a few properties in their area. Those households who believe a boom is due to rent growth expect the boom to continue – they forecast momentum in price growth. In contrast, those who believe the boom is due to lower discount rates – for example due to lower financial frictions – expect mean reversion and forecast lower price growth.

Our explanation for forecast differences is centered around this signal extraction problem. By 2010, Germany had seen a long period of decline in price-rent ratios, most recently reinforced by the global financial crisis. At the same time, rent growth had been sluggish. A reasonable initial belief thus had low rent growth and a high discount rate both contributing to low prices. As the economy recovered and the housing boom got underway, owners and renters interpreted price movements differently. Renters were aware of strong rent growth leading price growth and attributed less of the boom to low discount rates. Both effects led them to high price growth forecasts. In contrast, owners paid less attention to rents and hence attributed more of the boom to a recovery, which made them more pessimistic about its continuation.

Most of our empirical work is based on the 2014 Panel on Household Finances (PHF), a representative survey by the Bundesbank that asks detailed questions not only about household income and balance sheets, but also about expectations and behavior. In addition to measuring forecast distributions, we use PHF data to check a number of premises of our theory. First, we show that housing is indeed a special asset in that non-owners feel confident forming an opinion about prices. Indeed, when the survey elicits stock price expectations, about one half of non-owners take the "don't know" option. In contrast, the overwhelming majority of renters – close to 85% – provide a quantitative forecast of house price growth.

We further show that owners are aware of actual price movements, so their forecasting mistakes cannot be attributed to a lack of attention to the market. Owners' nowcasts of how much their own homes have been appreciating are close, on average, to the actual growth rates. Confidence bounds around these nowcasts are large, consistent with imperfect information among owners. Moreover, forecast differences between owners and renters are not due to whether they have recently moved or plan to move in the near future, or whether renters plan to buy or not. These findings suggest that endogenous information acquisition motivated by plans to trade is not behind our findings, and so we do not incorporate it into our model.

Some households own their primary residence as well as other real estate. While such real estate investors are owners, they have more information about rents than other owners. Our model predicts that these real estate investors should have forecasts that are higher than forecasts of other owners, and more similar to renter forecasts. This indeed is the case in the data.

As a piece of direct evidence on learning, we proposed a question on information acquisition to the Bundesbank's 2019 Online Panel - Households (BOP-HH). In particular, we elicit the importance of various sources of information households use when forecasting house prices. Consistent with our approach, the most important source of information – mentioned by more than 80% of households – is direct observation of prices, a noisy signal of local conditions. Other sources that aggregate information, such as classical and social media or financial advisors, play a smaller role. In the cross section, owners look more at prices, and renters look more at rents, as one would expect. Moreover, more renters indicate that talking to family and friends is important. Altogether, these findings support our approach of modeling expectation formation based on noisy signals that reflect subjective experience.

Formally, our model describes the joint dynamics of prices and rents perceived by households. We restrict beliefs to respect a standard asset pricing equation that expresses prices as a discounted stream of future rents. The equation extends the familiar user cost model of housing and can be interpreted as the first order condition of a developer who is active in both housing and rental markets. We further assume that households are aware of the two persistent factors driving prices, rent growth and the discount rate. Households observe prices and rents with error, meant to capture their sampling of a few properties that produces an imperfect signal of current market conditions. Household types – owners and renters – are identified by the magnitude of the errors, or equivalently the precision of their signals. They arrive at forecasts by forming conditional expectations via Bayes' rule. For a distribution of initial beliefs and a sequence of actual prices and rents, the model then generates a sequence of forecast distributions for owners and renters.

We quantify the model by inferring parameters of households' subjective belief and information structure from data on prices and price forecasts. We assume that, at the start of the boom in 2010, the average owner and renter agree on the price-rent ratio as well as the underlying discount and rent growth rates. We then identify parameters by matching means and dispersion of survey forecasts in our 2014 survey data, as well as unconditional moments of prices and rents in the data. The success of the quantitative exercise is that the observation of a few years of boom prices can generate the observed divergence between average owner and renter beliefs, even though signals contain enough noise to generate observed dispersion in forecasts.

Related Literature. We follow a long tradition of studying house price expectations through the lens of a present-value relationship. Glaeser and Nathanson (2015) survey applications of the user cost model. The classic challenge is to find a mechanism for expectation formation about rent growth so that the present value of rents satisfies key properties of regional house prices: short term momentum (Case and Shiller, 1989, Guren, 2018), long term reversal (Cutler, Poterba and Summers, 1991, Head, Lloyd-Ellis and Sun, 2014), as well as excess volatility (Glaeser, Gyourko, Morales and Nathanson, 2014). Many authors have turned to extrapolation from recent price observations, often via non-Bayesian updating schemes (see, for example, Gelain and Lansing, 2014, Glaeser and Nathanson, 2017 or DeFusco, Nathanson and Zwick, 2018.) This feature also helps generate extrapolation in booms, such as documented in survey data from US cities by Case and Shiller (2003) and Case, Shiller and Thompson (2012), for example. The learning mechanism in our model relies instead on agents' confusion of persistent rent growth, which generates momentum – as in many user cost models – and the discount rate, which generates reversal. We share the emphasis on discount rate volatility with a broader literature in asset pricing.[2]

At the heart of our model is differential updating of beliefs based on experience, in our case for renters and owners who see prices and rents with different precision. The relevance of experience for both expectations and choice has been established by a large empirical literature. Early work focused on stock return expectations. Vissing-Jorgensen (2003) and Malmendier and Nagel (2011) documented cohort effects on stock price expectations in the short and long run, respectively. This evidence is consistent with experience-driven learning models of stock prices such as Cogley and Sargent (2008). In the context of housing, Malmendier and Steiny (2021) relate inflation experience to home ownership. Bailey, Cao, Kuchler and Stroebel (2018) show that individuals' home purchase decisions are related to the experience of their friends on social networks. Kuchler and Zafar (2019) document that households' views about both first and second moments of future house prices are affected by recent local price observations. Armona, Fuster and Zafar (2019) describe an information experiment, in which individuals revise forecasts when they are told the actual recent price experience in their region. These results are consistent with our assumption that households do not have a firm view of "the market price of housing", but instead base their forecasts on noisy signals.

There is limited work on heterogeneity of beliefs between renters and owners. Favara

[2]The typical statistical decomposition of price-dividend ratios for equity in Campbell and Shiller (1988) attributes the overwhelming majority of variation to discount rate news. For an analogous result for US aggregate housing indices, see Campbell, Davis, Gallin and Martin (2009). Our exercise does not assume properties of this statistical decomposition but instead infers a subjective version of it. Here we follow De la O and Myers (2019) who use analyst forecasts to derive a subjective decomposition of price-dividend ratio movements for equity; their main result is that most variation comes from cash flow news. See Lewellen and Shanken (2002) for a Bayesian learning model that is consistent with this result.

and Song (2014) develop a theoretical model of the housing market with asymmetric information: in equilibrium more optimistic agents sort into ownership, whereas renters are more pessimistic. Our results suggest that asymmetric information is indeed important, but that sorting does not occur right away during a boom, perhaps because of transaction costs. It is plausible, however, that the belief differences we document lead to trade; since renter-owner transitions are a large share of housing market volume, they may thereby have a large effect on transaction prices.[3] Adelino, Schoar and Severino (2018) measure households' perceptions of house price risk and show that renters view housing as riskier than owners. In our model, renters' subjective variance of house prices is also higher than that of owners. Indeed, while renters have better signals of rent, they rely on noisier signals of prices. As a result, they perceive high uncertainty about the discount rate, which accounts for the bulk of price volatility.

Finally, our work relates to an emerging literature that incorporates survey evidence on expectations into quantitative models of housing choice and house prices. Landvoigt (2017) as well as Kaplan, Mitman and Violante (2020) have emphasized the role of expectations for the 2000s housing boom in the US. Leombroni, Piazzesi, Schneider and Rogers (2020) study how heterogeneous inflation expectations affected house prices and interest rates in the 1970s. Ludwig, Mankart, Quintana, Vellekoop and Wiederholt (2020) show how heterogeneous beliefs drive price dynamics in the recent Dutch housing boom-bust cycle; they also rely on a dataset with detailed income, balance sheet and expectations information at the household level. In this paper we approach the relationship between survey expectations and prices from the opposite side: rather than establish an effect of the distribution of beliefs on prices, we derive distributions of beliefs from price (and rent) histories.

The paper is structured as follows. Section 2 presents historical data on the housing market in Germany. Section 3 introduces the survey data and documents key stylized facts about the cross section of house price growth forecasts. Section 4 provides direct evidence on the information sets of renters and owners. Section 5 presents the model and its quantitative implications. The appendix contains details about the data, empirical results, and model derivations.

[3]Piazzesi and Schneider (2009) study a search model with heterogeneous beliefs and short sale constraints and show how entry of a small share of optimistic agents can have a large price impact. See Gao, Sockin and Xiong (2020) for a model where information frictions in housing affect migration and local capital accumulation.

2 Housing in Germany

In this section, after describing the institutional features of the German housing and rental market, we provide an overview of recent price and rent movements in Germany. In the aftermath of the financial crisis, Germany emerged from a decades-long housing slump. The boom saw an initial acceleration in rents followed by a boom in prices and price-rent ratios. New regional price data further show substantial regional heterogeneity, with a stronger boom starting earlier in major metropolitan areas.

2.1 Institutional Background

The German housing market is characterized by a low homeownership rate. Between 2010 and 2017 the share of households owning their main residence ranged around 44 percent, compared with about 60 percent in the Euro area as a whole. Germany has high transfer taxes on buying real estate, no mortgage interest tax deductions for owner-occupiers, and strong renter protections, see for example Kaas, Kocharkov, Preugschat and Siassi (2020) and Weber and Lee (2018). Overall, the rental market is very well developed and households are able to find rental housing in essentially all locations and quality segments.

The German rental market offers housing units of different qualities. Compared to the U.S., where the markets for owner-occupied and rental units are highly segmented, the German housing markets are far more integrated. For example, there are rental units among single family homes and townhouses in the top quality segment. As a result, renting is a choice made by a broad part of the population. In fact, in 2014 the homeownership rate in the top income decile was only 71 percent, while the CPS reports a number of almost 84 percent for the US. Even in the top net wealth quintile, there are 12 percent of households who rent their primary residence in Germany.

2.2 Average house price and rent data

Our housing data come from bulwiengesa AG, a leading German real estate data and consulting firm. For the period since 2005, they provide average transaction prices and market rents in each of the 401 German counties, measured in Euros per square meter.[4] For each county, we observe separate price indices by dwelling type (single-family house, town house as well as existing and newly built apartments), while for rentals we have

[4]Throughout, we use the term "county" to denote both "Kreise" – the administrative subdivision above a municipality – and "kreisfreie Staedte" – cities that are not part of any "Kreis" but provide the services of a "Kreis" at the municipality level.

separate rent indices for existing and newly built apartments. To compare prices to survey forecasts below, we would like to measure the value of the typical house in a region – the subject of the survey question – as opposed to the average transaction price. We would also like to aggregate price data to larger regions for which we have enough data to meaningfully measure average forecasts.

We thus build price indices for Germany as a whole as well as four regions that differ in average growth rates. We aggregate county level indices to larger regions – for example, all of Germany – using survey weights from the 2014 Bundesbank's Panel on Household Finances (PHF), a representative country-wide survey that we also use to measure forecasts (described in more detail below).[5] For every household in the survey, we know county and dwelling type. We can thus compute, for any region, the weighted average local rent or house price using the information on dwelling types of all survey respondents living in a region. The advantage of this approach – over weighting prices by transactions – is that the composition of our price indices resembles as closely as possible the composition of average forecasts in the survey. If the survey is indeed representative, then so are our actual price indices and average forecasts. In Appendix A we provide more information on the weighting procedures.

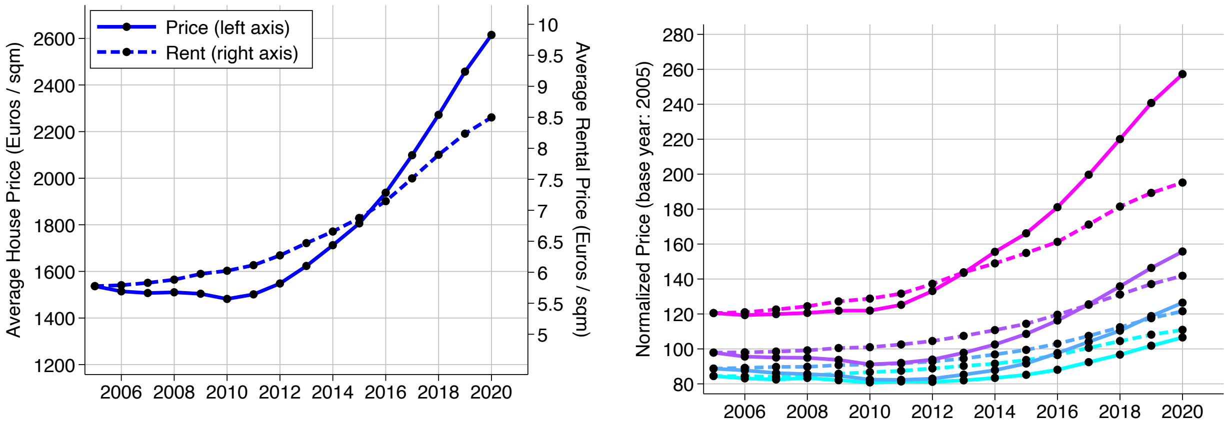
Figure 1 summarizes our data on house prices and rents. The left panel shows the behavior of Germany-wide aggregates: average house prices (solid line, left axis) and rental prices (dashed line, right axes) since the year 2005. Both price series are in Euros per square meter. During the years 2005-2020, house prices increased by 68 percent, while rental prices only increased by 46 percent. Consequently, the average price-rent ratio increased by almost 40 percent. We arranged the axes so that the price series start at exactly the same point in 2005 and growth rates are comparable between the two series. The graph shows that house prices stagnated until 2010, while rents were already growing since 2005. This combination led to a fall in the price-rent ratio from 2005 to 2010. After 2010, the growth rate of house prices outpaced the growth rate of rents, so that the price-rent ratio started to increase.

2.3 House price and rent data by region

Different regions in Germany have fared very differently during the housing boom. To document this regional heterogeneity, we divide the 401 counties into four "growth regions", that is, quartiles according to trend nominal house price growth over the decade 2010-2020. Here trend growth is the slope of a linear regression of the log house price on time. The right panel of Figure 1 shows the evolution of house prices (solid lines) across

[5] Reassuringly, when we weight county indices using Census data on county population sizes in the year 2014 instead of survey weights, we obtain very similar price indices.

Figure 1: House and Rental Prices Across Regions



(a) Average house prices and rents

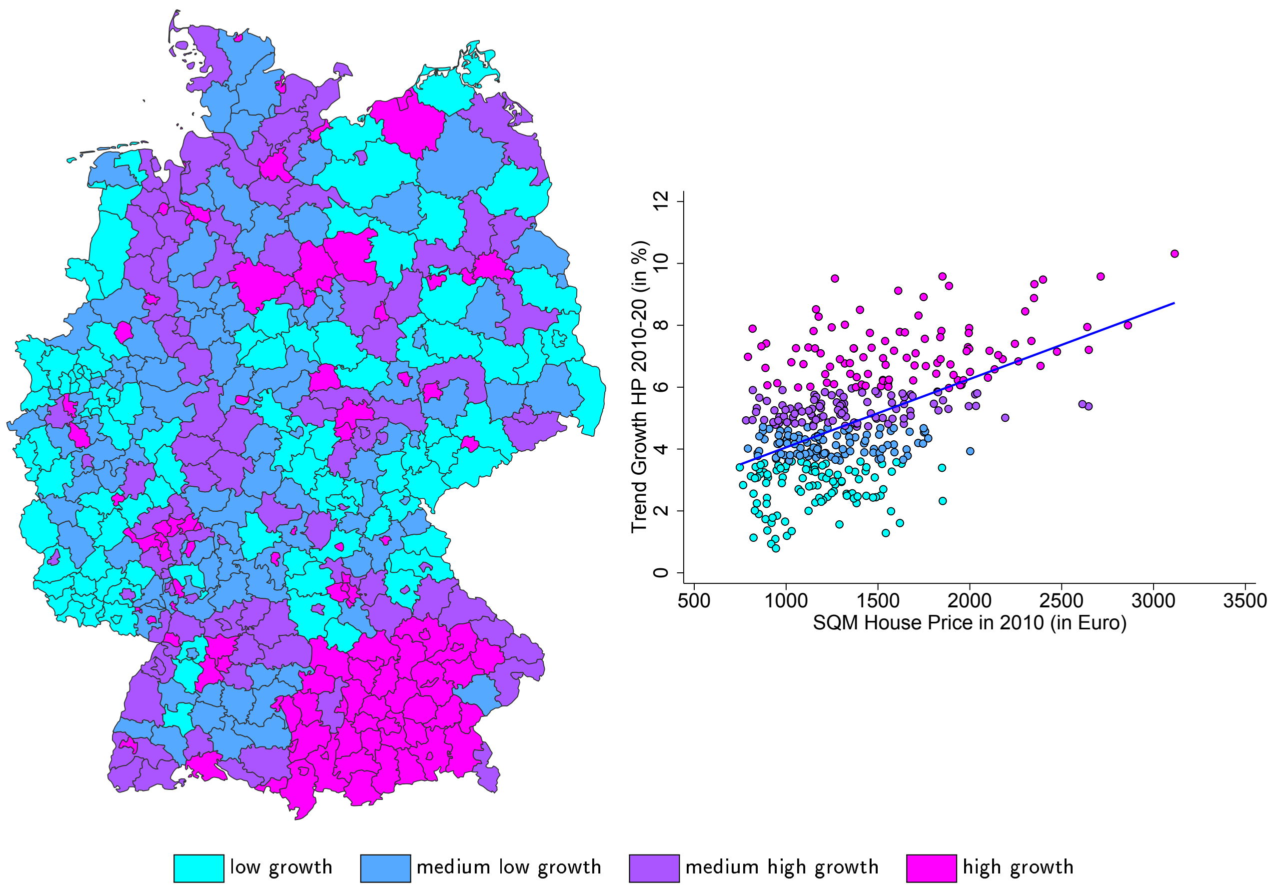
(b) High versus low growth regions

House prices are solid lines, rental prices are dotted lines. Panel (a) measures the national average of house prices on the left axis, rents on the right axis both in Euros per square meter. Panel (b) shows house prices in four growth regions (counties sorted into quartiles based on their trend growth over the years 2010-2020) normalized so that 100 corresponds to the national average in 2005. The rental series are normalized to the same value as the corresponding price series in 2005. *Source:* Own calculations based on data from bulwiengesa AG, PHF and OECD.

growth regions since 2005, normalized so that a level of 100 corresponds to the national average in 2005. There are large regional differences in the strength of the boom: in the high growth region, house prices more than double during the boom, while in the low growth region house prices increase by merely 25%. Moreover, the boom worked its way slowly through the regions: it started earlier in the highest growth region, with the first strong growth rate in 2011, whereas in the lowest growth region prices began to pick up only in 2014. The figure also shows the evolution of rental prices (dotted lines). Here we have normalized the rent series for each region to start in 2005 at the same values as the house price in the same region, in order to allow for a simple comparison of price and rent within regions. In all regions, rents began to grow before prices, but have now been overtaken in all but the lowest growth region.

To understand which regions belong to the faster growing regions, we produce a map of German counties in the left panel of Figure 2, with different colors indicating differences in the extent of trend house price growth. (The color coding is the same as in the right panel of Figure 1). High house price growth is concentrated in larger cities, and in particular the metropolitan area of München and surrounding Bavarian cities such as Ingolstadt. Low growth is prevalent in more rural areas, but also the Ruhr area, the largest historical coal mining region in Germany now in decline. In the right panel of Figure 2, we correlate trend house price growth (measured on the vertical axis) with the house price at the beginning of the housing boom (in the year 2010, measured on the horizontal

Figure 2: House Prices and Growth Across Germany



Source: Own calculations based on data from bulwiengesa AG and PHF.

axis). Many regions that have been growing the most began with relatively high prices in the years before the boom. Thus, the boom was not a period of convergence for different regions, but instead amplified regional differences. Since the relative ranking in housing affordability of different areas has not changed during this decade, the new media often focuses its coverage on the seven most expensive cities, which are referred to as the "Big Seven" (in English language for emphasis) and colored in pink in our map: München, Frankfurt, Hamburg, Berlin, Stuttgart, Düsseldorf, and Köln.

2.4 The house price boom in historical perspective

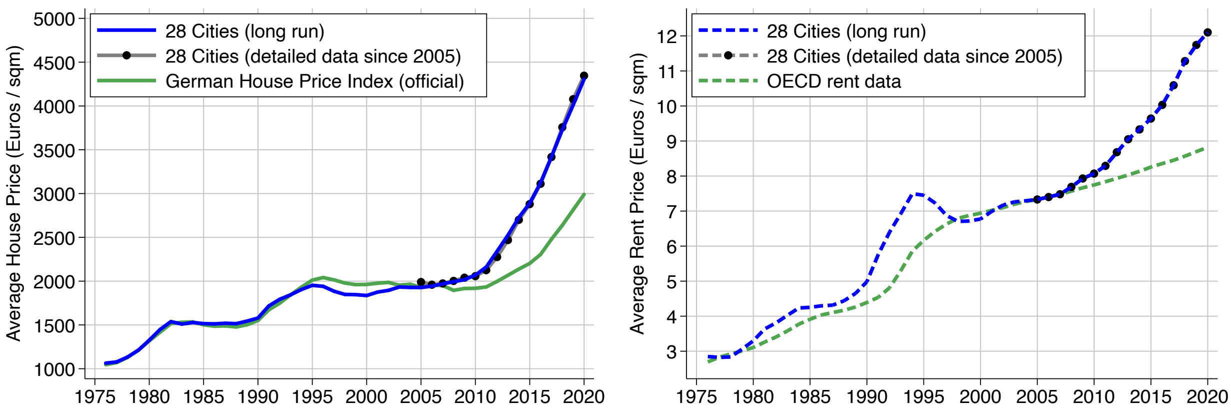
The detailed county-level dataset for Germany covers a relatively short sample. To provide a historical perspective of housing markets in Germany, the green line in the left panel of Figure 3 shows a Germany-wide house price index going back to the 1970s. This "official" index is calculated by the Deutsche Bundesbank and also used, for example, by the OECD and the BIS. The underlying data consists of house prices in a sample of 50 German cities during the 1970s and 1980s. After reunification, the number of sample cities was extended to 100 and successively to 125. Finally, starting from 2005, the Bundesbank uses data on the full set of 401 counties from bulwiengesa AG and combines it with data from other providers. Because of these changes in methodology, there are several structural breaks in the series, which make it rather difficult to interpret its evolution over time.

To obtain a more consistent series, we obtain house price data for 50 German cities since the 1970s from bulwiengesa AG, and select those 28 cities that are in the highest growth region during the current house price boom. The house price data only has prices for newly built apartments going back to 1970s, which are always higher than prices for existing properties. We therefore use the county-level data to compute a short house price series for all properties in the 28 cities for 2005-2020. We then rescale our long series for newly built apartments so that it matches the short series on average (over the years 2005-2020). The left panel of Figure 3 shows the long house price data on newly built apartments (in blue) and the short series based on detailed data (in grey with black dots.) We see that the difference between the two series is only in levels. Rescaling the long-run newly build price series makes the two lines almost indistinguishable. It is reassuring that the long series is actually a good proxy for the overall housing market performance in these 28 cities. Our long series is similar to the official German aggregate house price index during the 1970s and 1980s, but the two series diverge somewhat starting in the 1990s and show a remarkably different behavior starting in 2005.

With our house prices index for the same 28 cities, we can now interpret the evolution of house prices over time. The blue line in the left panel of Figure 3 shows that the current house price boom was an unusual development. Nominal house prices approximately doubled over the two decades between 1975 and 1995 and then stagnated for about 15 years. In the recent house price boom starting in 2010, house prices more than doubled over a single decade. This is an exceptional event in the history of German house prices. Note that, when looking at the official price series, one might come to a different conclusion, namely that the current house price growth is merely a return to long-run trend growth and not an exceptional boom. However, the consistent series for 28 cities in the highest growth region over the course of 45 years brings us to a different conclusion.

The right panel of Figure 3 shows the long-run rent data from bulwiengesa AG (for newly built and existing apartments) which coincides with the detailed data we have available for our 28 sample cities since 2005. The green line in this figure is the OECD's rent price index, where we normalized prices to the same level in 2005. The two rent series historically show quite some similarity, albeit the fact that bulwiengesa uses transaction data while the OECD calculates rents based on the rent component of the CPI. Only in the

Figure 3: Long house price and rent data



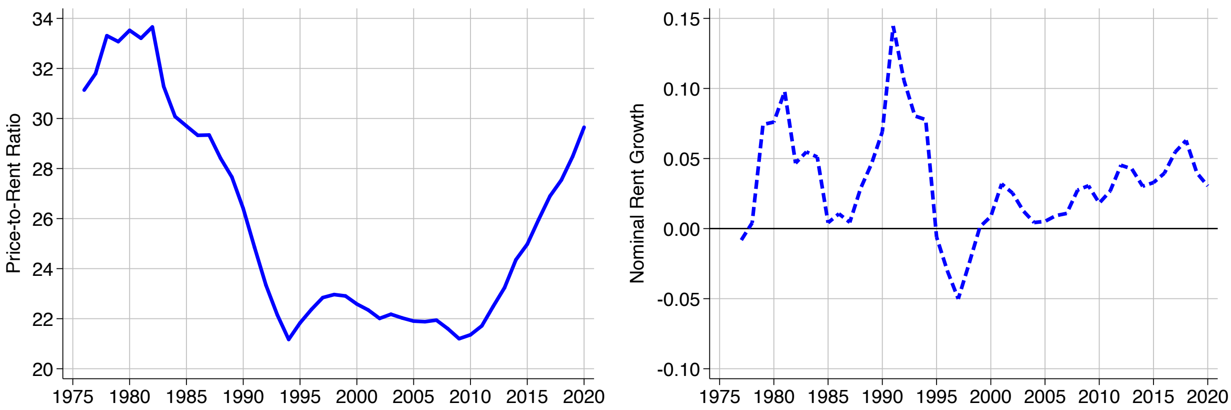
(a) House price (b) Rent

Source: Own calculations based on data from bulwiengesa AG, PHF and OECD.

current boom do we see a strong divergence between the series, which is quite consistent with what we observe for prices. The reason is that our 28 cities are a select sample of what happened in Germany as a whole. However, they are a consistent sample over the course of 45 years, allowing us to analyze the strength of the German house price boom.

Figure 4 shows the house price-rent ratio (left panel) as well as rent growth (right panel) calculated from the price and rent series of the 28 cities presented above. During the inflation of the 1970s and '80s, the price-rent ratio was extraordinarily high (as in many other countries, see Figure 2 in Piazzesi and Schneider (2008)). Starting in the mid 1980s, the price-rent ratio began to fall continuously, owing to the poor performance in house price growth. This long episode of a falling price-rent ratio is only briefly interrupted by some turbulence in the rental market after German reunification in the 1990s. The price-rent ratio reached its trough at around 21.5 in 2010 and then increased to almost 30 in 2020, about a 40% increase. While extraordinary in speed, the 2010s boom in house prices still represents a recovery from a slump in the price-rent ratio since the Great Inflation. Rent growth, in the right panel of Figure 4, spiked once in the 1980s and then again quickly after reunification, as migrants from the East pushed into West German cities. This last episode of high rent growth, however, was accompanied by a "correction" period of negative nominal rent growth in the late 1990s. After the financial crises, rents again picked up growth, much earlier than house prices.

Figure 4: House Price-Rent Ratio and Rent Growth in Germany



(a) House price-rent ratio

(b) Rent growth

Source: Own calculations based on data from bulwiengesa AG and PHF.

3 Household forecasts of house price growth

Our main source for households' price expectations is the *Panel on Household Finances (PHF)*, a representative survey of German households conducted every three years by the Deutsche Bundesbank. The survey collects data on household portfolios, consumption and income; the extent of detail on financial decisions resembles the US Federal Reserve Board's Survey of Consumer Finances. We mostly focus on the second survey wave which sampled 4,168 households in early 2014. We also perform some robustness checks drawing on 2017 data.

Our analysis makes use of two key strengths of the PHF. First, recent waves of the survey ask not only about actual decisions, but also contain a large number of questions on expectations of asset prices, as well as on households' attitudes and investment plans. Second, we can study expectations by region. Indeed, the survey question at the center of our analysis asks households about *regional* house price growth over the next 12 months. We have access to a restricted version of the dataset that allows us to match households to the counties they live in, and hence the growth regions defined in Section 2.

Beliefs about regional house prices are elicited via a two-part question. First, respondents are asked to give a qualitative view about the direction of the housing market (question `dhni0900`):

> *What do you think, how will real estate prices in your area change in the next twelve months?*

There are six candidate answers: (i) increase significantly, (ii) increase somewhat, (iii) stay about the same, (iv) decrease somewhat, (v) decrease significantly, and (vi) don't know/no answer. Second, respondents who give a positive or negative direction, that is, they answer (i), (ii), (iv) or (v) are then asked the follow-up quantitative question (`dhni0950`):

> *What do you think, by what percentage will real estate prices rise / fall in your area over the next 12 months?*

We focus on households who have an opinion about the housing market and hence drop everyone who responds "don't know" to the first question; we come back to these households when discussing information sets in Section 4 below. Our quantitative measure of price expectations then codes response (iii) as zero and uses the quantitative answer for the second question for all other households. After dropping the bottom and top 1 percent of house price forecasts from the distribution in order to guard against outliers, we are left with a total of 3,647 observations.

3.1 The cross section of household forecasts

In this section, we run regressions of household forecasts on household characteristics. The idea is to find out whether there are systematic differences of opinion between households along observable dimensions. We start with simple linear regressions and then proceed to nonlinear specifications. The dependent variable in all specifications is our quantitative measure of one-year-ahead expected house price growth in the region where the household lives. Average expected price growth among households in early 2014 was 3.1%, substantially below the subsequently realized growth rate of 5%.

Our choice of regressors fall into three broad categories. First, we consider characteristics that typically serve as state variables in life cycle models of household behavior, in particular, age, wealth, income and housing tenure. Second, the PHF asks questions that elicit risk aversion, financial literacy and patience. In economic models, these behavioral traits would typically correspond to features of preferences. Finally, to account for heterogeneity in local housing markets, we include variables that capture geography and house quality. Appendix B describes all variables used in regression tables and figures in detail.

Table 1 reports results from linear regressions. Different columns correspond to different sets of regressors. To save space, we list regressors here only if their coefficients are significantly different from zero at the one percent level in at least one specification. Full results, together with clustered standard errors, are in Appendix C. Column (1) shows that typical drivers of savings and portfolio choice – age, income and wealth – are weakly correlated with household expectations. The regressors are dummies indicating ten year

Table 1: The cross section of house price growth forecasts in 2014

	(1)	(2)	(3)	(4)	(5)
Demographics, Income, Wealth					
Age Group 30–39	1.365**	1.315**	0.957	0.711	0.371
	(0.591)	(0.587)	(0.578)	(0.566)	(0.582)
1st Net Wealth Quartile	1.903***	1.906***	0.132	0.024	–0.274
	(0.494)	(0.503)	(0.567)	(0.552)	(0.553)
2nd Net Wealth Quartile	0.939**	0.980**	–0.197	–0.149	–0.159
	(0.448)	(0.441)	(0.467)	(0.454)	(0.459)
Behavioral Traits		yes	yes	yes	yes
Tenure					
Renter			2.438***	2.342***	2.087***
			(0.382)	(0.371)	(0.378)
Growth Region					
Low				–2.063***	–1.559***
				(0.439)	(0.427)
Medium Low				–1.578***	–1.450***
				(0.433)	(0.434)
High				1.368***	1.019**
				(0.416)	(0.399)
Housing/Regional Characteristics					
City Center $\geq$ 500k Inh.					1.762**
					(0.720)
Sqm size/100					–1.619***
					(0.617)
(Sqm size/100)^2					0.423***
					(0.127)
Number of Cases	3647	3646	3646	3646	3598
R^2	0.038	0.042	0.063	0.119	0.142

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

age bins for the age of the household head, income quartiles and net worth quartiles, as well as the number of household members and whether the household head is college educated. The main finding is that young (below age 40) and poor (in the bottom 25% by net worth) households expect about 1.5-2 percentage point (abbreviated pp below) higher price growth compared to older and richer households.

Column (2) introduces three behavioral traits. The PHF measures self-assessed *risk aversion* by asking households to answer "Are you in general a risk-taking person or do you try to avoid risks?" on an eleven point scale from 0 ("not at all ready to take risks") to 10 ("very willing to take risks"). Similarly, *patience* is the response to "Are you in general a person who is patient or do you tend to be impatient?" again on an eleven point scale from “very patient” to “very impatient”. Finally, financial literacy is assessed by asking households three common questions that test their understanding of compound interest, nominal versus real rates of return, and portfolio diversification. We aggregate answers into a score equal to the number of correct answers, from zero to three. In a linear regression, adding behavioral traits marginally improves explanatory power, but otherwise does not change coefficients.

Column (3) introduces housing tenure, which is a strong predictor of price growth expectations. Indeed, households who rent the property they live in forecast almost 2.5pp higher price growth than homeowners. Moreover, much of the explanatory power from other demographics in columns (1) and (2) was due to the fact that young and poor households are more likely to rent – coefficients on age and net worth in column (3) are smaller in magnitude and lose significance. In fact, if we regress respondents' price growth forecasts on tenure alone, excluding all other variables, we also find a highly significant difference of about 2.5pp, as well as an R^2 of 0.05, close to the .063 in column (3). We conclude that housing tenure is a sufficient statistic for forecasts, given household characteristics like age, income, wealth, and behavioral traits.

A simple candidate explanation for this result is composition. Suppose that renters live predominantly in larger cities that also experience higher price growth. Suppose further that all residents of a city share the same opinion about local markets. Tenure might then appear like a good predictor of forecasts, because it is the best proxy for region, better than say, age or wealth. It is therefore important to control for region in order to understand the effect of tenure. An analogous composition effect might arise *within* regions. We know that housing booms often feature heterogeneity of capital gains by market segment – for example, lower quality houses, might become relatively scarce and increase more in price than the top end of the market. If those properties are in areas with more renters, we might naturally see higher price forecasts by renters. This effect calls for another set of controls that capture house quality.

Column (4) shows that tenure generates large significant differences in forecasts even controlling for region. We include dummies for three of our four growth regions, with the medium high growth region as base level. We do find that price growth forecasts align with local housing market conditions: residents of regions that have seen higher growth also forecast higher growth. The difference in average forecasts between the lowest and highest growth region is 3.4pp, about half the 6.25pp difference in realized growth rates between these regions over the year 2013, right before the survey was taken. Regional variation is important for explaining variation in forecasts: the R^2 now increases to 0.12. At the same time, the coefficient on tenure barely changes: unlike age and wealth, tenure is not a proxy for region. Appendix C shows similar results for a specification with more granular controls for the household's past house price growth experience.

Column (5) shows that the result persists when we add additional regional and housing characteristics. In particular, we control for the time span the household already lived in its current residence, the community size on a 10 point scale, whether the household lives in a city center or in the periphery, whether the building the household lives in needs renovation and a general rating of the dwelling quality on different levels in between "Very Simple" and "Exclusive". The one variable that plays a sizeable role in accounting for forecast variation is whether the household lives at the center of a very large city. Large city dwellers forecast higher price growth, and since they often rent, this reduces the coefficient on tenure slightly, to just above 2pp.

We also have data on the square meter size of the house or apartment, and column (5) includes both size and its squared value to accommodate possible nonlinear effects. While both coefficients are significant, the overall relevance of size if minor. The average size of a residence in our sample is around 100 square meters with most variation between 50 and 200 square meters. The coefficients thus imply a significant negative relationship between the size of the household's residence and house price forecasts. At the same time, size adds little explanatory power. If we include only square meter size and its squared value into the regression on top of the regressors in column (4) the R^2 only increases to 0.126, indicating that the predictive power of the size of the residence is limited.

3.2 The role of tenure for expectations

The previous section has shown that, in a linear regression setting, only two variables are relevant for explaining households' forecasts: tenure and location. To understand further the economic mechanism behind the role of tenure, this section goes beyond linear regressions. We report specifications that interact tenure with household characteristics, in order to understand whether the difference between renters and owners is driven by particular subgroups of households.

Table 3 again reports regressions of forecasts on predictors. Column (1) reproduces the last column of Table 2 that includes tenure, growth region as well as many other controls. These variables enter all specifications in Table 3, too. In addition, columns (2)-(4) interact tenure with age, risk aversion and financial literacy, respectively. In each case, we divide the characteristic into two bins and define dummies for one of them. We have run similar regressions with finer bins as well as with all the other characteristics included in Table 2 above. The three variables shown here are the ones for which we obtained quantitatively large, but not necessarily significant, coefficients.

Column (2) shows that the forecast differences between owners and renters are driven by mostly young and middle aged households, but are much weaker for households over 70.

The baseline renter below 70 forecasts 2.3pp higher house price growth than an owner under 70. In contrast, a renter above 70 predicts only about one percent higher growth than an owner over 70. Results on age are interesting also because there is existing evidence that forecasts – for example of inflation – are systematically related to experience. We already know from Table 2 that age by itself does not play an important role. A new point here is that there is a systematic difference between young and old renters; the latter forecast one 1pp lower growth. At the same time, we do not see a significant difference for owners. In light of these findings, we do not pursue the role of age as a predictor in our model below.

Columns (3) and (4) consider the role of risk aversion. We first consider the generic measure of self-assessed risk attitude introduced above (in Section 3.1). The finding here is that forecast differences between renters and owners are less pronounced among households with below median risk aversion. Indeed, the baseline in column (3) is an owner with above median risk aversion, and a renter with above median risk aversion forecasts 2.8pp higher price growth. In contrast, the difference between owners and renters with below median risk aversion is only about 1pp. A possible explanation is that forecasts are guided by fear: renters worry about higher house prices whereas owners worry about lower prices. As a result, renters respond more to news about high prices that are bad for them, which leads to high forecasts, whereas owners incorporate gloomy information into gloomy forecasts. If the effect is moreover stronger for more risk averse households, this is consistent with a larger gap between renters and owner for that group.

To further investigate the "fear hypothesis", we consider a second measure of risk aversion that more directly asks households about the risk-return tradeoff in an investment context: *If savings or investment decisions are made in your household: Which of the statements on list 5.9 best describes the attitude toward risk? Try to characterize the household as a whole, even if it is not always easy.* Households are asked to select one out of the five statements: "We take significant risks and want to generate high returns.", "We take above-average risks and want to generate above-average returns", "We take average risks and want to generate average returns", "We are not ready to take any financial risks". We again split households into two bins at the median.

Column (4) reruns the regression with this second measure, labeled "Financial risk aversion". Coefficients are now small and not significant. The two measures thus appear to measure different concepts. The result is puzzling since the coefficient on the first, generic risk aversion measure is significant only for owners, so the role of risk aversion appears to be more relevant for owners, who make investment decisions – not simple goods purchase decisions – in the housing market. One would thus expect to find some effect also for considerations of the risk-return tradeoff picked up by the second measure. Relatedly, we show in Section 4.3 below that owners' price growth forecasts do not depend on the

Table 2: Interactions with Age, Risk Aversion, and Financial Literacy

	(1)	(2)	(3)	(4)	(5)	(6)
Tenure						
Renter	2.087***	2.372***	2.800***	2.212***	2.153***	3.277***
	(0.378)	(0.402)	(0.448)	(0.470)	(0.371)	(0.543)
Tenure × Age						
Renter × ≥ 70		−0.988				−1.139
		(0.632)				(0.637)
Owner × ≥ 70		0.275				0.292
		(0.453)				(0.453)
Tenure × Risk Aversion						
Renter × Below Median			−0.820			−0.905
			(0.501)			(0.503)
Owner × Below Median			0.961***			1.106***
			(0.297)			(0.309)
Tenure × Financial Risk Aversion						
Renter × Below Median				−0.352		−0.183
				(0.511)		(0.502)
Owner × Below Median				−0.074		−0.262
				(0.277)		(0.284)
Tenure × Financial Literacy						
Renter × Very Low					0.766	0.782
					(1.179)	(1.161)
Owner × Very Low					3.829	3.910
					(2.741)	(2.753)
Growth Region						
Low	−1.559***	−1.558***	−1.556***	−1.571***	−1.538***	−1.536***
	(0.427)	(0.426)	(0.427)	(0.427)	(0.421)	(0.419)
Medium Low	−1.450***	−1.437***	−1.470***	−1.442***	−1.423***	−1.433***
	(0.434)	(0.433)	(0.432)	(0.434)	(0.432)	(0.430)
High	1.019**	1.001**	1.012**	1.000**	1.044***	1.022***
	(0.399)	(0.399)	(0.395)	(0.400)	(0.398)	(0.395)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of Cases	3598	3598	3598	3594	3598	3594
R-Square	0.142	0.143	0.147	0.142	0.143	0.152

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

time owners plan to remain in their current residence; if answers were guided by fear, then owners who plan to sell sooner should worry more about low price growth. In sum, we note that while there is some interesting interaction between risk attitude and tenure, the results are not strong enough to assign a special role for risk attitude in our modeling exercise.

Column (5) shows that forecast differences between owners and renters are driven by households who are reasonably financially literate. We observe a stark discrepancy in forecasts between households who did not answer any of the three test questions correctly, labeled "very low literacy".[6] Illiterate owners have average price growth forecasts that are 3.9pp higher than for other owners, while illiterate renters have forecasts that are only 0.9pp lower than illiterate owners. Financially illiterate households as a whole thus have unusually high growth expectations regardless of tenure. This result suggests that the difference between owners and renters cannot be attributed entirely to unsophisticated reasoning. Instead, the mechanism behind it must also apply to the most literate households.

3.3 The distribution of forecasts by growth region

In this section, we take a closer look at distributions – mean and dispersion – of renter and owner forecasts by growth region. The figures here interact the key predictor of forecasts established in the previous sections. They also provide us with targets that our quantitative model below will be required to match.

Figure 5 shows mean forecasts by region and how they compare to realized price growth. The wide red bars in the left panel are average renter forecasts in the four regions; whiskers indicate the 95 percent confidence interval. Narrow bars represent realized price growth in the respective growth region. Any differences in the height of bars for a region reflect forecast errors made by the average renter in 2014. The right panel repeats the exercise for owners, whose average forecasts are wide yellow bars; the narrow bars represent the same realizations as in the left panel.

The figure summarizes three robust patterns. First, households generally underpredict house price growth. With the exception of renters in the lowest growth region, all price growth forecasts lie below realized price growth, and most of them are significantly different from realized growth. Second, households' forecasts are consistent with regional differences in the sense that forecasts in low growth regions are lower than those in high growth regions, regardless of tenure status. Third, within each of the different growth regions, we find that renters make higher forecasts than owners.

Figure 6 considers the cross sectional mean squared error in households' forecasts. Again the left panel shows renters and the right panel shows owners, each with four bars for the four growth regions. An individual household's forecast error is defined as the squared difference between realized growth – common to all individuals in the region – and the individual forecast. The mean squared forecast error for a group of households can there-

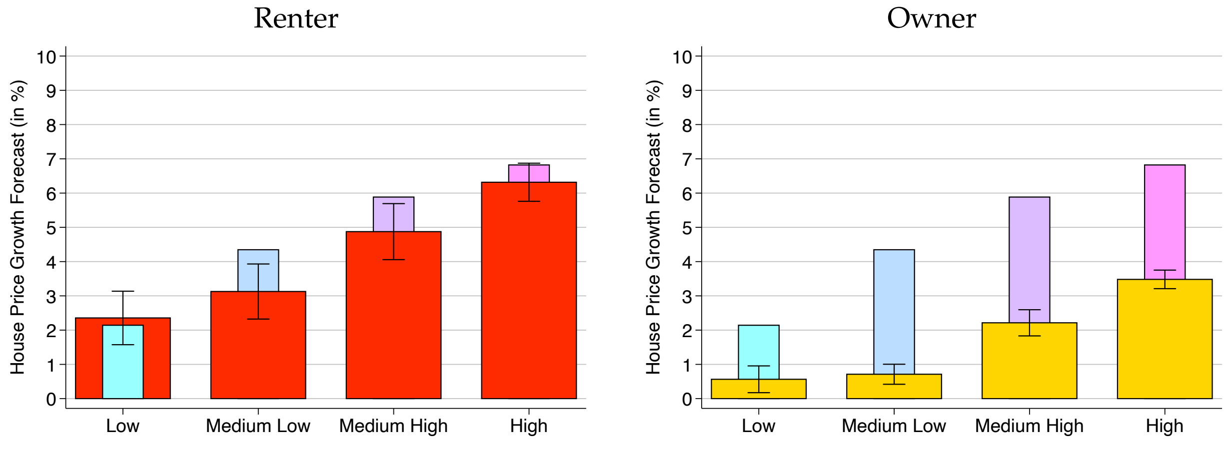
[6] Note that such households only constitute 3.9 percent of our total sample.

fore be decomposed into two parts: the squared average forecast error – indicated by light colors in the figure – and the (cross sectional) variance of the forecasts. The squared average forecast error reflects the mistake made by the average owner or renter, as shown already in Figure 5. The variances reflect differences of opinions within the groups of renters and owners.

In all growth regions, renters exhibit a larger mean squared forecast error than owners. The result is entirely driven by the wide dispersion in renters' individual forecasts. As we have seen above, renters' average forecast is in fact closer to the actual growth realization everywhere. At first sight, the finding might speak against an informational explanation for the forecast differences – indeed, in a simple model where agents predict an unknown parameter from noisy signals, the unconditional mean squared error of a better-informed agent (that is, an agent with a more precise signal) is always below that of a less-informed agent. However, the result here is about errors conditional on a particular realization. Renters, even though their information is more noisy, can thus be "in the right place at the right time" during a boom that featured strong rent growth. Our model below formalizes this point and shows how learning from different information can jointly rationalize means and dispersions of forecasts.

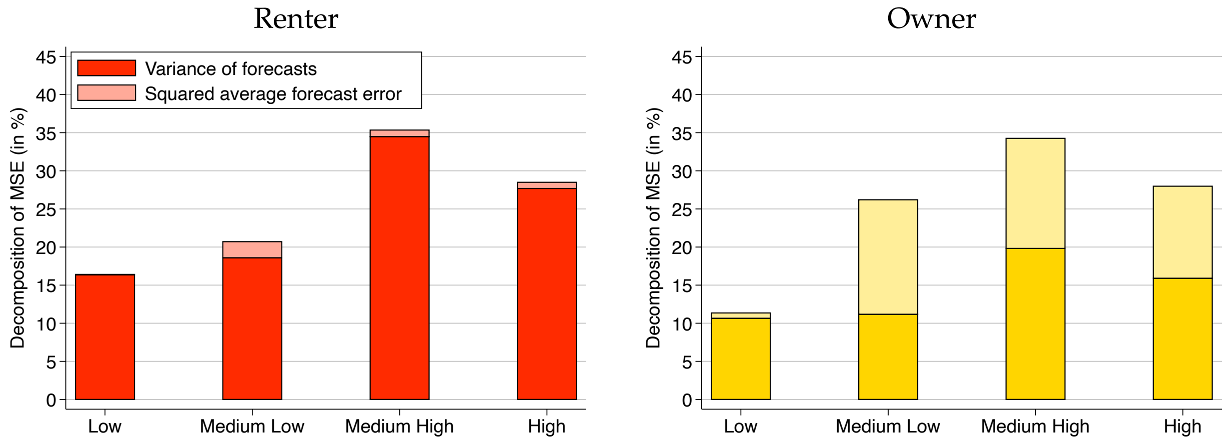
As a final point, we describe how we aggregate findings by region into single numbers for Germany as whole. We find this useful because the patterns on differences between renters and owners we have shown in this section – as well as others that follow below – are qualitatively very similar across growth regions. We can thus streamline the exposition by presenting summary numbers at the national level, rather than always showing each region separately. However, aggregation must take into account that forecasts reflect regional growth, and renters and owners are not equally distributed across Germany.

Figure 5: House Price Growth Forecasts by Tenure and Growth Region in 2014



Source: Own calculations based on data from bulwiengesa AG and PHF.

Figure 6: Cross sectional MSE of House Price Growth Forecasts in 2014



Source: Own calculations based on data from bulwiengesa AG and PHF.

The first two columns of Table 3 show the distribution of households across growth regions by tenure type using our sample weights: renters are relatively more likely to live in cities where house price growth has been high and that therefore belong to the higher growth regions. Even if owners and renters in each region made the exact same forecasts, but differed by region, a simple average would therefore show relatively higher forecasts from renters. We do not want this composition effect to inflate national level forecast differences.

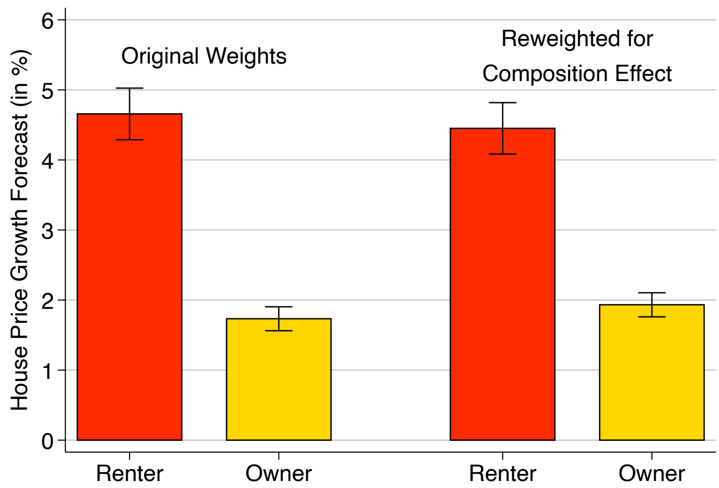
Table 3: Distribution across growth regions by tenure type

	PHF Sample Weights			*Reweighted for Composition Effect*	
Growth Region	Renter	Owner	Total	Renter	Owner
Low Growth	16.52	21.34	18.84	18.84	18.84
Medium Low Growth	20.42	27.62	23.88	23.88	23.88
Medium High Growth	25.17	29.07	27.04	27.04	27.04
High Growth	37.89	21.97	30.24	30.24	30.24
Sample Share	51.95	48.05	100.00	51.95	48.05

We aggregate forecasts across growth regions by reweighting: we scale the sample weights for households in a tenure cell and growth region so that the distribution of households across growth regions becomes the same for renters and owners. In particular, both distributions become equal to the distribution of all households across regions shown in the third column; since the construction of regions did not weigh counties by populations, the high growth region that contains larger cities is more populated. To get from the original to the reweighted distribution, involves, for example, reducing (increasing) the sample weight of renters (owners) in the top region.

The reweighted distribution does not generate the above composition effect: if all renters and owners in each region made the same forecast but differed by region, then the average forecast for Germany would also be equal. Figure 7 reports the full sample averages of house price growth forecasts by tenure status using both the original sample weights as well as the weights that control for household composition. We find a relatively small difference between the two weighting schemes; nevertheless we employ our reweighting scheme in what follows to guard against composition effects.

Figure 7: House Price Growth Forecasts by Tenure



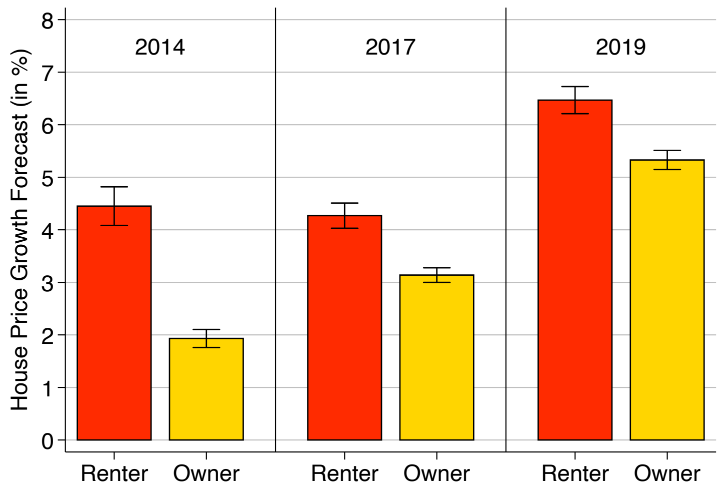
Source: Own calculations based on data from PHF.

3.4 Expectations of renters and owners over time

So far, we only looked at results from wave 2 of the PHF in 2014. Yet, there are additional data available that allow us to track the differences between households of different housing tenure over time. Wave 3 of the PHF asked households the same forecasting questions in the year 2017, where the house price boom had already arrived in all German regions. In addition, we can draw on data from the *Bundesbank Online Survey - Households* (BOP-HH), a pilot survey that was initiated by the Deutsche Bundesbank in 2019. This representative survey again asks respondents about their forecasts of regional price growth, collects their tenure status and can be matched to our local house price growth database.[7]

Figure 8 shows the average house price growth forecasts of renters and owners over time. The first thing we see is that price forecasts increase especially in the group of owners, but also in the group of renters. As house prices grow for an extended period of time, households seem to adapt their expectations about the future accordingly. While the gap between forecasts of renters and owners narrows a bit between 2014 and 2017, there is still a sizable difference of more than one percentage point left in 2017 and 2019. Hence, the fact that renters make significantly higher price growth forecasts than owners persists over time and across different surveys. In Appendix C we provide additional data and sensitivity checks for this result. Most importantly, we clarify that the forecast difference between renters and owners is not driven by few extreme observations. Summing up, this section leaves us with the robust stylized fact that across regions of different house price growth and across time, renters make significantly higher price growth forecasts than owners.

Figure 8: House Price Growth Forecasts by Tenure Over Time



Source: Own calculations based on data from PHF and BOP-HH.

4 Learning about housing cost: direct evidence

The facts presented in the previous sections lead us to explore differences in information sets as a possible explanation. In particular, our theory postulates that renters have better

[7]The survey has much less details compared to the PHF when it comes to household characteristics, income and wealth. The questions about house price expectations, however, were framed in exactly the same way. We process the data so that it is comparable to the PHF. Details can be found in Appendix C.

information about housing dividends than owners. We now use additional Bundesbank data to provide more direct evidence on information sets. These facts guide specific assumptions we make in our model below. We proceed in five steps. Section 4.1 shows that housing differs from equity in the ability of non-owners of the asset to form an opinion about its price. Section 4.2 provides evidence that owners are well aware of current house price growth, but with large confidence bounds. Section 4.3 shows that plans to trade do not matter for differences of opinion between owners and renters; their forecasts do not depend on whether they plan to buy or sell in the future, or whether they plan to move. Because of this evidence, our model does not feature endogenous information acquisition. Section 4.4 provides evidence that real estate investors make forecasts that are comparable to renter forecasts. Finally, section 4.5 provides evidence on how households get their information about housing, emphasizing the role of direct price and rent observations.

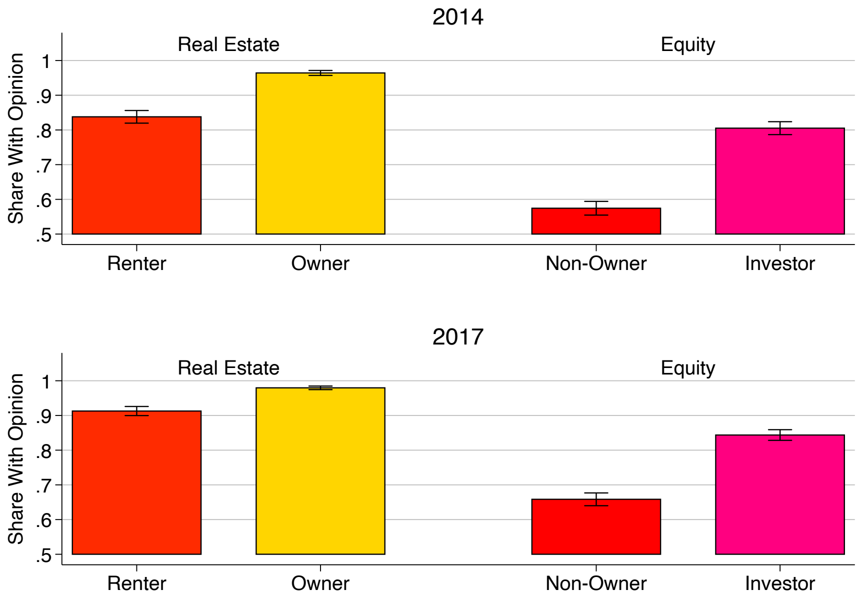
4.1 Opinions about future prices: real estate vs. equity

The premise of our theory is that the housing market is special among asset markets in that even agents who do not participate – that is, renters – have information about dividends. We now show that this feature indeed differentiates housing from the other major long term assets in modern economies, equity. We make use of the two part structure of the PHF expectations questions: for the question on housing described in Section 3, the first part gives households the option of responding "don't know" if they do not want to voice any opinion on the direction of the housing market. An analogous question is available for the stock market.

Housing is a special asset with regard to the share of non-owners who feel confident forming an opinion about price movements. In fact, the vast majority of renters can make a forecast of future house prices. This stands in stark contrast to equity, where we see that a large fraction of non-owners is not capable of making stock price forecasts, which suggests that obtaining signals about house prices is relatively cheap for renters. Figure 9 shows the fraction of opinionated households – who do not answer "don't know" by participation status for both asset markets. The top panel provides results based on the 2014 survey, while the bottom panel uses the 2017 survey. For equity, "investors" comprise not only households who directly invest in stocks but also those who invest only indirectly via mutual funds or pension funds.

In 2014, the overwhelming majority of the survey population – on average about 90 percent – is willing to provide a qualitative forecast of future house prices. This number is substantially larger than the 70% who opine on equity. The difference is particularly striking among non-participants: about 85% of renters have developed a view about the housing market, whereas less than 60% of households who do not hold stocks have a view about the stock market. The bottom panel of Figure 9 shows that the same pattern was present in 2017, although the differences between non-owners and owners shrink. Like most countries, Germany has been experiencing an extended period of zero interest rates. Being traditionally mostly active in the risk-free savings market, German households now have to search for investment alternatives, which raises the incentives for non-owners of real estate or equity to familiarize with those assets.

Figure 9: Opinion formation about prices: real estate vs. equity



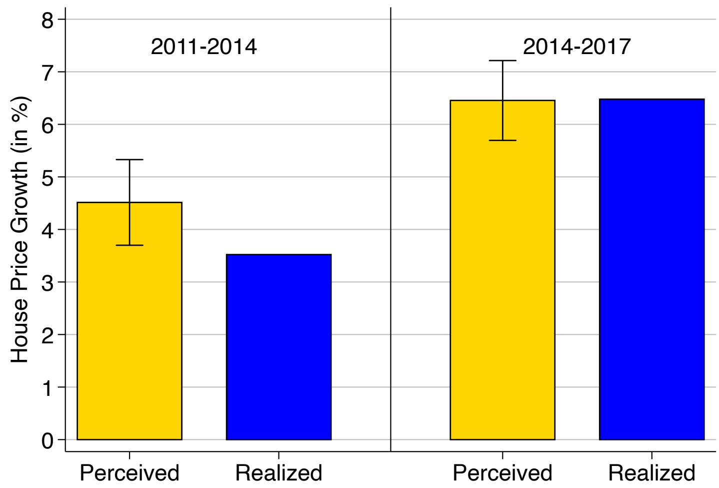
Source: Own calculations based on data from PHF.

4.2 Do owners keep track of house prices?

A possible explanation for owners' forecasting mistakes during the boom is that they simply do not pay attention to prices. We can shed some light on this possibility by checking owners' perceptions of the value of their own residence over time. Indeed, a subset of the PHF survey is organized as a panel, following households across waves from 2011 to 2014 and again from 2014 to 2017. In each wave, households are asked to estimate the hypothetical sales price of their current main residence. We match those price estimates for all panel households who did not move between two survey waves. This allows us to calculate a *perceived home price growth rate* for all owners over both the years 2011-2014 and 2014-2017.

Figure 10 compares annualized mean perceived home price growth with realized growth

Figure 10: Information quality of opinionated owners



Source: Own calculations based on data from PHF.

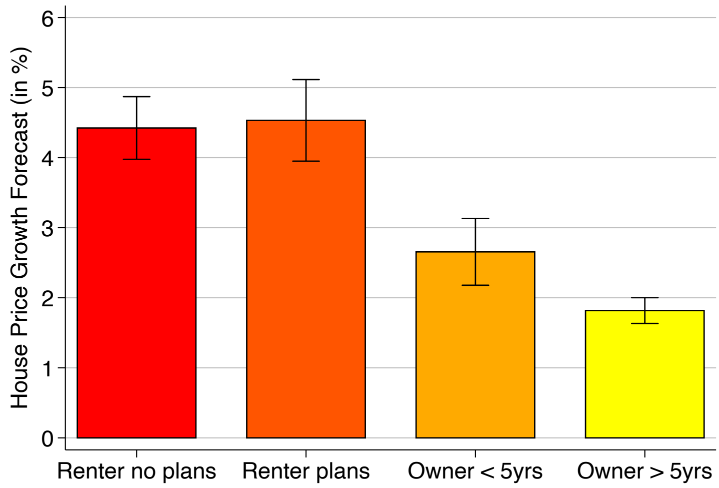
rates in the owner's region over the same time span. The left panel shows 2011-14 and the right panel shows 2014-17; in both cases, whiskers on the left hand yellow bars indicate 95% confidence intervals for the mean perceived home price growth rate. Confidence intervals are large, consistent with imperfect information on the part of owners. At the same time, for both time spans, the realized growth rate lies within or at least very close to the confidence band of the mean perceived growth rate. We thus take away that the *average* owner is aware of local housing market conditions, and that an explanation for differences between owners and renters should not rely on owners simply receiving no market signals.

4.3 Forecasts and plans to buy, sell or move

Our theory assumes that renters obtain better information about rents because they observe their own rent as well as that of related properties, and use this information to reason about prices. An interesting alternative hypothesis is that households pay attention to prices in an illiquid market only rarely, namely when they buy or sell. If this were the case, there would be nothing special about tenure per se, but tenure would be an imperfect proxy for incentives to trade.

The PHF includes three questions that speak to this hypothesis. First, it asks renters *"Do you intend to buy or build a house or flat for your own accommodation?"*. Second, owners are asked about the date they moved into their current residence. Third, in the 2017 wave both renters and owners were asked how long they plan to remain in their current residence. If information improves when households trade, we should expect more accurate forecasts from renters who plan to buy, from owners who have recently bought, as well as from owners or renters who plan to move in the near future.

Figure 11: Price forecasts and incentives



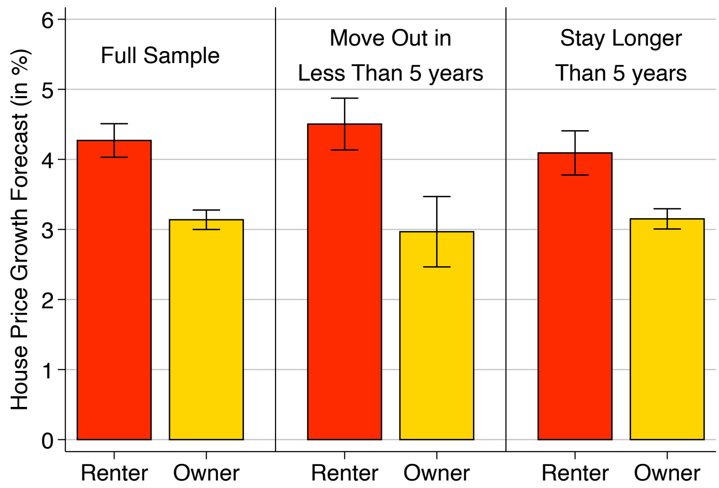
Source: Own calculations based on data from PHF.

Figure 11 shows average price forecasts for four types of households: renters with and without plans to buy, as well as owners who moved less than or more than five years ago. The differences among the two renter types are negligible: the high accurate forecast we observe by the average renter is not driven by renters who are planning to buy. For owners, there is a small and borderline significant difference: recent owners make slightly higher forecasts. However, the difference between owner types is small relative to the overall difference between owners and renters.

Figure 12 compares average forecasts for owners and renters in the full sample to average forecasts for those households who plan to move within five years as well as those who plan to remain in their current residence longer. Among renters, there is some qualitative support for the idea that agents who plan to shop for a new place soon make higher (and hence more accurate) forecasts. However, for both renters and owners, differences across groups are small and insignificant. Overall, we take away that incentives to trade are not a major factor in driving price forecasts. Our model thus focuses on information advantages that simply reflect tenure, and does not feature endogenous information acquisition.[8]

[8] The results of Figure 12 also speak to the hypothesis that households report forecasts of what they fear. If this were the case, we would expect renters who plan to move soon to make higher price growth forecasts, as they fear price or rent increases, whereas owners who plan to move soon should forecast particularly low prices. We do not see relevant differences in the figure.

Figure 12: Forecasts and expected time to move



Source: Own calculations based on data from PHF.

4.4 Forecasts by real estate investors

Our model assumes that renters receive better information about rents than owners. Now, households who own their main residence may also own other real estate that they rent out. While these households are owners, they have more information about rents than other owners, similar to renters. From the perspective of our model, the forecasts for these owner-landlords should be higher than owner forecasts and comparable to renter forecasts. In this subsection, we present evidence that supports these predictions from the model.

It is not obvious how to best identify owner-landlords in the survey. Some households who own their main residence also own a vacation home that they rent out occasionally. Other households own property that they leave vacant, or they only rent out to family members at below market rents. We therefore define owner-landlords to be those households who own their primary residence, own at least one additional property (apartment, single-family home, or townhouse), and receive some minimal rental income (more than 650 Euros per month). Below, we use alternative definitions that try to identify more serious real estate investors.

Column (1) in Table 4 repeats our baseline finding from Table 1 about the importance of tenure and region for the cross section of forecasts. To save space, we omit the estimated coefficients on the growth region dummies in Table 4. Column (2) in Table 4 shows that owner-landlords indeed make higher house price growth forecasts (by 0.42 percentage points) than other owners. However, the coefficient is not precisely estimated. One issue is that there are not many households who own their main residence and also other real estate. In the survey, only 12 percent of households are owner-landlords.

To identify serious real estate investors, we consider the following alternatives. Column (3) defines investors as owner-landlords who have an above median loan-to-value ratio on their investment property. The estimated coefficient is comparable to the coefficient on the renter dummy, but again standard errors are high. We find that for other owners, a high loan-to-value ratio leads to a negative coefficient, which is also imprecisely estimated. Column (4) shows that owner-landlords who have more than one rental property also make higher forecasts. Columns (5), (6) and (7) isolate owner-landlords for whom the rental income is an increasing share of total household income, ranging from at least 30, 40 to 50 percent. Investors with higher rental income shares make higher forecasts. Moreover, the estimated coefficients for investors with a rental income share above 40 percent are comparable to those of renters, and also significant. We conclude that forecasts for serious real estate investors are quantitatively in the same ballpark as renter forecasts.

Table 4: Forecasts by real estate investors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Tenure							
Renter	2.087***	2.107***	2.125***	2.109***	2.099***	2.097***	2.095***
	(0.378)	(0.380)	(0.381)	(0.380)	(0.381)	(0.380)	(0.380)
Is Owner-Landlord							
Yes		0.422	0.021	0.115	0.092	−0.053	−0.256
		(0.523)	(0.585)	(0.533)	(0.403)	(0.386)	(0.415)
Is Owner-Landlord × *Has above median LTV on other property*							
Yes × Yes			1.893				
			(1.253)				
Is Owner-Landlord × *Has more than one rental property*							
Yes × Yes				0.958			
				(1.214)			
Is Owner-Landlord × *Rental income share*							
Yes × ≥ 0.3					1.096		
					(1.410)		
Yes × ≥ 0.4						1.963	
						(1.667)	
Yes × ≥ 0.5							3.586**
							(1.693)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Cases	3598	3598	3598	3598	3598	3598	3598
R-Square	0.142	0.142	0.142	0.142	0.142	0.143	0.144

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

4.5 Sources of information

To find out more directly how households acquire information about housing markets, we draw on a new question we proposed for the Bundesbank's Online Survey – Households (BOP-HH). The BOP-HH was fielded in spring 2019; it elicits less detail on household income and balance sheet, but instead focuses on expectation formation. Question 306 of the BOP-HH survey reads:

> *How important are each of the following sources of information for you to evaluate future house prices?*

Respondents are presented the following seven potential information sources:

1. Relatives, friends and neighbors
2. Classical media (newspapers, tv, etc.)
3. Social media (like Facebook and Twitter)
4. Online real estate platforms
5. Financial consultants
6. Direct observations of rents in your neighborhood
7. Direct observations of house prices in your neighborhood

For each candidate source, households can check one of four intensities, "Not important at all", "Somewhat important", "Quite important", and "Very important". To summarize average choices by a single numerical score, we code these intensities as 0, 33, 66, and 100, respectively. Table 5 lists the seven answers in order of importance, measured by the fraction of all respondents who labeled an answer either "Quite important" or "Very important", reported in the first column. The second and third columns report the average score of the information source among renters and owners, respectively, and the final column measures the difference.

The primary source of information for households to forecast house price growth is the direct observation of prices. In fact, more than 80 percent of households look to rents and almost 80 percent to house prices. Online real estate portals, friends, advisors and especially social media are much less important sources of information. The magnitude of the scores for the seven sources are broadly similar across renters and owners. There are however two significant differences. First, while all households rely on direct observations, renters look more at rents, whereas owners look more at prices. Second, renters rely more on social media, online real estate portals and especially family and friends when gathering information. We view the results as broadly supportive of an approach that emphasizes different information sets that relate to current market experience.

Table 5: Sources of information by tenure

Source	Perceived Important by	Coded average Renters	Coded average Owners	Difference
Direct observation of rents	83.34	72.02	69.25	-2.773^{***} (1.061)
Direct observation of prices	78.24	66.04	69.83	3.785^{***} (1.148)
Classical Media	73.53	60.32	60.37	0.057 (0.993)
Online Real Estate Portals	66.76	57.72	55.29	-2.432^{***} (1.108)
Family & Friends	52.50	52.62	48.11	-4.508^{***} (1.101)
Financial Advisors	46.39	44.68	45.47	0.791 (1.158)
Social Media	12.26	23.81	20.86	-2.963^{***} (0.943)

5 A model of learning about housing cost

In this section, we develop a simple model that describes the joint distribution of prices, rents, and household forecasts. Its goal is to show how learning with different information sets can naturally lead to the large differences in forecast distributions that we see in the data. We do not explicitly model household decisions, but only specify the information sets of owners and renters and compute their conditional expectations. We do impose one piece of structure: there is a connection between rents and prices because a developer sector arbitrages between rental and owner occupied houses. Appendix D derives the equations below.

Asset pricing. We consider the valuation of houses in a region. Developers can sell housing units at a price P_t or rent them out at the rental price R_t, both denominated in Euros. Prices and rents should be thought of as regional averages, as in our data presented above. At any date t, developers are indifferent between selling a house at date t today or holding it for one period and receiving rent as "dividend". The equilibrium house price P_t

thus satisfies a standard intertemporal Euler equation

$$P_t = E_t \left[\tilde{M}_{t+1} \left(P_{t+1} + R_{t+1} \right) \right], \tag{1}$$

where E_t is the conditional expectation operator and $\tilde{M}_{t+1}$ is the stochastic discount factor of the developer.

We emphasize that the Euler equation (1) allows for a wide range of frictions and behavior by developers (or investors in developer firms). The special case where $\tilde{M}_{t+1}$ is perfectly foreseen one period in advance is the familiar "user cost model" for frictionless housing markets under rational expectations, $M_t := E_t \left[\tilde{M}_{t+1} \right]$ is the one-period ahead nominal bond price. More generally, the stochastic discount factor may capture (i) risk attitude of developers, (ii) financial frictions that affect the developer's cost of capital, or (iii) differences in beliefs between developers who value houses and outside observers who know the current price and rent. Large literatures have documented the relevance of time variation in $(i) - (iii)$ on many asset prices. We do not take a stand on what force is most important; all that matters below is that households evaluate the Euler equation with expectations E_t conditional on their time-t information, and they contemplate movements in $\tilde{M}_{t+1}$ as a source of price volatility.

To deal with trends, it is helpful to work with rent growth and price-rent ratios. We denote the gross growth rate of rents by $G_t = R_t / R_{t-1}$ and define the price-rent ratio as $V_t = P_t / R_t$. We can then rewrite (1) as

$$V_t = E_t \left[\tilde{M}_{t+1} \left(V_{t+1} + 1 \right) G_{t+1} \right]. \tag{2}$$

We follow standard practice in assuming that $\tilde{M}_t$ and G_t are jointly stationary, and focus on stationary solutions V_t to the difference equation (2). In other words, rents and prices follow a stochastic trend, but they are cointegrated so the price-rent ratio is stationary. Movements in the stochastic discount factor or the growth rate may therefore lead to divergence of price and rent growth in the short or medium run, but not in the very long run.

To implement our learning model, we use a log-linear approximation of (2) around its deterministic steady state. Suppose that the mean stochastic discount factor is M and the mean growth rate is G, with $MG < 1$. The steady state price-rent ratio is then $V = MG / \left(1 - MG\right)$. Denoting logarithms by small letters, the steady state log growth rate is $g = \log G$ and the log price-rent ratio is $v = \log V$. Denoting log deviations from steady state by hats, we obtain the linear difference equation

$$\hat{v}_t = \hat{m}_t + E_t \left[MG \hat{v}_{t+1} + \hat{g}_{t+1} \right], \tag{3}$$

where $m_t = E_t [\tilde{m}_{t+1}]$ is the predictable component of the log discount factor and $\hat{m}_t$ is its deviation from steady state. Appendix D.1 contains a detailed derivation of this equation.

The dynamics of prices and rents. We choose functional forms for the stochastic discount factor and growth rate that capture key features of the data presented in Section 2 but at the same time allow for easy application of Bayes' rule. In particular, deviations from the mean in both the growth rate of rents and the predictable component of the discount factor are described by Gaussian AR(1) processes

$$\begin{aligned} \hat{g}_t &= \alpha_g \hat{g}_{t-1} + \varepsilon_t^g, \\ \hat{m}_t &= \alpha_m \hat{m}_{t-1} + \varepsilon_t^m. \end{aligned} \tag{4}$$

The innovations ε_t^g and ε_t^m are serially as well as mutually uncorrelated and normally distributed with mean zero. Both processes are persistent but stationary, $\alpha_g, \alpha_m < 1$. The distribution of rent growth can be directly estimated from the data. Persistence in the growth rate allows in particular for an acceleration of rent growth that pushes rents to a permanently higher level.

The role of the persistent discount factor shock ε_t^m is to allow for "excess volatility" in house prices: the price-rent ratio can move even if there is no news about current or future rents (that is, no "cash flow news"). To see this, we solve (3) by the method of undetermined coefficients to obtain the stationary solution:

$$\hat{v}_t = \beta_m \hat{m}_t + \beta_g \hat{g}_t, \tag{5}$$

where $\beta_m = 1/(1 - \alpha_m MG)$ and $\beta_g = \alpha_g / \left(1 - \alpha_g MG\right)$ are positive coefficients. Prices are high relative to rents when developers either (i) discount the future at a lower rate or (ii) expect unusually high growth in rents. Since both types of fluctuations are mean-reverting, their impact on prices depends on their persistence relative to the duration of houses, captured by MG.

Putting together the trend and fluctuations around it, we now summarize the joint dynamics of (log) rents and prices in the region by

$$\begin{aligned} p_t &= v + r_t + \hat{v}_t, \\ r_t &= g + r_{t-1} + \hat{g}_t. \end{aligned} \tag{6}$$

Both prices and rents grow on average at the rate g. Movements in $\hat{g}_t$ induce transitory deviations from trends in both variables. In contrast, movements in $\hat{m}_t$ can drive prices to move above or below trend even if rents simply grow at the trend growth rate. The model thus allows for two types of booms discussed in the literature: swings in $\hat{m}_t$ capture changes in interest rates, credit conditions or investor sentiment, whereas swings in $\hat{g}_t$ capture changes in actual rent growth.

To study household learning, it is helpful to have concise vector notation for the dynamics of prices and rents. We thus define a state vector $x_t = (\hat{m}_t, \hat{g}_t, r_t, 1)^\top$ that contains the predictable component of the discount factor as well as the stochastic components of rents and the growth rate. We can then represent the distribution of rents and prices as

$$\begin{pmatrix} p_t \\ r_t \end{pmatrix} = Bx_t; \quad x_t = Ax_{t-1} + C\varepsilon_t, \tag{7}$$

for some matrices A, B and C, where ε_t is a 2×1 vector of iid standard normal innovations see Appendix D for a detailed description of the state space system.

Information structure and forecasts. Consider now owner and renter households who answer survey questions. We want to capture the idea that they sample the rents and prices of a few individual dwellings and hear about others from friends or neighbors. We thus assume that an individual household i who is either an owner (type $h = o$) or a renter (type $h = r$) observes a vector of noisy signals of the current average (log) price and rent in the region

$$s_t^{i,h} = \begin{pmatrix} p_t \\ r_t \end{pmatrix} + w_t^{i,h}, \qquad h = r, o. \tag{8}$$

Here $w_t^{i,h}$ is a 2×1 vector of idiosyncratic Gaussian shocks with mean zero that are mutually uncorrelated and iid in the cross section of individual households. Their covariance matrix $\mathcal{W}^h$ depends on the household type h: for example, owners may receive less precise (or more noisy) signals about rents relative to prices, and vice versa for renters.

We assume that households know the distribution of our average price and rent data (p_t, r_t) and their own signals $s_t^{i,h}$. We identify an individual's survey forecast with the conditional expectation of average price growth given that individual's history of signals and their initial view of the state. Household i believes that the initial state is $x_0^{i,h}$. This initial nowcast is drawn from a normal distribution that depends on the household type h through its mean $\bar{x}_0^h$ (the average initial nowcast of type h agents) and the covariance matrix Ω_0^h (their cross sectional dispersion). The price growth forecast at date t is then

$$f_t^{i,h} = E\left[\Delta p_{t+1} | s_t^{i,h}, s_{t-1}^{i,h}, .. s_1^{i,h}, x_0^{i,h}\right]. \tag{9}$$

Since the data, signals and the initial belief are all jointly normally distributed per (4)-(8), subsequent beliefs are also normally distributed and forecasts can be computed via the Kalman filter.

Given the structure of the system (7), we can choose the variance Σ_0^h of the initial nowcast error $x_0 - x_0^{i,h}$ such that the forecast error variance of type h households is time invariant. Intuitively, this works because households track a persistent hidden state by observing noisy signals. Every signal contains information that lowers uncertainty about the state, but also adds additional noise. When the two forces balance, uncertainty about the state as well as forecast error variances are constant. Standard results further imply that this choice of initial variance is what one would obtain if agents had seen an infinite sequence of past signals. We make this choice throughout in our quantitative application below – it captures in a parsimonious way the idea that agents are uncertain even at the beginning of our sample.

Characterizing the distribution of nowcasts and forecasts. Given an initial cross sectional distribution of beliefs about the state – that is, a cross section of $x_0^{i,h}$s – as well as a realization of the data (p_t, r_t), our model generates a panel of forecasts that we can match to our survey data. The key to understanding the dynamics of forecasts is the evolution of individual "nowcasts" of the current state, denoted $x_t^{i,h} := E\left[x_t | s_t^{i,h}, s_{t-1}^{i,h}, ..s_1^{i,h}, x_0^{i,h}\right]$. Indeed, (7) implies that this nowcast is a sufficient statistic for forecasting future prices given an agent's past information. In particular, up to a constant, the price growth forecast is $f_t^{i,h} = B_{1\bullet}(A - I)x_t^{i,h}$, where $B_{1\bullet}$ is the first row of the matrix B.

The law of motion of the nowcast takes the standard form

$$x_t^{i,h} = Ax_{t-1}^{i,h} + \Gamma^h \left(s_t^{i,h} - BAx_{t-1}^{i,h}\right), \tag{10}$$

where the "gain matrix" Γ^h is constant because of our choice of initial conditions. Nowcasts – and hence forecasts – are updated according to a time invariant rule. To arrive at a nowcast for date t, agents start from the date $t-1$ forecast of the state $Ax_{t-1}^{i,h}$. Upon receiving signals, they make an adjustment depending on the last forecast error, the term in parentheses. We note that errors occur not only because of the new realization of the data (p_t, r_t) that is common to all agents, but also because of the noise in agents' signals.

What is the *average* forecast made by agents of type h? Since forecasts are linear in nowcasts, they depend on the evolution of the average nowcast $\bar{x}_t^h$ of type h agents. By the law of large numbers, the noise in type-h signals washes out in the average, and we obtain a recursion for the average type-h nowcast:

$$\bar{x}_t^h = A\bar{x}_{t-1}^h + \Gamma^h \left(\begin{pmatrix} p_t \\ r_t \end{pmatrix} - BA\bar{x}_{t-1}^h \right). \tag{11}$$

Given an initial average nowcast $\bar{x}_0^h$ for type h, the current average nowcasts – and hence also current rent and price growth forecasts – are deterministic functions of the data (p_t, r_t). This is the relationship we use below to link observed average forecasts to the observed house price and realizations.

What is the *cross sectional dispersion* of nowcasts? Suppose the date $t-1$ cross sectional variance of nowcasts $x_{t-1}^{i,h}$ is Ω_{t-1}^h. From (10), the date t variance is then

$$\Omega_t^h = \left(I - \Gamma^h B\right) A \Omega_{t-1}^h A^\top \left(I - \Gamma^h B\right)^\top + \Gamma^h \mathcal{W}^h \Gamma^{h\top}. \tag{12}$$

The first term reflects the adjustment of nowcasts due to the information conveyed by prices and rents. Since the same data realizations affect all agents, this tends to reduce the dispersion of nowcasts. The second term reflects new date t noise which increases dispersion. We can choose as our initial dispersion Ω_0^h the fixed point of (12) at which the two forces balance. As a result, the dispersion of nowcasts (and hence forecasts) is constant over time.[9] The idea is to parsimoniously capture disagreement about nowcasts among individuals at the beginning of our sample, much like our initial variance for individual beliefs captures uncertainty initially perceived by an individual.

Information in prices and rents. Why do renters make higher price growth forecasts in the current German housing boom? To see how learning accounts for this fact, consider price growth under the subjective belief of household i of type h. We can decompose it into a forecast – based on elements of the nowcast vector $x_t^{i,h}$ – as well as an orthogonal forecast error

$$\Delta p_{t+1} = g + \frac{\alpha_m - 1}{1 - \alpha_m MG} \hat{m}_t^{i,h} + \frac{(\alpha_g)^2 (1 - MG)}{1 - \alpha_g MG} \hat{g}_t^{i,h} + u_{t+1}^{i,h}.$$

Here the forecast error reflects both the nowcast error due to imperfect learning up to date t and the new innovations that affect actual prices and rents at date $t+1$.

The two forces that can generate housing booms in our model thus affect price growth forecasts in opposite directions. Indeed, the coefficient on the nowcast of the discount factor $\hat{m}_t^{i,h}$ is negative: when agents perceive a housing slump due to financial frictions, say, they expect mean reversion and hence forecast high price growth. At the same time, the coefficient on the growth rate $\hat{g}_t^{i,h}$ is positive: agents also predict high price growth when they perceive rents to be rising. Averaging across agents, we have that renters have higher price growth forecasts than owners if they perceive higher rent growth or a lower discount factor.

Now consider an owner at the end of the financial crisis, who initially believes the dis-

[9]The resulting cross sectional variance of nowcasts is lower than the posterior variance of any individual's nowcast. This is because the latter incorporates uncertainty about the time series evolution of the hidden state, whereas the former reflects only noise in signals.

count factor and the growth rate of rents are both below average, so both forces contribute to a low price-rent ratio by (5). The owner thus expects prices to rebound from the crisis. For concreteness, suppose further that the owner perfectly observes prices, but not rents. He thus initially sees a recovery with low price growth. As a result, he adjusts his nowcast in the direction dictated by the forecast error – towards lower rent growth and a higher discount factor which both imply a lower price forecast. Contrast this with a renter who also observes rent. As he sees strong rent growth, he attributes the sluggish movement in prices to a low discount factor. Both high rent growth and a low discount factor lead the renter to forecast higher price growth.

What changes with noise? The argument so far has only used average responses to price and rent realizations. A version of it will thus go through when agents observe rents and prices with error; all that changes is that average responses to signals are weaker when signals are less precise. The new feature with noise is that the model can also speak to dispersion of forecasts. There are two opposing forces here: less precise signals make responses weaker, but noise also makes signals more dispersed. Higher noise can therefore make forecasts more or less dispersed. In fact, at the extremes of perfect and completely uninformative signals, forecast dispersion is zero, whereas it is positive at intermediate dispersion.

We also want to understand how renters make more accurate forecasts on average even though as group they make larger mean squared forecasts errors. The forecast error differs from (minus) the forecast only by a constant common to all agents, namely the new price growth realization. The mean squared error can therefore be decomposed into the dispersion of forecasts and the squared bias, that is, the average forecast error:

$$MSE_t^h\left(\Delta p_{t+1}\right) = B_{1\bullet}\left(A-I\right)\Omega^h\left(A-I\right)^\top B_{1\bullet}^\top + \left(\Delta p_{t+1} - B_{1\bullet}\left(A-I\right)\bar{x}_t^h\right)^2, \tag{13}$$

where Ω^h is the posterior variance of nowcasts, the fixed point of (12). The first term depends on the information structure only. The second term, in contrast, depends on the specific realization of the two factors driving prices and rents. In particular, if the boom is driven by rent growth that is better observed by renters, then it is possible for the average renter to be closer to the actual realizations, even though he has less precise signals about prices.

Quantitative analysis. We now explore whether the mechanism sketched above can quantitatively account for the mean and variance of the cross section of price growth forecasts in the 2014 survey. We focus on learning in the top growth region, which represents the biggest challenge for our model. We use the data going back to the 1970s on 28 German cities in the top growth region for which we have prices and rents available, see the discussion in Section 2.4. Based on the time series in Figure 4, we estimate the distribution of price-rent ratios and rent growth. For the recent house price boom, we use our high quality price and rent data (p_t, r_t) from bulwiengesa AG to construct prices and rents for these cities. Our focus on the top growth region has the advantage that its boom was leading the rest of Germany, so the assumption that households learn from local prices and rents is appropriate. In other regions, there may also be regional spillovers in learning.

We pick parameters for the agents' perceived distribution of prices and rents as well as for their information sets in two steps. A first set of parameters can be fixed up front. The model period is a year and we focus on the early boom years starting at the onset in 2010 up to 2014, where the first expectations data is available. We set the average (gross) rent growth rate to $G = 1.033$, the long-run average computed from the data on the top growth region in Figure 4. We then choose the average discount factor $M = 0.932$ to obtain the average price-rent ratio in the data of 25.8. Estimating an AR(1) process for rent growth in the top German cities delivers an AR(1) parameter $\alpha_g = 0.72$ and an innovation volatility of 2.58%. We thus impose in particular that agents correctly perceive the unconditional volatility of rent growth in the data which is 3.72%. Furthermore, we assume that agents have identical beliefs about the persistence of shocks to rents and shocks to the discount factor $\alpha_m = \alpha_g$. Therefore, the process for the price-to-rent ratio is also AR(1) with an autoregression parameter of 0.72. With these parameter choices, the coefficient on rent growth in (5) is pinned down at $\beta_g = 2.36$ and the coefficient on discount factor shocks is $\beta_m = 3.27$. As a result, households' beliefs are consistent with the stylized fact of "excess volatility" of house prices.

We select a second set of parameters by matching a set of target moments. For beliefs, we need to choose noise variances as well as a distribution of priors for each household type. To discipline our model, we make two assumptions. First, we assume priors at the beginning of the year 2010 are such that average households of both types (i) agree on the nowcast of the rent growth rate and (ii) correctly nowcast the beginning of 2010 price-rent ratio. The assumption that the means $\bar{x}_0^h$ are identical across types h implies that any average disagreement between types observed in 2014 must be explained by learning over the boom years 2010 to 2014, and not by disagreement at the beginning. Second, our baseline specification assumes that renters observe rent without noise, and that owners do not observe rents at all. In other words, the noise variances on rent for renters and owners are zero and infinity, respectively. This leaves two noise variances for house price observations of owners and renters, respectively.

Parameters and targets for the moment matching exercise are collected in Table 6. In addition to the parameter values for the prior and noise – the last three lines – we need to find the innovation volatility for the discount factor m. We match the average forecasts and the cross sectional dispersion of forecasts in 2014 of all sample households living in the 28 cities in the top growth region from which we calculate prices and rents. Consequently, renters make higher forecasts than owners while at the same time their forecast volatility is higher. The most subtle point perhaps is how forecast moments help identify the perceived volatility of the discount factor. The argument here is that observed signals can only lead to sufficient disagreement if they carry enough news about the future. This leads the calibration to a perceived discount factor that has relatively large innovations at about 12%.

Table 6: Model Results

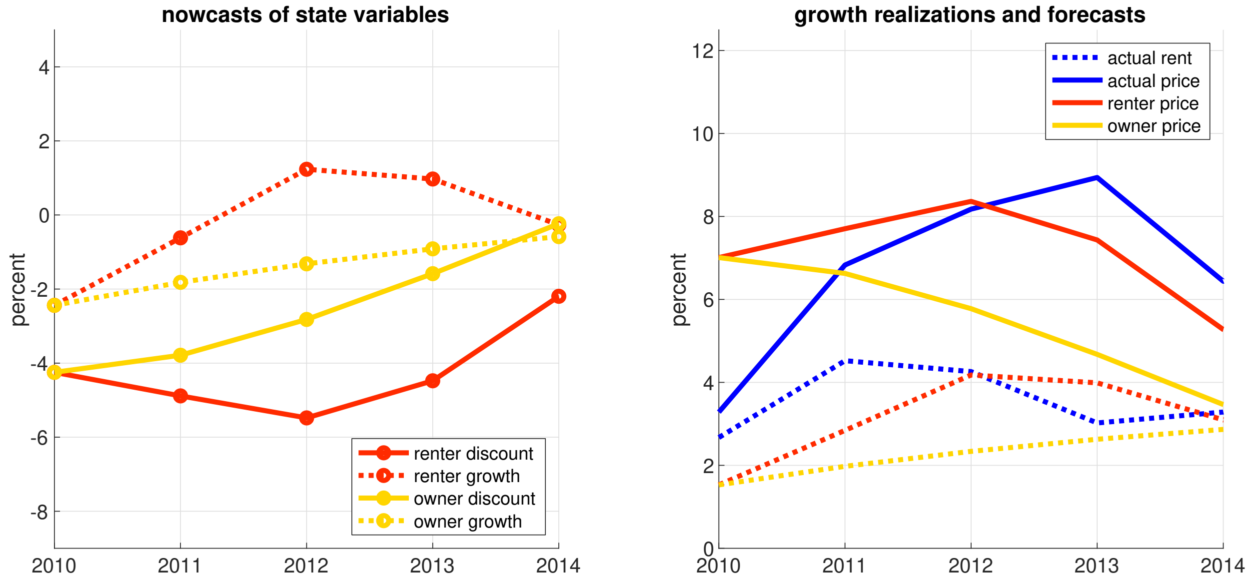
Parameters	Baseline values	Targets	data = model
volatility of discount innovations ε^m_{t+1}	0.121	avg forecast renter	5.27%
initial average rent growth nowcast $\bar{g}_0$	0.009	avg forecast owner	3.47%
noise volatility for renter	0.217	vol forecast renter	4.83%
noise volatility for owner	0.205	vol forecast owner	3.68%

Consider the calibrated parameter values that speak to the information set. The initial average nowcast of rent growth is low, 0.9%. In order for the average agent to correctly nowcast an initial price-rent ratio that is almost 20% below steady state requires an average nowcast $\bar{m}_0$ of the discount factor of -4.3%. The coefficient on the discount factor in (5) is 3.27 so the discount factor does most of the work in rationalizing low prices. Renters obtain slightly more noisy signals than owners, although both deal with substantial noise of 21.7% and 20.5%, respectively.

Figure 13 illustrates the belief dynamics that lead the model to account for the data. The color coding is red for renters and yellow for owners. The left panel shows both types' average nowcasts of rent growth (dotted line) and the discount factor (solid line), starting from their common initial condition in 2010. The right panel shows both types' forecasts for rent growth (dotted line) and price growth (solid line) as well as the subsequent actual realizations (in blue), aligned so the realized forecast error is the vertical distance. We thus obtain a quantitative version of the narrative described above. The initial common forecast is based mostly on the common belief that the discount factor coming out of the financial crisis is very low. As the boom starts to develop, opinions diverge: owners perceive a quick rebound of the discount factor, whereas renters take into account rent growth and believe that the rebound of the discount factor is sluggish. Price growth forecasts thus decline, but substantially more so for owners. In 2014, both types make forecasts that are too conservative, with owners about 1.8 percentage points below renters on average, as in the data.

Overall, our calibration strategy is guided by the stated goal of this section: to illustrate that learning about housing cost can lead to large divergence of expectations. The stark

Figure 13: Model Results



difference in information sets in our baseline serves to make our mechanism more powerful. At the same time, assuming initial agreement between average owners and renters also asks it to deliver a lot. An alternative strategy might introduce more or less noise in the rent observations of renters and owners, respectively, and might also allow for some initial disagreement between the average renter and the average owner – after all, it is plausible that the information frictions we emphasize were in place also before the financial crisis. Since we do not have direct evidence on beliefs in this early period, we cannot provide accurate detail here. The quantitative insight we take away is that our mechanism can have first order effects on beliefs.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino, "Perception of House Price Risk and Homeownership," *NBER Working Paper*, 2018, *25090.*

Armona, Luis, Andreas Fuster, and Basit Zafar, "Home Price Expectations and Behaviour: Evidence from a Randomized Information Experiment," *Review of Economic Studies*, 2019, *86* (4), 1371–1410.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel, "The Economic Effects of Social Networks: Evidence from the Housing Market," *Journal of Political Economy*, 2018, *126* (6), 2224–2276.

Campbell, John Y. and Robert J. Shiller, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1988, *1* (3), 195–228.

Campbell, Sean, Morris Davis, Joshua Gallin, and Robert F. Martin, "What Moves Housing Markets: A Variance Decomposition of the Rent-Price Ratio," *Journal of Urban Economics*, 2009, *66* (2), 90–102.

Case, Karl E. and Robert J. Shiller, "The Efficiency of the Market for Single-Family Homes.," *American Economic Review*, 1989, *79* (1), 125–137.

_ and _ , "Is there a bubble in the housing market," *Brookings Papers on Economic Activity*, 2003, *2*, 299–362.

_ , _ , and Anne Thompson, "What have they been thinking? Home buyer behavior in hot and cold markets," *National Bureau of Economic Research No.*, 2012, *w18400.*

Cogley, Timothy and Thomas J. Sargent, "The market price of risk and the equity premium: A legacy of the Great Depression?," *Journal of Monetary Economics*, 2008, *55*, 454–476.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, "Speculative Dynamics," *Review of Economic Studies*, 1991, *58* (3), 529–546.

De la O, Ricardo and Sean Myers, "Subjective Cash Flow and Discount Rate Expectations," 2019. Working Paper, Stanford.

DeFusco, Anthony A., Charles G. Nathanson, and Eric Zwick, "Speculative Dynamics of Prices and Volume," 2018. Working Paper, Kellogg.

Favara, Giovanni and Zheng Song, "House price dynamics with dispersed information," *Journal of Economic Theory*, 2014, *149*, 350–382.

Gao, Zhenyu, Michael Sockin, and Wei Xiong, "Learning about the Neighborhood," 2020. Forthcoming *Review of Financial Studies*.

Gelain, Paolo and Kevin J. Lansing, "House prices, expectations, and time-varying fundamentals," *Journal of Empirical Finance*, 2014, *29*, 3–25.

Glaeser, Edward L. and Charles G Nathanson, "Housing Bubbles," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5 Elsevier 2015, pp. 701–751.

_ , Joseph Gyourko, Eduardo Morales, and Charles G. Nathanson, "Housing dynamics: An urban approach," *Journal of Urban Economics*, 2014, *81*, 45–56.

Glaeser, Edwards L. and Charles G. Nathanson, "An Extrapolative Model of House Price Dynamics," *Journal of Financial Economics*, 2017, *126* (1), 147–170.

Guren, Adam M., "House Price Momentum and Strategic Complementarity," *Journal of Political Economy*, 2018, *126* (3), 1172–1218.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun, "Search, Liquidity, and the Dynamics of House Prices and Construction," *American Economic Review*, 2014, *104* (4), 1172–1210.

Kaas, Leo, Georgi Kocharkov, Edgar Preugschat, and Nawid Siassi, "Low homeownership in Germany – a quantitative exploration," *forthcoming: Journal of the European Economic Association*, 2020.

Kaplan, Greg, Kurt Mitman, and Gianluca Violante, "The Housing Boom and Bust: Model Meets Evidence," 2020. Forthcoming *Journal of Political Economy*.

Kuchler, Theresa and Basit Zafar, "Personal Experiences and Expectations about Aggregate Outcomes," *Journal of Finance*, 2019, *74* (5), 2491–2542.

Landvoigt, Tim, "Housing Demand during the Boom: The Role of Expectations and Credit Constraints," *Review of Financial Studies*, 2017, *30* (6), 1865–1902.

Leombroni, Matteo, Monika Piazzesi, Martin Schneider, and Ciaran Rogers, "Inflation and the Price of Real Assets," *NBER Working Papers*, 2020, *26740.*

Lewellen, Jonathan and Jay Shanken, "Learning, Asset Pricing Tests, and Market Efficiency," *Journal of Finance*, 2002, *57* (3), 1113–1145.

Ludwig, Alexander, Jochen Mankart, Jorge Quintana, Nate Vellekoop, and Mirko Wiederholt, "House Price Expectatons and Housing Choice," 2020. Working Paper, University of Frankfurt.

Malmendier, Ulrike and Alexandra Steiny, "Rent or Buy? Inflation Experiences and Homeownership within and across Countries," 2021. Working paper, Berkeley.

_ and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *Quarterly Journal of Economics*, 2011, *126* (1), 376–416.

Piazzesi, Monika and Martin Schneider, in John Y. Campbell, ed., *Inflation Illusion, Credit, and Asset Pricing*, Chicago University Press, Chicago, IL 2008, pp. 147–181.

_ and _ , "Momentum Traders in the Housing Market: Survey Evidence and a Search Model," *American Economic Review*, 2009, *99* (2), 406–411.

Vissing-Jorgensen, Annette, "Perspectives on Behavioral Finance: Does "Irrationality" Disappear with Wealth? Evidence from Expectations and Actions," *NBER Macroeconomics Annual*, 2003.

Weber, Jan Philip and Gabe Lee, "On the Measure of Private Rental Market Regulation Index and Its Effects on Housing Rents: Cross-Country Evidence," IREBS Beiträge zur Immobilienwirtschaft, Heft 21 2018.

A House Price Data

A.1 Constructing house price and rent indices for Germany

The detailed house price and rent dataset from bulwiengesa A we use to construct price indices since the year 2005 contains the following price series for houses in all 401 German counties:

- single-family homes
- apartments (newly built, "Erstbezug")
- apartments (already existing, "Wiederbezug")
- town houses (not available for all counties)

As only 7.5 percent of households in the PHF live in town houses and the price series is not available for all German counties, we ignore these building types when constructing our house price index.

We merge the two price series on apartments assuming that about 15 percent of apartments are new and 85 are already existing. This is the official number used by the Deutsche Bundesbank, and it is based on estimates of the German Federal Statistical Office. This leaves us with two price series: one for single-family homes and the other one for apartments. We combine the two series to one using data from the PHF. For each growth region, we calculate the share of households living in single-family homes out of the share of households that live in either single-family homes or apartments (more than 90 percent of the population). The respective data is shown in Table A.1. Not surprisingly, single-family homes are less common in high growth regions, as the high growth region contains a lot of cities.

Table A.1: Types of housing in different growth regions

	Growth Region			
	Low	Medium Low	Medium High	High
Single-Family Home	49.50	43.21	41.22	16.12
Apartment	50.50	56.79	58.78	83.88

As for rents, there are two series available:

- apartments (newly built, "Erstbezug")
- apartments (already existing, "Wiederbezug")

According to the PHF, more than 80 percent of renters are living in apartments. A rental market, especially for single-family homes, hardly exists. As a result, data on rented houses is scarce. To construct the series for rent prices, we again combine the two price series assuming that about 15 percent of apartments are new.

A.2 Long-run Data on House Price and Rents

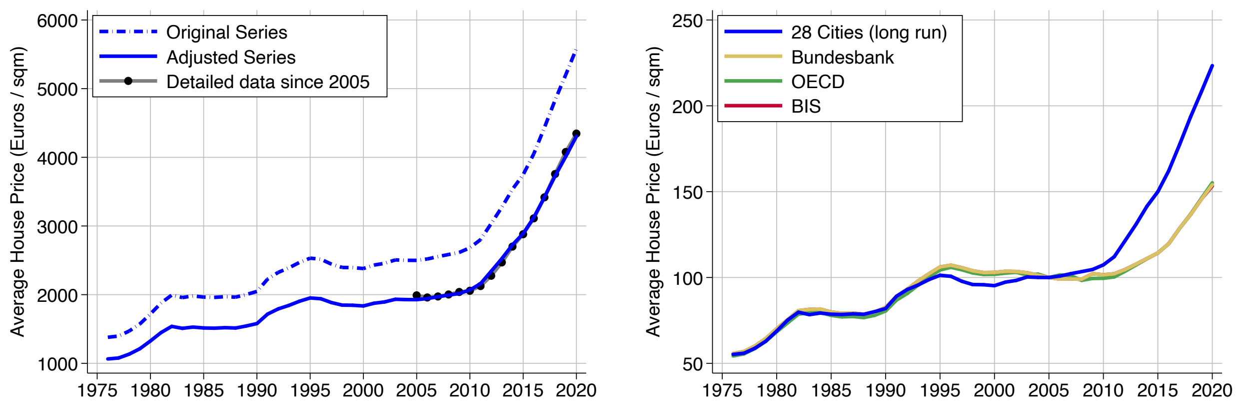
The Bundesbank provided us with the following data for the time span between 1976 and 2020 collected by bulwiengesa AG:

- Purchase prices of apartments (newly built, "Erstbezug")
- Rents of apartments (newly built, "Erstbezug")
- Rents of apartments (already existing, "Wiederbezug")

The data is available for a total of 50 (West German) cities. The series for West-Berlin is discontinued in 1990. We select from the remaining 49 cities those, who are in the highest growth region in the period between 2010 and 2020. This leaves us with a total of 28 cities.

Leveling house price data For the long house price series, we only know the purchase price of newly built apartments, which is consistently higher than the average purchase price of a house in the highest growth region. To make the long price series comparable with the 2005-2020 bulwiengesa data, we re-level the data as shown in the left panel of Figure A.1. The dash-dotted line shows the original long price series for newly built

Figure A.1: Leveling of house price data



apartments. The gray series with black dots is the price series determined for exactly the same 28 cities/counties from the detailed bulwiengesa data we have available since 2005. For the time frame in which both series are available, they move in parallel, but newly built apartments are just more expensive. To compensate for this effects, we calculate the mean ratio between the two price series at any date in between 2005 and 2020. We then normalize the original series by this ratio. The solid line in the left panel of Figure A.1 shows the resulting adjusted time series which is hardly distinguishable from the detailed data series between 2005 and 2020. This makes us confident that we can use the newly built apartment price series to extrapolate the evolution of prices backward in time until 1976.

In the right panel of Figure A.1, we compare the long-run price series with publicly available time series data for house prices in Germany from the Deutsche Bundesbank, the OECD and the BIS. In the 1970s to early 1990s, all of these data are based on the same common data source that also underlies our price series. The Deutsche Bundesbank uses data on the purchase prices of newly built apartments for all 50 cities, not the 28 we selected, but the price dynamics seem to be fairly similar. With German reunification, the official sample of the Bundesbank was extended to a total of 100 cities and in 1995 to 125 cities, also covering East Germany. This is the first time where our series departs from official sources as we want to provide data for one set of counties over time, while the Bundesbank seeks to approximate house price dynamics all over Germany. The second structural break in the official data series is in 2005, when the detailed bulwiengesa data starts. From this point onward, the Bundesbank uses data covering all German counties from different data sources. As we have seen before, the house price boom is much stronger in the highest growth region than in Germany as a whole. Not surprisingly, our price series then departs from official ones at the onset of the German house price boom.

Rents Turning to rent prices, the long-run data contains exactly the same information as the detailed bulwiengesa data. Hence, when applying the same weighting as described in the previous section, we obtain the same price series from 2005 onwards, see the right panel of Figure 3. When combining the rent and the price series, we obtain the time series for the price-rent ratio shown in Figure 4.

B Documentation of Variables and Sample Selection

While we already provide detailed information on both the house price and rent data as well as on the assessment of households' price expectations in the main text of the paper, this appendix documents all other survey variables we used in the regressions and figures reported in this paper as well as our sample selection procedure.

B.1 Explanatory Variables Used in Regressions

The following variables are used as explanatory variables in regressions. We use data from multiple waves of the Panel on Household Finances (PHF) as well as the Online Survey – Households. As variable names might differ across those surveys, but the survey methodology is the same, we only provide the variable name for our baseline survey, the second wave (2014) of the Panel on Household Finances. A full set of variable names and Stata codes are available upon request.

B.1.1 Demographics, Income, Wealth

Age (`ra0300`): Age refers to the age of the household head. The household head is the "Financially knowledgeable person" of the interview, i.e. the person that knows best about the households' finances and answers the household questionnaire including all details on wealth holdings for the entire household.

Number of Household Members (`anzhhm`): Number of all household members currently living in the household, including the household head.

Net Wealth (`constructed from several variables`): We sum up all safe assets (bonds, pension accounts, life insurances), equity (stocks, mutual funds, businesses), real estate investments and non-interest bearing assets (cash, vehicles, private assets) at the household level and subtract any outstanding debts for the entire household.

Net Household Income (`dhi0600`): Household heads are asked for the total net disposable household income, consisting of wages, salary, income from self-employment, retirement benefits or pensions, income from public aid, income from renting, income from leasing, housing allowance, child benefits, and other income.

Education (`pa0200`): This variable describes the highest education level the household head has achieved. Low education refers to having no school degree or at most a lower level secondary school degree (corresponding to the German "Realschule"), Upper Secondary education comprises all upper secondary level degrees ("Abitur", "Fachhochschulreife") and tertiary education refers to university or equivalent degrees.

B.1.2 Behavioral Traits

Financial Literacy (`dhnm0100`, `dhnm0200`, `dhnm0300`): Household heads are asked three questions in the areas of interest rate compounding, inflation and diversification, which we use to measure their financial literacy. In particular, those questions are:

Question 1: "Let us assume you have a balance of 100 euros in your savings account. This balance bears interest at an annual rate of 2%, and you leave it there for 5 years. What do you think: How high is your balance after 5 years?"
Possible answers: 1 – Higher than 102 euros / 2 – Exactly 102 euros / 3 – Lower than 102 euros / -1 – Don't know / -2 – No answer

Question 2: "Let us assume that the interest paid on your savings account is 1% per year and the inflation rate is 2% per year. What do you think: After a year, will you be able to buy just as much, more or less than today with the balance in your savings account?"
Possible answers: 1 – More / 2 – Just as much / 3 – Less than today / -1 – Don't know / -2 – No answer

Question 3: "Do you agree with the following statement: "The investment in the stock of a single company is less risky than investing in a fund with stock in similar companies?"
Possible answers: 1 – I agree / 2 – I do not agree / -1 – Don't know / -2 – No answer

We form a literacy index out of the answers to these three questions. Literacy is "Very Low" in case the household head gives a wrong or no answer to all three questions; it is "Low" if the respondent was able to answer exactly one question correctly; it is "Medium" if the respondent was able to answer exactly two questions correctly; it is "High" if the respondent gave correct answers to all three questions.

Patience (`zi105`): The patience measure comes from a Likert scale where respondents are asked to assess their level of patience. The exact wording of the question is: "How do you view yourself personally: Are you in general a person who is patient or do you tend to be impatient?" Answers can be given on a scale from 0 to 10 where zero refers to being "Very patient" and 10 to "Very impatient". We form three categories for our variable of patience with "High patience" including all respondents who answered at most three on the scale, "Average patience" between 4 and 6 of the scale and "No patience" if respondents rated their patience to be between 7 and 10 on the scale. We chose the division into three groups according to the frequency of answers. Consequently, each of the three groups is populated approximately by an equal share of households.

Risk aversion (`zi103`): The survey assesses risk aversion in the same way as patience. Households are asked about how risk averse they view themselves. They can again provide a score on a scale from 0 to 10. We split the sample in two groups of about equal size, those "Below Median" and those "Above Median" risk aversion.

Financial risk aversion (`hd1800, dhd2800`): An alternative measure of risk aversion – financial risk aversion – asks about the household's investment behavior. The exact wording of the question we use is: "If savings or investment decisions are made in your household: Which of the statements best describes the attitude toward risk?"

Possible answers:

1 – We take significant risks and want to generate high returns.
2 – We take above-average risks and want to generate above average returns.
3 – We take average risks and want to generate average returns.
4 – We are not ready to take any financial risks.
5 – No uniform classification is possible for the household as a whole.

If the household head chooses answer 5, he or she is asked what investment behaviour he/she has for him-/herself. We again split the sample in two groups of about equal size, those "Below Median" and those "Above Median" risk aversion.

B.1.3 Tenure and Growth Region

Tenure (`dhb0200a, dhb0200b, hb0500`): We say a household is an "Owner" if the household owns at least 50 percent of the property the household members are living in, the so-called primary residence. All other households are classified as "Renters".

Growth Region: Our house price data set is based on German counties (Kreise and kreisfreie Staedte). There are 401 such counties in Germany, out of which approximately one half is covered by the Panel on Household Finances. To define growth regions, we take the 401 counties and calculate their "trend growth" over the course of the past 10 years, i.e. starting from the beginning of the house price boom in 2011 to the last year available in the data (2020). To this end, we regress the log of the house price in each county on a year variable. The regression coefficient is our trend growth variable. We divide the 401 counties into four quartiles with respect to trend growth. This leaves us with four growth regions, i.e. those with Low, Medium Low, Medium High and High trend growth in between 2011 and 2020.

B.1.4 Housing/Regional Characteristics

Time Household has been Living in Current Residence (`intjahr, dhb0120`): Number of years since household has moved into their current primary residence (measured at the time of the interview).

Size of the Local Community (`bikgk10`): While the county ("Kreis") a household is living in is the unit on which we measure house prices, this is an even finer description of the size of the local municipality. The official community size measure in Germany is the so-called "BIK Scale" that classifies community sizes along ten categories:

1 – $<$ 2,000 inhabitants
2 – 2,000 to 4,999 inhabitants
3 – 5,000 to 19,999 inhabitants
4 – 20,000 to 49,999 inhabitants
5 – 50,000 to 99,999 inhabitants and BIK structure type 2 (city region) / 3 (second tier towns) / 4 (third tier towns)
6 – 50,000 to 99,999 inhabitants and BIK structure type 1 (metropolitan area)
7 – 100,000 to 499,999 inhabitants and BIK structure type 2/3/4
8 – 100,000 to 499,999 inhabitants and BIK structure type 1
9 – $\geq$ 500,000 inhabitants and BIK structure type 2/3/4
10 – $\geq$ 500,000 inhabitants and BIK structure type 1

Building in Which Household Lives Needs Renovation (`sc0400`): After each interview, the interviewer is asked to fill out a questionnaire on the interview with the household, including questions about the atmosphere, the willingness of the respondent to answer, the living conditions and the state of the house. From these para-data we take the following question to assess the quality of the building that households live in: "Please describe the condition of the building."

Possible answers:

1 – Clean and well maintained
2 – A few small cracks in the facade, and isolated cases of peeling paint
3 – Badly in need of renovation
4 – Dilapidated

Our indicator of renovation needs refers to all answer categories greater than 1.

Quality of the Household's Residence (`sc0200`): The interviewer is also asked to give a rating of the house. The exact wording of the question is: "Please rate the building."

Possible answers:

1 – Exclusive
2 – Very good
3 – Satisfactory
4 – Simple
5 – Very simple

Size of Households Residence (`hb0100`): This is a simple measure of the square meter size of the household's residence. This sqm number refers to the living space of the house or the apartment, not the size of the full property.

B.2 Variables Used in Figures

The following variables are used as explanatory variables in figures.

Equity Investors (`constructed from several variables`): We say a household is an equity investor, if the household's portfolio contains any of the following assets: directly held shares, certificates, other securities, mutual funds and private pensions primarily invested in stocks, hedge funds, other risky assets, silent partnerships as well as managed accounts. In our definition, equity also contains the value of self-employment business.

Perceived House Price Growth (`hb0900`): We can match a subsample of households (the panel households) across the 2011 and 2014 waves as well as the 2014 and 2017 waves of the PHF, respectively. To construct a household's perceived house price growth (see Figure 10), we look at those households who were owners of the same residence in two consecutive waves. For these households, the PHF asks household heads for a price quote of their current primary residence in both waves. Using these price quotes, we calculate each household's perceived annualized house price growth rate and contrast it with the average house price growth rate in the county the household is living in.

Renters with Plans to Buy a House (`dhb3000`): Renters in the PHF are asked about their intention to buy a property. The exact working of the question is: "Does your household intend to buy or build a house or flat for your own accommodation?". The possible answers are "Yes", "No" and "Don't Know".

Owner's Tenure Time (`hb0700`): Owners are asked about the year in which they became owner of the property they currently live in. Note that this date doesn't have to coincide with the date the household moved into the property. We subtract the date the household became owner of the property from the date of the interview to get to the household's tenure time.

Intentions to Move (`dhb0125`): In wave 3 of the PHF, all respondent households are asked about their intentions to stay in their current primary residence. The exact working of the question is: "How long do you expect you/at least one member of the household will continue to live in your primary residence?" The answer is given on a numerical scale indicating "At least another ... years". Alternatively the household can choose "Forever" or "Don't Know".

B.3 Sample Preparation and Sample Selection in the PHF

In this section, we describe our sample preparation and selection procedure for the PHF, our primary dataset. We proceed in a similar way with the Online Survey – Households. Yet, since the BOP-HH is less rich, the sample preparation and selection is much simpler. A description is available upon request.

In the PHF, we use only the first implicate of the full dataset for simplicity. Since hardly any of the variables we use in this paper are imputed variables, this doesn't play very much of a role. We then proceed as follows to prepare the sample:

1. **Matching between PHF and house price data:** We first match households to their specific county. This is done via the "Kreiskennziffer", a unique identifier for each of the 401 German counties. Note that the scientific use files of the PHF do not contain this information, neither do they provide any house price or rent price data on the county level. Users who wish to get access to this data should get in contact with the Research Data and Service Center of the Deutsche Bundesbank.[10]

2. **Classification of Household Portfolios:** We classify household portfolios according to the following schedule:
   - *Safe Assets:* This category contains deposits, directly held bonds as well as mutual funds and private pensions consisting predominantly of money market investments and/or bonds. We subtract from this mortgages as well as any unbacked bank credit.
   - *Real Estate:* This category contains mutual funds and private pensions consisting predominantly of real estate investment, the value of the household's main

[10] See `https://www.bundesbank.de/en/bundesbank/research/rdsc` for further details.

residence as well as of all secondary residences, all direct real estate investments made for the purpose of renting them out to others or keeping them empty to enjoy capital gains, as well as all business related real estate.

- *Equity:* This category contains mutual funds and private pensions consisting predominantly of stocks, hedge funds or any other risky investments, silent partnerships, managed accounts, directly held shares, certificates and other securities as well as other financial assets. We also include the value of self-employment business in this category.
- *Non Interest-Bearing:* This category contains money other households owe to the observed household, cash, vehicles and valuables. We subtract from this all money the observed household owes to any other household.

The sum of safe assets, real estate, equity and non interest-bearing assets yields total household net wealth. We derive portfolio shares for the four asset categories by dividing their sum on the household balance sheet by net wealth.

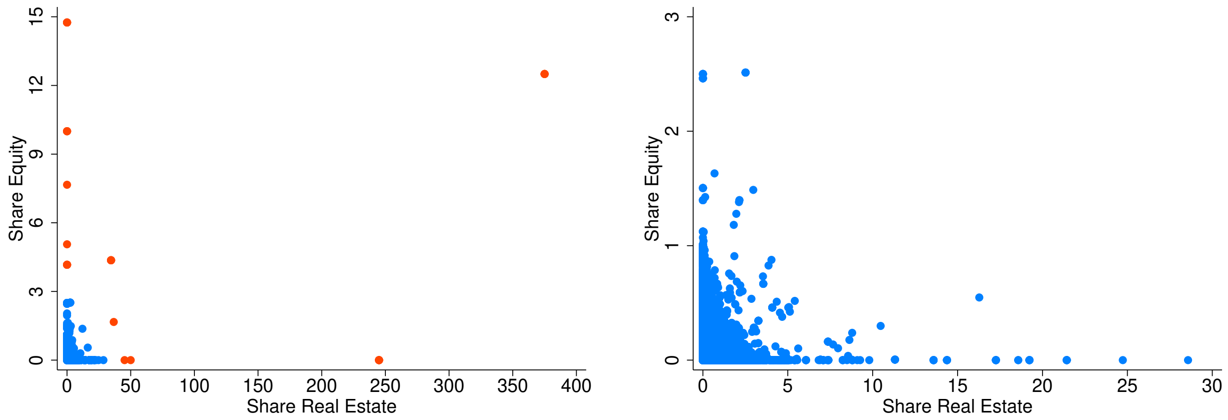
3. **Household Expectations:** The PHF asks households about their expectations towards (i) the deposit rate one year ahead of the time of the interview, (ii) house price growth over the next 12 months, (iii) the growth in equity price as measured by the German Stock Market Index (DAX) as well as (iv) growth in the general price level. Except for the question about the expected deposit rate, all expectation questions are asked in the two-step format discussed in section 3. After having coded all households' numerical expectations, we cut off the top and the bottom 1 percent of reported expectations by "missing" them to missing. This is done in order to deal with outliers.

With the full sample at hand, we clean the sample and identify outliers as follows:

- We drop all households who report a non-positive net wealth balance in their portfolio. The majority of those households holds no or only a tiny amount of real estate and equity, but typically has a lot of unbacked consumer credit and credit card debt.

- We then look at portfolio shares. The left panel of Figure B.2 shows a scatter plot of portfolio shares of households. We drop all observations represented by a red dot in the figure, which are rather extreme. This means that we drop households with a more than 300 percent equity share or a 3000 percent real estate share in their portfolio with the understanding that this most likely represents coding or data collection errors. The right panel of Figure B.2 shows the resulting distribution of portfolio shares across households.

Figure B.2: Portfolio Shares and Outlier Identification



C Additional Datawork

The aim of this appendix is twofold. On the one hand, we provide the full regression tables for the regressions present in short in Tables 1 and 2. On the other hand, we provide additional robustness checks using alternative explanatory variables as well as data from the 2017 version of the PHF.

C.1 The Cross Section of House Price Forecasts in 2014

Table 1 conveyed as simple but important message for our paper: In a linear regression setting, only two variables are relevant for explaining households' forecasts: tenure and location. While this table only listed variables that were significantly different from zero in at least one regression, Table C.2 shows the entire regression results.

Table C.2: The Cross Section of House Price Forecasts in 2014 (see Table 1)

	(1)	(2)	(3)	(4)	(5)
Demographics, Income, Wealth					
Age					
< 30	1.159	1.125	0.701	0.530	0.092
	(0.707)	(0.692)	(0.695)	(0.663)	(0.676)
30 - 39	1.365**	1.315**	0.957	0.711	0.371
	(0.591)	(0.587)	(0.578)	(0.566)	(0.582)
40 - 49	−0.227	−0.245	−0.227	−0.224	−0.296
	(0.451)	(0.450)	(0.451)	(0.432)	(0.443)
60 - 69	0.337	0.329	0.377	0.424	0.354
	(0.467)	(0.467)	(0.462)	(0.453)	(0.474)
≥ 70	−0.171	−0.221	−0.153	−0.198	−0.213
	(0.418)	(0.429)	(0.427)	(0.420)	(0.424)
Number of Household Members					
1	0.576	0.540	0.253	−0.070	−0.475
	(0.434)	(0.434)	(0.437)	(0.425)	(0.435)
3	0.314	0.293	0.456	0.337	0.325
	(0.479)	(0.481)	(0.470)	(0.456)	(0.457)
4	0.371	0.301	0.506	0.425	0.336
	(0.665)	(0.655)	(0.655)	(0.647)	(0.611)
≥ 5	0.186	0.203	0.599	0.625	0.731
	(0.551)	(0.548)	(0.547)	(0.555)	(0.580)
Net Wealth Quartiles					
1st Quartile	1.931**	1.930**	0.135	0.073	−0.266
	(0.478)	(0.522)	(0.683)	(0.629)	(0.572)
2nd Quartile	0.939**	0.980**	−0.197	−0.149	−0.159
	(0.448)	(0.441)	(0.467)	(0.454)	(0.459)
4th Quartile	0.279	0.296	0.547	0.099	0.059
	(0.292)	(0.296)	(0.300)	(0.295)	(0.311)

Net Household Income Quartiles					
1st Quartile	0.107	0.125	0.302	0.404	0.623
	(0.560)	(0.559)	(0.559)	(0.547)	(0.554)
2nd Quartile	0.030	0.050	0.074	0.165	0.138
	(0.438)	(0.437)	(0.439)	(0.430)	(0.422)
4th Quartile	−0.208	−0.226	−0.331	−0.381	−0.491
	(0.320)	(0.324)	(0.320)	(0.310)	(0.302)
Education					
Low	−0.249	−0.256	−0.398	−0.383	−0.426
	(0.566)	(0.569)	(0.564)	(0.562)	(0.567)
Tertiary	0.394	0.402	0.270	0.134	0.062
	(0.305)	(0.315)	(0.311)	(0.301)	(0.302)
BEHAVIORAL TRAITS					
Financial Literacy					
Very Low		1.619	1.659	1.690	2.070
		(1.351)	(1.397)	(1.334)	(1.435)
Low		−0.680	−0.755	−0.883	−0.752
		(0.741)	(0.731)	(0.739)	(0.776)
Medium		−0.163	−0.127	−0.294	−0.212
		(0.345)	(0.343)	(0.338)	(0.351)
Patience					
Not Patient		0.350	0.279	0.287	0.415
		(0.385)	(0.384)	(0.370)	(0.364)
Very Patient		0.118	0.055	0.026	0.011
		(0.338)	(0.336)	(0.330)	(0.325)
Risk Aversion					
Below Median		−0.062	−0.120	−0.001	0.112
		(0.300)	(0.299)	(0.294)	(0.296)
TENURE AND LOCATION					
Tenure					
Renter			2.438***	2.342***	2.087***
			(0.382)	(0.371)	(0.378)
Growth Region					
Low				−2.063***	−1.559***
				(0.439)	(0.427)
Medium Low				−1.578***	−1.450***
				(0.433)	(0.434)
High				1.368***	1.019**
				(0.416)	(0.399)
HOUSING/REGIONAL CHARACTERISTICS					
Time Household has been Living in Current Residence (Years)					
$t < 5$					0.293
					(0.456)
$5 \leq t < 10$					0.573
					(0.445)
$10 \leq t < 15$					0.191
					(0.525)

Size of the Local Community (Inhabitants)					
2k < 5k					−1.591 (1.303)
5k < 20k					−0.129 (0.795)
20k < 50k					−0.117 (0.764)
50k < 100k, periphery					0.680 (0.981)
50k < 100k, center					0.884 (0.984)
100k < 500k, periphery					0.706 (0.745)
100k < 500k, center					0.607 (0.767)
≥ 500k, periphery					1.112 (0.828)
≥ 500k, center					1.762** (0.720)
Building in Which Household Lives Needs Renovation					
Yes					−0.421 (0.389)
Quality of the Household's Residence (Interviewer Rating)					
Exclusive					−0.028 (0.625)
Satisfactory					−0.338 (0.387)
Simple					−0.448 (0.566)
Very simple					−0.084 (1.008)
Size of Households Residence					
Sqm size/100					−1.619*** (0.617)
Sqm size/100 squared					0.423*** (0.127)
Constant	1.749*** (0.473)	1.708*** (0.540)	1.366** (0.542)	2.102*** (0.663)	2.903** (1.159)
Number of Cases	3647	3646	3646	3646	3598
R-Square	0.038	0.042	0.063	0.119	0.142

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

C.2 Clustered standard errors

The standard errors in Table 1 are not clustered. To check whether clustering affects our results, we divided households into 30 clusters according to local house price growth over the last decade, with a reasonable number of households in each cluster (> 100). Our results do not change much; tenure and region dummies are still highly significant. We also clustered by Bundesland but do not report results here; they are also consistent with those in Table C.3.

Table C.3: Regressions from Table 1 with 30 clusters by house price growth region

	(1)	(2)	(3)	(4)	(5)
Demographics, Income, Wealth					
Age Group 30–39	1.365***	1.315***	0.957**	0.711	0.371
	(0.458)	(0.460)	(0.455)	(0.448)	(0.449)
1st Net Wealth Quartile	1.903***	1.906***	0.132	0.024	−0.274
	(0.531)	(0.550)	(0.588)	(0.567)	(0.562)
2nd Net Wealth Quartile	0.939	0.980**	−0.197	−0.149	−0.159
	(0.474)	(0.468)	(0.473)	(0.443)	(0.415)
Behavioral Traits		yes	yes	yes	yes
Tenure					
Renter			2.438***	2.342***	2.087***
			(0.455)	(0.416)	(0.362)
Growth Region					
Low				−2.063***	−1.559***
				(0.402)	(0.248)
Medium Low				−1.578***	−1.450***
				(0.440)	(0.396)
High				1.368***	1.019**
				(0.463)	(0.400)
Housing/Regional Characteristics					
City Center $\geq$ 500k Inh.					1.762
					(0.952)
Sqm size/100					−1.619***
					(0.505)
(Sqm size/100)^2					0.423***
					(0.119)
Number of Cases	3647	3646	3646	3646	3598
R^2	0.038	0.042	0.063	0.119	0.142

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

C.3 Growth Region Controls vs. Past House Price Growth Experience

In the regression presented in Table 1 – and in more detail in Table C.2 – we use "Growth Region" as a control for the evolution of the regional housing market a survey respondent is living in. We constructed this variable by dividing all 401 counties for which we have house price data available into four groups, according to their trend house price growth over the period 2011 to 2020. Pooling about 100 counties into one group is convenient for the data work shown in this paper, as it creates consistency between regressions and figures. Yet, one concern is that this approach is too coarse in the sense that the four controls for regional housing market performance (Low, Medium Low, Medium High and High) do not adequately reflect households' actual regional housing market experience.

To test the robustness of our results with respect to this concern, we use more granular controls for a survey respondent's regional house price growth experience. In particular, we calculate house price growth on the county level over the last year (2013-2014), the last three years (annualized between 2011-2014) and the last five years (annualized between 2009-2014). We then use these variables instead of the "Growth Region" to control for a household's regional housing market experience in columns (2), (3) and (4) of Table C.4.

Controlling for a more fine-grained house price growth experience delivers the same core result we already saw in Table 1: In a linear regression setting, only two variables are relevant for explaining households' forecasts: tenure and location, or in this case the regional housing market performance. This is true for almost all measures of past house price growth on the county level. Only the one-year measure is not significant at the five percent level, but on the edge. All other variables, and especially the dummy variable for tenure, are of equal size and significance, regardless of the growth experience measure we use. The same is true for the R-square value. Overall, these findings show that controlling for regional housing market performance on a four region scale is enough and does not bias our results into a particular direction.

Table C.4: Growth Region Controls vs. Past House Price Growth Experience

	(1)	(2)	(3)	(4)
TENURE, LOCATION AND PAST HOUSE PRICE GROWTH EXPERIENCE				
Tenure				
Renter	2.087**	1.929**	1.915**	1.984**
	(0.389)	(0.360)	(0.429)	(0.432)
Growth Region				
Low	−1.559***			
	(0.211)			
Medium Low	−1.450***			
	(0.180)			
High	1.019***			
	(0.074)			
Past House Price Growth Experience				
Last Year		0.213		
		(0.085)		
Three Year Back Average			0.394***	
			(0.015)	
Five Year Back Average				0.463***
				(0.039)
DEMOGRAPHICS, INCOME, WEALTH				
Age				
< 30	0.092	0.208	0.157	0.105
	(0.363)	(0.328)	(0.354)	(0.353)
30 - 39	0.371	0.545	0.436	0.414
	(0.153)	(0.208)	(0.210)	(0.193)
40 - 49	−0.296	−0.350	−0.306	−0.332
	(0.412)	(0.361)	(0.410)	(0.409)
60 - 69	0.354	0.359	0.373	0.321
	(0.582)	(0.536)	(0.558)	(0.560)
≥ 70	−0.213	−0.246	−0.184	−0.211
	(0.340)	(0.326)	(0.318)	(0.315)
Number of Household Members				
1	−0.475	−0.485	−0.512	−0.569
	(0.504)	(0.513)	(0.546)	(0.546)
3	0.325	0.333	0.296	0.306
	(0.546)	(0.519)	(0.538)	(0.503)
4	0.336	0.449	0.340	0.363
	(0.588)	(0.493)	(0.568)	(0.575)
≥ 5	0.731	0.771	0.640	0.634
	(0.780)	(0.750)	(0.772)	(0.723)
Net Wealth Quartiles				
1st Quartile	−0.274	−0.187	−0.194	−0.337
	(0.640)	(0.671)	(0.677)	(0.665)
2nd Quartile	−0.159	−0.119	−0.109	−0.197
	(0.483)	(0.441)	(0.446)	(0.445)
4th Quartile	0.059	0.149	0.037	0.037
	(0.327)	(0.310)	(0.374)	(0.341)

Net Household Income Quartiles				
1st Quartile	0.623	0.749	0.636	0.671
	(0.478)	(0.527)	(0.559)	(0.584)
2nd Quartile	0.138	0.159	0.115	0.146
	(0.200)	(0.310)	(0.324)	(0.332)
4th Quartile	−0.491	−0.490	−0.446	−0.433
	(0.351)	(0.322)	(0.336)	(0.373)
Education				
Low	−0.426	−0.512	−0.480	−0.455
	(0.620)	(0.565)	(0.617)	(0.594)
Tertiary	0.062	0.066	0.032	0.044
	(0.126)	(0.132)	(0.129)	(0.135)
BEHAVIORAL TRAITS				
Financial Literacy				
Very Low	2.070**	1.998**	2.012**	2.161**
	(0.559)	(0.603)	(0.362)	(0.396)
Low	−0.752	−0.713	−0.724	−0.631
	(0.325)	(0.444)	(0.417)	(0.381)
Medium	−0.212	−0.165	−0.252	−0.176
	(0.327)	(0.269)	(0.277)	(0.280)
Patience				
Not Patient	0.415	0.390	0.364	0.385
	(0.343)	(0.354)	(0.381)	(0.368)
Very Patient	0.011	0.038	0.033	0.044
	(0.261)	(0.220)	(0.234)	(0.245)
Risk Aversion				
Below Median	0.112	0.013	0.005	0.016
	(0.202)	(0.183)	(0.214)	(0.224)
HOUSING/REGIONAL CHARACTERISTICS				
Time Household has been Living in Current Residence (Years)				
$t < 5$	0.293	0.281	0.303	0.315
	(0.404)	(0.429)	(0.402)	(0.366)
$5 \leq t < 10$	0.573	0.616	0.624	0.665
	(0.645)	(0.663)	(0.607)	(0.552)
$10 \leq t < 15$	0.191	0.219	0.268	0.246
	(0.462)	(0.485)	(0.511)	(0.521)

Size of the Local Community (Inhabitants)				
2k < 5k	−1.591	−0.425	−0.538	−0.052
	(0.626)	(0.523)	(0.508)	(0.572)
5k < 20k	−0.129	0.750	0.296	0.614
	(0.532)	(0.675)	(0.636)	(0.637)
20k < 50k	−0.117	0.387	−0.007	0.158
	(0.631)	(0.731)	(0.623)	(0.689)
50k < 100k, periphery	0.680	0.855	0.624	1.256
	(0.726)	(0.805)	(0.671)	(0.806)
50k < 100k, center	0.884	0.998	0.557	0.910
	(0.697)	(0.676)	(0.875)	(0.966)
100k < 500k, periphery	0.706	1.571**	1.181	1.761**
	(0.481)	(0.458)	(0.398)	(0.439)
100k < 500k, center	0.607	1.535	0.754	1.168
	(0.983)	(0.811)	(0.886)	(0.871)
≥ 500k, periphery	1.112	1.983	1.352	1.826
	(1.147)	(1.022)	(1.046)	(1.228)
≥ 500k, center	1.762**	3.008***	1.846***	2.247***
	(0.345)	(0.305)	(0.310)	(0.350)
Building in Which Household Lives Needs Renovation?				
Yes	−0.421	−0.330	−0.336	−0.309
	(0.429)	(0.400)	(0.453)	(0.445)
Quality of the Household's Residence (Interviewer Rating)				
Exclusive	−0.028	0.133	0.010	0.022
	(0.646)	(0.706)	(0.586)	(0.575)
Satisfactory	−0.338	−0.366	−0.363	−0.353
	(0.453)	(0.454)	(0.481)	(0.472)
Simple	−0.448	−0.647	−0.575	−0.611
	(0.448)	(0.480)	(0.575)	(0.519)
Very simple	−0.084	−0.276	−0.613	−0.543
	(1.228)	(1.293)	(1.071)	(1.157)
Size of Households Residence				
Sqm size/100	−1.619***	−1.929***	−1.829***	−1.843***
	(0.258)	(0.234)	(0.210)	(0.231)
Sqm size/100 squared	0.423***	0.483***	0.460***	0.468***
	(0.059)	(0.038)	(0.037)	(0.048)
Constant	2.903**	0.947	1.173	1.186
	(0.661)	(0.655)	(0.776)	(0.862)
Number of Cases	3598	3598	3598	3598
R-Square	0.142	0.130	0.144	0.140

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

C.4 The cross section of house price forecasts in 2017

In Table C.5 we test whether we can also replicate the results from Table 1 in the 2017 wave of the PHF. Using the 2017 data, we obtain the same picture as in our baseline regression: tenure and location are the single most important determinants of households' house price forecasts. Yet, there are also some differences. As we already showed in Figure 8, the difference in forecasts of renters and owners shrinks over time. Consequently, the coefficient on the "Renter" dummy variable is smaller in 2017 than in our baseline regression. In addition, it seems that controlling for community size becomes more important, which might reflect the divergence of local housing market performance over the course of the German house price boom. Last but not least, the R-squared value is smaller than in our baseline regression.

Table C.5: The cross section of house price forecasts in 2017

	(1)	(2)	(3)	(4)	(5)
Demographics, Income, Wealth					
Age					
< 30	−0.737	−0.819	−1.013	−0.970	−1.233
	(0.478)	(0.518)	(0.508)	(0.510)	(0.529)
30 - 39	0.174	0.139	−0.043	−0.115	−0.396
	(0.379)	(0.406)	(0.448)	(0.470)	(0.474)
40 - 49	−0.258	−0.307	−0.327	−0.324	−0.491***
	(0.099)	(0.128)	(0.128)	(0.131)	(0.057)
60 - 69	0.388***	0.376***	0.370***	0.424***	0.475**
	(0.042)	(0.049)	(0.041)	(0.053)	(0.112)
≥ 70	0.177	0.218	0.160	0.161	0.363
	(0.208)	(0.214)	(0.187)	(0.188)	(0.173)
Number of Household Members					
1	0.785***	0.818***	0.714**	0.516**	0.273
	(0.108)	(0.113)	(0.149)	(0.153)	(0.144)
3	0.418	0.412	0.493	0.505	0.578
	(0.413)	(0.365)	(0.369)	(0.352)	(0.413)
4	0.156	0.181	0.320	0.402	0.313
	(0.476)	(0.538)	(0.539)	(0.555)	(0.529)
≥ 5	0.587	0.667	0.840	0.901	1.150
	(0.497)	(0.492)	(0.506)	(0.543)	(0.491)
Net Wealth Quartiles					
1st Quartile	0.827	0.743	−0.316	−0.397	−0.681
	(0.643)	(0.678)	(0.516)	(0.525)	(0.607)
2nd Quartile	0.571	0.546	−0.264	−0.198	−0.249
	(0.411)	(0.388)	(0.377)	(0.400)	(0.431)
4th Quartile	−0.085	−0.145	0.105	−0.137	0.085
	(0.185)	(0.207)	(0.233)	(0.276)	(0.414)

Net Household Income Quartiles					
1st Quartile	−0.220	−0.213	−0.182	−0.108	−0.079
	(0.204)	(0.208)	(0.231)	(0.244)	(0.279)
2nd Quartile	−0.213	−0.220	−0.194	−0.155	−0.116
	(0.470)	(0.474)	(0.454)	(0.455)	(0.454)
4th Quartile	−0.297	−0.290	−0.391	−0.424	−0.482**
	(0.193)	(0.191)	(0.164)	(0.186)	(0.129)
Education					
Low	−0.095	−0.143	−0.262	−0.308	−0.532
	(0.800)	(0.810)	(0.848)	(0.835)	(0.617)
Tertiary	0.523	0.544	0.500	0.308	0.283
	(0.430)	(0.423)	(0.424)	(0.362)	(0.304)
BEHAVIORAL TRAITS					
Financial Literacy					
Very Low		−0.181	−0.033	0.074	−0.204
		(0.682)	(0.681)	(0.779)	(1.113)
Low		0.709	0.710	0.714	0.639
		(0.248)	(0.262)	(0.240)	(0.378)
Medium		−0.074	−0.005	0.030	0.039
		(0.434)	(0.432)	(0.455)	(0.485)
Patience					
Not Patient		0.604	0.543	0.524	0.529
		(0.392)	(0.414)	(0.407)	(0.303)
Very Patient		−0.064	−0.105	−0.120	−0.173
		(0.320)	(0.317)	(0.316)	(0.248)
Risk Aversion					
Below Median		0.191	0.134	0.109	0.070
		(0.297)	(0.310)	(0.325)	(0.202)
TENURE AND LOCATION					
Tenure					
Renter			1.492***	1.388**	1.028**
			(0.229)	(0.241)	(0.180)
Growth Region					
Low				−0.679***	−0.388
				(0.045)	(0.169)
Medium Low				−0.185**	−0.214
				(0.051)	(0.165)
High				1.103***	0.906***
				(0.061)	(0.036)
HOUSING/REGIONAL CHARACTERISTICS					
Time Household has been Living in Current Residence (Years)					
$t < 5$					0.877
					(0.300)
$5 \leq t < 10$					0.668
					(0.634)
$10 \leq t < 15$					0.256
					(0.284)

Size of the Local Community (Inhabitants)					
2k < 5k					0.492 (1.021)
5k < 20k					0.941 (0.801)
20k < 50k					1.137** (0.299)
50k < 100k, periphery					1.927 (0.660)
50k < 100k, center					1.045 (0.885)
100k < 500k, periphery					2.410** (0.500)
100k < 500k, center					1.291 (0.603)
≥ 500k, periphery					1.759*** (0.267)
≥ 500k, center					1.934** (0.458)
Building in Which Household Lives Needs Renovation?)					
Yes					−0.178 (0.213)
Quality of the Household's Residence (Interviewer Rating))					
Exclusive					−0.383 (0.487)
Satisfactory					−0.046 (0.267)
Simple					0.578 (0.308)
Very simple					0.256 (0.587)
Size of Households Residence)					
Sqm size/100					−1.044 (0.544)
Sqm size/100 squared					0.144 (0.076)
Constant	3.014*** (0.404)	2.793** (0.586)	2.561** (0.523)	2.629*** (0.386)	1.903** (0.502)
Number of Cases	4249	4247	4247	4247	4124
R-Square	0.020	0.027	0.039	0.061	0.087

Robust standard errors in parentheses. **: $p < 0.05$, ***: $p < 0.01$

C.5 Interactions with Age, Risk Aversion, and Financial Literacy

In Table 2 we only presented selected regression coefficients for the interaction of tenure with age, risk aversion and financial literacy. Table C.6 therefore shows the full set of regression coefficients.

Table C.6: Interactions with Age, Risk Aversion, and Financial Literacy (see Table 2)

	(1)	(2)	(3)	(4)	(5)	(6)
Tenure						
Renter	2.087***	2.372***	2.800***	2.212***	2.153***	3.277***
	(0.378)	(0.402)	(0.448)	(0.470)	(0.371)	(0.543)
Tenure × Age						
Renter × ≥ 70		−0.988				−1.139
		(0.632)				(0.637)
Owner × ≥ 70		0.275				0.292
		(0.453)				(0.453)
Tenure × Risk Aversion						
Renter × Below Median			−0.820			−0.905
			(0.501)			(0.503)
Owner × Below Median			0.961***			1.106***
			(0.297)			(0.309)
Tenure × Financial Risk Aversion						
Renter × Below Median				−0.352		−0.183
				(0.511)		(0.502)
Owner × Below Median				−0.074		−0.262
				(0.277)		(0.284)
Tenure × Financial Literacy						
Renter × Very Low					0.766	0.782
					(1.179)	(1.161)
Owner × Very Low					3.829	3.910
					(2.741)	(2.753)
Growth Region						
Low	−1.559***	−1.558***	−1.556***	−1.571***	−1.538***	−1.536***
	(0.427)	(0.426)	(0.427)	(0.427)	(0.421)	(0.419)
Medium Low	−1.450***	−1.437***	−1.470***	−1.442***	−1.423***	−1.433***
	(0.434)	(0.433)	(0.432)	(0.434)	(0.432)	(0.430)
High	1.019**	1.001**	1.012**	1.000**	1.044***	1.022***
	(0.399)	(0.399)	(0.395)	(0.400)	(0.398)	(0.395)

Demographics, Income, Wealth						
Age						
< 30	0.092	−0.062	0.190	0.148	0.106	0.069
	(0.676)	(0.683)	(0.677)	(0.684)	(0.675)	(0.685)
30 - 39	0.371	0.269	0.463	0.414	0.393	0.400
	(0.582)	(0.589)	(0.580)	(0.590)	(0.579)	(0.586)
40 - 49	−0.296	−0.324	−0.257	−0.296	−0.287	−0.279
	(0.443)	(0.443)	(0.439)	(0.442)	(0.443)	(0.438)
60 - 69	0.354	0.365	0.361	0.322	0.382	0.384
	(0.474)	(0.474)	(0.468)	(0.474)	(0.471)	(0.465)
≥ 70	−0.213		−0.246	−0.260	−0.192	
	(0.424)		(0.419)	(0.429)	(0.421)	
Number of Household Members						
1	−0.475	−0.456	−0.392	−0.432	−0.445	−0.301
	(0.435)	(0.435)	(0.428)	(0.436)	(0.434)	(0.428)
3	0.325	0.364	0.338	0.322	0.322	0.392
	(0.457)	(0.456)	(0.454)	(0.458)	(0.456)	(0.454)
4	0.336	0.385	0.394	0.315	0.336	0.463
	(0.611)	(0.612)	(0.607)	(0.613)	(0.604)	(0.599)
≥ 5	0.731	0.779	0.766	0.741	0.740	0.831
	(0.580)	(0.580)	(0.586)	(0.586)	(0.579)	(0.586)
Net Wealth Quartiles						
1st Quartile	−0.274	−0.224	−0.310	−0.350	−0.216	−0.234
	(0.553)	(0.554)	(0.551)	(0.558)	(0.557)	(0.563)
2nd Quartile	−0.159	−0.152	−0.183	−0.190	−0.169	−0.215
	(0.459)	(0.459)	(0.458)	(0.459)	(0.454)	(0.454)
4th Quartile	0.059	0.020	−0.015	0.077	0.058	−0.043
	(0.311)	(0.311)	(0.306)	(0.308)	(0.311)	(0.304)
Net Household Income Quartiles						
1st Quartile	0.623	0.600	0.598	0.601	0.578	0.524
	(0.554)	(0.554)	(0.547)	(0.553)	(0.551)	(0.542)
2nd Quartile	0.138	0.106	0.197	0.141	0.117	0.150
	(0.422)	(0.421)	(0.420)	(0.423)	(0.415)	(0.412)
4th Quartile	−0.491	−0.452	−0.485	−0.474	−0.479	−0.403
	(0.302)	(0.302)	(0.302)	(0.301)	(0.301)	(0.301)
Education						
Low	−0.426	−0.377	−0.462	−0.437	−0.396	−0.388
	(0.567)	(0.566)	(0.564)	(0.566)	(0.561)	(0.559)
Tertiary	0.062	0.072	0.035	0.081	0.072	0.071
	(0.302)	(0.302)	(0.301)	(0.302)	(0.303)	(0.301)
Behavioral Traits						
Financial Literacy						
Very Low	2.070	2.154	2.043	2.022		
	(1.435)	(1.437)	(1.431)	(1.439)		
Low	−0.752	−0.811	−0.803	−0.764	−0.766	−0.908
	(0.776)	(0.769)	(0.765)	(0.778)	(0.775)	(0.760)
Medium	−0.212	−0.201	−0.163	−0.222	−0.220	−0.167
	(0.351)	(0.351)	(0.346)	(0.353)	(0.350)	(0.345)

Patience						
Not Patient	0.415	0.445	0.389	0.429	0.445	0.465
	(0.364)	(0.365)	(0.360)	(0.367)	(0.368)	(0.364)
Very Patient	0.011	0.038	−0.008	0.010	0.037	0.058
	(0.325)	(0.327)	(0.322)	(0.327)	(0.324)	(0.325)
Risk Aversion						
Below Median	0.112	0.094			0.114	
	(0.296)	(0.295)			(0.294)	
HOUSING/REGIONAL CHARACTERISTICS						
Time Household has been Living in Current Residence (Years)						
$t < 5$	0.293	0.263	0.238	0.286	0.242	0.131
	(0.456)	(0.458)	(0.449)	(0.455)	(0.456)	(0.453)
$5 \leq t < 10$	0.573	0.582	0.494	0.562	0.527	0.435
	(0.445)	(0.444)	(0.442)	(0.445)	(0.448)	(0.445)
$10 \leq t < 15$	0.191	0.179	0.132	0.189	0.159	0.083
	(0.525)	(0.524)	(0.517)	(0.524)	(0.527)	(0.515)
Size of the Local Community (Inhabitants)						
2k < 5k	−1.591	−1.592	−1.511	−1.588	−1.562	−1.513
	(1.303)	(1.298)	(1.303)	(1.300)	(1.262)	(1.267)
5k < 20k	−0.129	−0.148	−0.125	−0.162	−0.015	−0.062
	(0.795)	(0.790)	(0.774)	(0.800)	(0.779)	(0.761)
20k < 50k	−0.117	−0.151	−0.094	−0.165	0.035	−0.029
	(0.764)	(0.758)	(0.749)	(0.777)	(0.744)	(0.738)
50k < 100k, periphery	0.680	0.614	0.710	0.680	0.809	0.758
	(0.981)	(0.978)	(0.964)	(0.967)	(0.968)	(0.941)
50k < 100k, center	0.884	0.815	0.918	0.874	0.956	0.891
	(0.984)	(0.984)	(0.967)	(0.984)	(0.968)	(0.956)
100k < 500k, periphery	0.706	0.679	0.765	0.689	0.801	0.807
	(0.745)	(0.741)	(0.730)	(0.745)	(0.724)	(0.708)
100k < 500k, center	0.607	0.540	0.701	0.589	0.699	0.694
	(0.767)	(0.760)	(0.749)	(0.778)	(0.754)	(0.737)
≥ 500k, periphery	1.112	1.069	1.188	1.096	1.226	1.241
	(0.828)	(0.826)	(0.816)	(0.836)	(0.810)	(0.801)
≥ 500k, center	1.762**	1.733**	1.790**	1.749**	1.845***	1.821***
	(0.720)	(0.715)	(0.702)	(0.726)	(0.707)	(0.690)
Building in Which Household Lives Needs Renovation						
Yes	−0.421	−0.415	−0.441	−0.398	−0.400	−0.404
	(0.389)	(0.390)	(0.385)	(0.387)	(0.389)	(0.382)
Quality of the Household's Residence (Interviewer Rating)						
Exclusive	−0.028	−0.018	0.013	−0.026	−0.017	0.037
	(0.625)	(0.627)	(0.615)	(0.627)	(0.625)	(0.617)
Satisfactory	−0.338	−0.383	−0.355	−0.353	−0.370	−0.456
	(0.387)	(0.388)	(0.382)	(0.382)	(0.385)	(0.379)
Simple	−0.448	−0.483	−0.420	−0.475	−0.479	−0.513
	(0.566)	(0.568)	(0.559)	(0.569)	(0.556)	(0.555)
Very simple	−0.084	−0.083	−0.248	−0.150	−0.011	−0.263
	(1.008)	(0.994)	(1.038)	(1.015)	(0.985)	(1.007)

Size of Households Residence						
Sqm size/100	−1.619***	−1.579***	−1.635***	−1.629***	−1.659***	−1.632***
	(0.617)	(0.609)	(0.609)	(0.619)	(0.609)	(0.587)
Sqm size/100 squared	0.423***	0.410***	0.415***	0.426***	0.427***	0.402***
	(0.127)	(0.124)	(0.123)	(0.127)	(0.126)	(0.116)
Constant	2.903**	2.810**	2.600**	2.997**	2.781**	2.413**
	(1.159)	(1.153)	(1.129)	(1.180)	(1.161)	(1.132)
Number of Cases	3598	3598	3598	3594	3598	3594
R-Square	0.142	0.143	0.147	0.142	0.143	0.152

Standard errors clustered on growth region level in parentheses.
**: $p < 0.05$, ***: $p < 0.01$

D Derivations of model equations

This appendix provides derivations for all the formulas presented in the paper.

D.1 The log-linearized first order condition

We first log-linearize the developer's first order condition to derive the linear stochastic difference equation (3) for the log price-rent ratio. The first order condition (2) is

$$V_t = E_t \left[\tilde{M}_{t+1}(V_{t+1} + 1)G_{t+1} \right].$$

This equation has a positive and finite steady state at

$$V = MG(V + 1) \quad \Leftrightarrow \quad V = \frac{MG}{1 - MG} \tag{D.1}$$

in case $0 < MG < 1$.

An alternative way to write the first order condition is

$$1 = E_t \left[\tilde{M}_{t+1} \frac{(V_{t+1} + 1)G_{t+1}}{V_t} \right] = E_t \left[\tilde{M}_{t+1} S_{t+1} \right], \tag{D.2}$$

where S_{t+1} is the return on housing investment for the developer:

$$S_{t+1} = \frac{P_{t+1} + R_{t+1}}{P_t} = \frac{(V_{t+1} + 1)G_{t+1}}{V_t}. \tag{D.3}$$

For any variable X_t, small letters denote logs and hats the deviation from steady state X

$$x_t = \log(X_t) \quad \text{and} \quad \hat{x}_t = \log(X_t) - \log(X) = x_t - x.$$

Taking logs of the return identity (D.3), we obtain

$$s_{t+1} = \log\left(1 + \exp(v_{t+1})\right) + g_{t+1} - v_t.$$

In the spirit of a Campbell-Shiller decomposition, we approximate the logarithmic term linearly in v_{t+1} around its steady state value $v = \log(V)$. This yields

$$\log\left(1 + \exp(v_{t+1})\right) \approx \log\left(1 + \exp(v)\right) + \frac{\exp(v)}{1 + \exp(v)}(v_{t+1} - v) = k + \frac{V}{1 + V} v_{t+1}$$

with $k = \log(1 + V) - \frac{V}{1+V} \log(V)$. Using the identity $MG = \frac{V}{1+V}$ from equation (D.1), we then obtain

$$s_{t+1} \approx k + MGv_{t+1} + g_{t+1} - v_t.$$

Acknowledging that this equation also needs to hold in the steady state, we can write the return identity in log-deviations from the steady state as

$$\hat{s}_{t+1} \approx MG\hat{v}_{t+1} + \hat{g}_{t+1} - \hat{v}_t. \tag{D.4}$$

Next, we turn to the first order condition (D.2)

$$1 = E_t\left[\tilde{M}_{t+1}S_{t+1}\right] = E_t\left[\exp(\tilde{m}_{t+1} + s_{t+1})\right].$$

Assuming that $\tilde{m}_{t+1}$ and s_{t+1} are jointly conditionally normally distributed, we obtain

$$\begin{aligned} 1 &= E_t\left[\exp(\tilde{m}_{t+1} + s_{t+1})\right] \\ &= \exp\Big(E_t[\tilde{m}_{t+1}] + E_t[s_{t+1}] + \frac{1}{2}\left[\mathrm{Var}_t\left(\tilde{m}_{t+1}\right) + \mathrm{Var}_t\left(s_{t+1}\right) + 2\mathrm{Cov}_t\left(\tilde{m}_{t+1}, s_{t+1}\right)\right]\Big). \end{aligned}$$

If $\tilde{m}_{t+1}$ and s_{t+1} are homoskedastic, we can subsume the conditional variances and covariances in a constant term

$$D = \frac{1}{2}\mathrm{Var}_t\left(\tilde{m}_{t+1}\right) + \frac{1}{2}\mathrm{Var}_t\left(s_{t+1}\right) + \mathrm{Cov}_t\left(\tilde{m}_{t+1}, s_{t+1}\right).$$

Taking logs and using our notation $m_t = E_t\left[\tilde{m}_{t+1}\right]$ for the predictable component of the log stochastic discount factor, we obtain

$$m_t + E_t[s_{t+1}] + D = 0,$$

or in terms of log-deviations from the steady state

$$\hat{m}_t + E_t[\hat{s}_{t+1}] = 0.$$

Substituting the approximate return identity (D.4), we can write

$$\begin{aligned} & \hat{m}_t + E_t\left[MG\hat{v}_{t+1} + \hat{g}_{t+1} - \hat{v}_t\right] \approx 0 \\ \Leftrightarrow \quad & \hat{v}_t \approx \hat{m}_t + E_t\left[MG\hat{v}_{t+1} + \hat{g}_{t+1}\right], \end{aligned}$$

which is the linear stochastic difference equation (3).

D.2 The solution for the log price-rent ratio

Suppose that $\hat{m}_t$ and $\hat{g}_t$ follow the Gaussian AR(1) dynamics (4). Our conjecture is that the solution to the linear stochastic difference equation (3) is

$$\hat{v}_t = \beta_m \hat{m}_t + \beta_g \hat{g}_t. \tag{D.5}$$

We plug this conjecture into equation (3)

$$\begin{aligned}
\beta_m \hat{m}_t + \beta_g \hat{g}_t &= \hat{m}_t + E_t \left[MG \left[\beta_m \hat{m}_{t+1} + \beta_g \hat{g}_{t+1} \right] + \hat{g}_{t+1} \right] \\
&= \hat{m}_t + MG \left[\beta_m \alpha_m \hat{m}_t + \beta_g \alpha_g \hat{g}_t \right] + \alpha_g \hat{g}_t \\
&= \left[1 + MG\beta_m \alpha_m \right] \hat{m}_t + \left[\alpha_g + MG\beta_g \alpha_g \right] \hat{g}_t,
\end{aligned}$$

where we used $E_t \left[\varepsilon^m_{t+1} \right] = E_t \left[\varepsilon^g_{t+1} \right] = 0$.

Comparing coefficients on both sides of the equation, we have

$$\begin{aligned}
\beta_m = 1 + MG\beta_m \alpha_m \quad &\Leftrightarrow \quad \beta_m = \frac{1}{1 - \alpha_m MG} \\
\beta_g = \alpha_g + MG\beta_g \alpha_g \quad &\Leftrightarrow \quad \beta_g = \frac{\alpha_g}{1 - \alpha_g MG}.
\end{aligned}$$

D.3 Aggregate state space system

We start from the definitions of the log price-rent ratio and rent growth, $\log(P_t/R_t) = \log(V_t) = v + \hat{v}_t$ and $\Delta r_t = r_t - r_{t-1} = g + \hat{g}_t$. Combined with the solution (D.5) and the dynamics (4), we can write

$$\begin{aligned}
p_t &= r_t + v + \hat{v}_t = r_t + v + \beta_m \hat{m}_t + \beta_g \hat{g}_t, \\
r_t &= r_{t-1} + g + \hat{g}_t = r_{t-1} + g + \alpha_g \hat{g}_{t-1} + \varepsilon^g_t.
\end{aligned}$$

To track the evolution of prices and rents within a state space system, the state vector x_t has to contain the current realizations of both AR(1) processes as well as the current rent level r_t (owing to the unit root nature of rents). In addition, we need to add a constant to the state vector. We can write the aggregate state space system as

$$\begin{pmatrix} p_t \\ r_t \end{pmatrix} = Bx_t \quad \text{and} \quad x_t = Ax_{t-1} + C\varepsilon_t \tag{D.6}$$

with

$$B = \begin{pmatrix} \beta_m & \beta_g & 1 & v \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad x_t = \begin{pmatrix} \hat{m}_t \\ \hat{g}_t \\ r_t \\ 1 \end{pmatrix} \tag{D.7}$$

as well as

$$A = \begin{pmatrix} \alpha_m & 0 & 0 & 0 \\ 0 & \alpha_g & 0 & 0 \\ 0 & \alpha_g & 1 & g \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad , \quad C = \begin{pmatrix} \sigma_m & 0 \\ 0 & \sigma_g \\ 0 & \sigma_g \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \varepsilon_t = N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]. \tag{D.8}$$

D.4 Characterizing the distribution of nowcasts and forecasts

On the level of the individual, the state space system reads

$$\begin{aligned} s_t^{i,h} &= Bx_t + w_t^{i,h} \quad \text{and} \\ x_t &= Ax_{t-1} + C\varepsilon_t, \end{aligned} \tag{D.9}$$

where $w_t^{i,h}$ is a 2×1 vector of idiosyncratic Gaussian shocks that are uncorrelated across agents and time, $w_t^{i,h} \sim N\left(0, \mathcal{W}^h\right)$. The covariance matrix depends on the type h of agent i. This state space system has the same state vector x_t and matrices as in (D.7) and (D.8). If we leave out the individual noise vector $w_t^{i,t}$, we again obtain the aggregate state space system (D.6). Our computations thus work with separate systems for each type. To ease notation, we suppress the dependence of $s^{i,h}$ and $x_t^{i,h}$ on type h in what follows.

Forecasts and posterior variance. To compute conditional expectations for the system (D.9), we apply the Kalman filter and thus proceed recursively. Agent i of type h computes his conditional expectation or nowcast x_t^i of x_t given the history of his observables s_t^i. It can be represented as

$$x_t = x_t^i + u_t^i,$$

where the nowcast error u_t^i is orthogonal to the history $s_t^i, s_{t-1}^i, ...$ and has conditional covariance matrix Σ_t^h. The initial nowcast x_0^i of agent i is drawn from a normal distribution with mean $\bar{x}_0^h$ and a (cross sectional) variance Ω_0^h. The mean and variance depend on the type h of agent i. Standard results on updating with normal distributions imply that x_t^i is also normally distributed.

Note that all agents share the same view on the state space system and use the same now-and forecasting rules. The only elements that are allowed to depend on type h are the initial mean $\bar{x}_0^h$ and variance Ω_0^h as well as the variance $\mathcal{W}^h$ of the signals. As a result, the covariance matrix Σ_t^h of nowcast errors is identical across individuals of the same type h, but may differ between owners and renters.

For an individual with a signal history $s_{t-1}^i, s_{t-2}^i, \ldots$ and a corresponding nowcast x_{t-1}^i, the conditional joint distribution of s_t^i and x_t one period ahead is

$$
\begin{aligned}
s_t^i\Big|_{x_{t-1}^i} &= B\left(A\left(x_{t-1}^i + u_{t-1}^i\right) + C\varepsilon_t\right) + w_t^i \\
x_t|_{x_{t-1}^i} &= A\left(x_{t-1}^i + u_{t-1}^i\right) + C\varepsilon_t.
\end{aligned}
$$

The vector $\left(s_t^{i\top}, x_t^\top\right)^\top\Big|_{x_{t-1}^i}$ is therefore conditionally normally distributed with known variance covariance matrix

$$
\begin{bmatrix}
B\left(A\Sigma_{t-1}^h A^\top + CC^\top\right)B^\top + \mathcal{W}^h & B\left(A\Sigma_{t-1}^h A^\top + CC^\top\right) \\
\left(A\Sigma_{t-1}^h A^\top + CC^\top\right)B^\top & A\Sigma_{t-1}^h A^\top + CC^\top
\end{bmatrix}.
$$

We define the gain matrix

$$
\Gamma_t^h = \left(A\Sigma_{t-1}^h A^\top + CC^\top\right)B^\top \left(B\left(A\Sigma_{t-1}^h A^\top + CC^\top\right)B^\top + \mathcal{W}^h\right)^{-1}.
$$

The gain again depends on the agent's type h via the noise variance $\mathcal{W}^h$ and its effect on the posterior, because agents only differ in their information set.

We can then write updating of the posterior mean vector and covariance matrix as

$$
x_t^i = Ax_{t-1}^i + \Gamma_t^h\left(s_t^i - BAx_{t-1}^i\right), \tag{D.10}
$$

$$
\Sigma_t^h = (I - \Gamma_t^h B)\left(A\Sigma_{t-1}^h A^\top + CC^T\right). \tag{D.11}
$$

Equation (D.10) is the stochastic difference equation (10) for the conditional mean x_t^i. Averaging across agents delivers (D.11) which is the covariance matrix (11). For stationary s_t^i, it has a stationary solution – this is the conditional expectation given the infinite history of signals. The second equation is a matrix Ricatti difference equation. Its positive definite fixed point is the posterior variance given an infinite history of signals. Once we know this fixed point, we also have a time invariant gain matrix Γ^h that summarizes how agents alter their estimates of hidden states in response to observations.

Distribution of forecasts and forecast errors. Prices and rents at date $t+1$ are Bx_{t+1} and their growth rates are

$$g_t = Bx_{t+1} - Bx_t = B\left(A - I\right)x_t + BC\varepsilon_{t+1}.$$

An individual agent's forecast of rent and price based on date t information is therefore BAx_t^i and forecasts of *growth* in price and rent are $f_t^i = B\left(A - I\right)x_t^i$. An individual agent's forecast error is

$$g_t - f_t^i = B\left(A - I\right)\left(x_t - x_t^i\right) + BC\varepsilon_{t+1}.$$

Forecast errors arise when either (i) agents are surprised by the innovation $C\varepsilon_{t+1}$ or (ii) agents have a mistaken nowcast of the hidden state based on their date t information.

To understand the cross sectional distribution of forecasts, we consider the evolution of nowcasts. We can write

$$\begin{aligned} x_t^i &= \left(I - \Gamma_t^h B\right)Ax_{t-1}^i + \Gamma_t^h s_t^i, \\ &= \left(I - \Gamma_t^h B\right)Ax_{t-1}^i + \Gamma_t^h B\left(Ax_{t-1} + C\varepsilon_t\right) + \Gamma_t^h w_t^i, \end{aligned} \tag{D.12}$$

We note that the first and third term vary in the cross section, but are orthogonal, whereas the middle term is aggregate and does not vary across individual agents.

To compute average forecasts, we use the fact that signals in equation (D.12) are unbiased. The cross sectional mean of nowcasts $\bar{x}_t^h$ for agents of type h thus evolves according to

$$\bar{x}_t^h = \left(I - \Gamma_t^h B\right)A\bar{x}_{t-1}^h + \Gamma_t^h B\left(Ax_{t-1} + C\varepsilon_t\right).$$

Agent types have different nowcast paths if they respond differently to the same news due to their using of different gain matrices.

The cross sectional distribution of forecasts (or forecast errors) depends only on the cross sectional distribution of the nowcasts (or nowcast errors). The cross sections of forecasts and forecasts errors differ only by a constant.

Denote by Ω_t^h the cross sectional variance of the nowcast for type h at date t. From (D.12), it evolves according to

$$\Omega_t^h = \left(I - \Gamma_t^h B\right)A\Omega_{t-1}^h A^\top \left(I - \Gamma_t^h B\right)^\top + \Gamma_t^h \mathcal{W}^h \Gamma_t^{h\top},$$

a Lyapunov-style difference equation. If we substitute the time invariant gain matrix Γ^h derived above, we can find a fixed point Ω^h that describes the cross sectional variance of nowcasts after an infinite sequence of signals.

Finally, let's turn back to forecasts errors. The average forecast error is

$$g_t - \bar{f}_t^h = B(A-I)(x_t - \bar{x}_t^h) + BC\varepsilon_{t+1}.$$

For the cross sectional variance of forecasts we obtain

$$\begin{aligned} E_t\left[(f_t^i - \bar{f}_t^h)(f_t^i - \bar{f}_t^h)^\top\right] &= E_t\left[B(A-I)(x_t^i - \bar{x}_t^h)(x_t^i - \bar{x}_t^h)^\top (A-I)^\top B^\top\right] \\ &= B\,(A-I)\,\Omega^h\,(A-I)^\top\,B^\top. \end{aligned}$$

Consequently, we can write the mean squared error of forecasts as

$$\begin{aligned} MSE_t^h &= E_t\left[(g_t - f_t^i)(g_t - f_t^i)^\top\right] \\ &= \underbrace{B\,(A-I)\,\Omega^h\,(A-I)^\top\,B^\top}_{\text{cross sectional variance of forecasts}} \\ &\quad + \underbrace{\Big(B(A-I)(x_t - \bar{x}_t^h) + BC\varepsilon_{t+1}\Big)\Big(B(A-I)(x_t - \bar{x}_t^h) + BC\varepsilon_{t+1}\Big)^\top}_{\text{squared average forecast error}}. \end{aligned}$$