Department of Economics
Working Paper No. 299

Tracking Owners' Sentiments: Subjective Home Values, Expectations and House Price Dynamics

Anthony Lepinteur
Sofie R. Waltl

May 2020

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

Tracking Owners' Sentiments: Subjective Home Values, Expectations and House Price Dynamics

Anthony Lepinteur[a] and Sofie R. Waltl*[b,c]

[a]University of Luxembourg
[b]Luxembourg Institute of Socio-Economic Research (LISER)
[c]Vienna University of Economics and Business (WU)

May 27, 2020

Abstract

Economic theory predicts that expectations on future house price growth are related to the current price of a house. We test this relationship for the supply side of the secondary housing market using micro data that links individual expectations to a subjective owner estimated value (OEV). We find a strong causal relationship that optimistic expectations indeed imply higher OEVs as compared to neutral or pessimistic expectations. We find qualitatively and quantitatively consistent results for Italy and the US as well as for booming and gloomy years. Our results survive ample robustness checks.
Since we use subjective data on house prices, we first show that OEVs are indeed a valid source to study house price dynamics by performing three types of convergent validity tests. We find that price dynamics derived by either combining OEVs and dwelling characteristics, or making use of repeatedly provided OEVs by the same owner over time reproduce objectively measured market trends strikingly well – even over decades. In contrast, OEVs and objective data tend to differ in levels – potentially due to psychological bias. These results hold for a large set of countries. We hence conclude that the "wisdom of the home-owner crowd" is sufficient to study house price dynamics but OEVs are less suited for measuring the level of market prices.

Keywords: Housing Markets; Expectations; Heterogeneous Beliefs; Subjective Data; Convergent Validity
JEL codes: C43; D9; G4; R31

Notes and Acknowledgements: We would like to thank Francisco Amaral, Ronan C. Lyons, Giorgia Menta and Martin Eiglsperger for providing us with additional data and information, as well as participants in the *2019 EUROSTAT International Conference on Real Estate Statistics*, the *2020 NOeG Conference* and research seminars at LISER, the Vienna University of Economics and Business, the University of Strathclyde, the Johannes Kepler University Linz and the University of Melbourne for valuable comments.
We are grateful for having received the 2020 NOeG Young Economist Award for this article.
This article uses data from the *Household Finance and Consumption Survey* (European Central Bank), the *Survey on Household Income and Wealth* (Banca d'Italia), the *American Housing Survey* (U.S. Department of Housing and Urban Development and U.S. Census Bureau), as well as the *Survey of Consumer Expectations* (Federal Reserve Bank of New York). The results published, and the related observations and analyses may not correspond to results or analyses of the data producers.

*Corresponding author: Sofie R. Waltl; Maison des Sciences Humaines; 11, Porte des Sciences; 4366 Esch-sur-Alzette/Belval; Luxembourg; sofie.waltl@liser.lu; www.sofiewaltl.com

"If molecules could talk, would physicists refuse to listen?"

Alan Blinder[1]

1 Introduction

We study how expected future changes in house prices affect currently perceived market prices. A standard model for house prices – the user cost model – predicts that, *ceteris paribus,* optimistic expectations about future capital gains justify high prices today, and *vice versa* for pessimistic expectations. In the extreme case, excessive expectations may breed a (rational) housing bubble.

We test this model for the supply side of the secondary housing market by studying homeowners' sentiments. Inline with standard theory, we find a significant positive effect of an optimistic outlook on perceived current prices. We analyze data for the US and Italy, and find that optimistic owners report on average a 6–7% higher price than owners with neutral expectations. At odds with the user cost model but consistent with the notion of *loss aversion* (Genesove and Mayer, 2001; Einiö et al., 2008; Anenberg, 2011; Chan et al., 2016), we show that pessimistic owners do not report significantly lower prices than neutral owners. While the share of optimists, pessimists and owners with neutral expectations fluctuates dramatically over the business cycle, the link we establish between expectations and perceived current prices is persistent.

Furthermore, our results survive an ample set of robustness checks ruling out the possibility that our findings may be driven by *dispositional optimism* or a *hedonic recall bias*. Using the panel structure of the Italian data, we also confirm *ex post* that optimistic owners had no extra objective information translating into higher appreciation rates. Additionally, we argue that the direction of causality is clearly one of expectations affecting estimated prices, and *not* the other way round.

We rely on survey micro data. Surveys, in particular wealth and housing surveys, usually ask homeowners to provide an estimated current market price of their home, which we call *owner-estimated value* (OEV). We select surveys that additionally ask for expected future changes in the value of the home as well as a reasonable set of structural and locational house characteristics. These characteristics allow us to control for fundamentally justified differences in house prices in a hedonic fashion (Rosen, 1974).

While we motivate our work as a test for a fundamental theory in economics, one could also put the cart before the horse: assuming the theory holds, we use micro data to test whether stated expectations on future house price growth are not simply random noise. Put that way, we confirm that stated expectations on future changes in the value of the home are indeed economically meaningful and may be used to predict behaviour.

In short, we contribute to several streams of literature. In particular, we first establish an empirical link between expected changes in the value of one's home and current prices, and thus empirically confirm a major feature of the user cost model. Second, we carefully design three types of *convergent validity tests* that demonstrate how and when OEVs are useful to study house price dynamics and for which kinds of questions using OEVs are doomed to fail. As a by-product we, third, develop tools to construct subjective house price indices, which are firmly rooted in state-of-the-art price index methodology but adapted to fit the specifities

[1] See Senik (2014).

of survey data. Since for some countries the resulting indices are even longer than currently existing objective counterparts, we provide new information on long-term trends in housing markets. Fourth, we confirm that stated expectations elicited in a survey carry economically meaningful information.

The remainder of this article is composed as follows: section 2 designs and performs three types of convergent validity tests and presents subjective house price indices for a large set of countries. In section 3, we recap the theoretical foundations of a link between expectations, and derive an econometric model fit for testing this link. Subsequently, section 4 presents the data and shows empirical results. This section also contains an ample set of robustness checks. Finally, section 5 concludes.

A comprehensive appendix provides more details about the subjective index construction techniques developed for this purpose and building on state-of-the-art index theory, all data used in this article, as well as a simulation study supporting our convergent validity findings, and additional tables.

2 Convergent Validity

2.1 Transaction Data versus OEVs

Convergent validity means that two indicators constructed to *theoretically* measure the same phenomenon are *in fact* related to each other. While, strictly speaking, transaction prices are also an imperfect data source to measure house price dynamics representative for the stock of houses, it is the closest objective piece of information to benchmark dynamics implied by OEVs. Transaction-based measures – and thereof mainly house price indices – are, despite any potential shortcomings, widely used in this context by academics, home-owners and potential buyers, professional and private investors, (monetary) policy makers, real estate agents and developers, and hence also constitute an obvious choice from a practical point of view.

We start by characterizing and discussing potential shortcomings of both, indicators based on transaction data and indicators based on OEVs.

Traditionally, housing markets have mainly been studied relying on transaction data, i.e., one observes the outcome of a bargaining process and data is only recorded once a deal is sealed. Transaction data is usually not representative for the stock of housing (see also Wallace and Meese, 1997, for the well-known Akerlof-type lemons' bias in real estate transaction data) and the volume of home transactions fluctuates substantially over the business cycle (Leamer, 2007, 2015).

In contrast to transaction data, there is no volume cycle in subjective survey data. Above, survey data comes with ample information on the property *and* the owner, is often augmented with questions on beliefs and expectations, and – importantly – is much more representative for the housing stock. Hence, survey data describes the supply side of the secondary private housing market well.[2]

Yet, survey data does not contain a well-defined and objectively justified market value of the home. The OEV reflects the owner's subjective beliefs – and maybe wishes. Specifically, in a (wealth or housing) survey, the owner is asked to anticipate the outcome of a bargaining process: she is not asked to estimate some type of "fundamental value" but really what she

[2]Ideally, we would also like to include privately owned rental property. Unfortunately, the surveys do not contain sufficient characteristics referring to other dwellings than the main residence nor is there a question on expected price changes.

thinks she could earn when the dwelling was put on sale at the day of the interview. This is anything but an easy task – in particular for non-experts.[3]

Next to a lack of knowledge and skills, responses may be affected by psychological bias. In particular, previous studies have documented a general tendency of homeowners to *over-report* the value of their home (Kish and Lansing, 1954; Goodman Jr and Ittner, 1992; Kiel and Zabel, 1999; Agarwal, 2007; Benítez-Silva et al., 2015; Gallin et al., 2018; Molloy and Nielsen, 2018). This may hint towards an *endowment effect* (Kahneman et al., 1991): owners may assign a higher value to their home than could realistically be achieved on the market – simply because they own it.

While this observed pattern is consistent with an endowment effect, these studies do not provide a causal proof. In general, only few studies explicitly examine the effects of psychological biases in the context of housing markets. Scott and Lizieri (2012) show (in a stylized lab experiment with students who did not own a house) that anchoring strongly influences judgments about house prices. There is also evidence that *loss aversion* has a significant impact on prices (Genesove and Mayer, 2001; Einiö et al., 2008; Anenberg, 2011; Chan et al., 2016): homeowners whose dwellings lost value, demand significantly higher prices. Stephens and Tyran (2012) use Danish panel data to show that *nominal* loss aversion has a severe impact on the evaluation of the advantageousness of transactions, whereas *real* losses do not play an important role. Van der Cruijsen et al. (2014) relate the accuracy of OEVs reported in a Dutch survey to the duration of ownership – which they argue tests an *endowment effect* – and the loan-to-value ratio to test for loss aversion. They report strong effects. Robert Shiller documents the importance of psychological effects for the build-up of housing bubbles as popularized in Shiller (2015).

All these effects and potential problems question the appropriateness of using OEVs to study housing markets. Understanding in which context their use is justified is hence an important prerequisite for every empirical analysis using OEVs.

2.2 Three Aspects of OEVs

We assess three aspects of OEVs, namely whether objectively measured dynamics in house prices are reproduced when *(i)* linking OEVs and dwelling characteristics in a "hedonic" way, *(ii)* combining repeatedly provided OEVs by the same owner at different points in time, and *(iii)* linking subjective and objective prices available for different points in time.

All three tests follow a *wisdom of the crowd principle.* We aim to test whether the collective knowledge and perceptions of many individuals reproduce objectively observable market trends.

The first aspect is fundamental whenever dwelling characteristics are used to control for fundamentally justified differences in house prices. The theoretical link between prices and characteristics follows from a hedonic valuation approach (Rosen, 1974) and is empirically well established for transaction data. The hedonic valuation approach assumes that differentiated products are completely described by a vector of objectively measured characteristics and that observed prices of these goods are composed of implicit or "hedonic" prices associated with these characteristics. Due to the large degree of heterogeneity among dwellings, hedonic valuation approaches are widely used whenever dealing with house prices. In this context, the relevant characteristics describe both, the major features of the structure and the attractiveness of the dwelling's location.

[3]When a person plans to buy or sell property, she is likely to seek advise from real estate experts. It is, however, more than unlikely that such costly actions are taken when preparing for a survey interview.

The second aspect is crucial when relying on repeatedly provided OEVs by the same owner to guarantee that like is compared with like. Relying on repeated observations is beneficial as dwelling characteristics – and in the case of OEVs also individual characteristics of the owner – remain (largely) the same over time. Hence, repeated OEVs constitute price pairs of comparable goods, which are not confounded by differences in characteristics. The repeat-sales index construction methodology builds upon this idea (Bailey et al., 1963; Case and Shiller, 1987, 1989). This second convergent validity test checks whether market price dynamics are reproduced by repeated OEVs.

The third aspect establishes an individual link between transaction prices, namely the acquisition price and OEVs. If the two types of data are comparable in levels, then linking the two can be considered equivalent to linking repeated transaction prices as in the repeat-sales methodology. The findings of previous research documenting owners' tendency to over-report as compared to market prices creates some prior doubts on the success of this test. An acquisition price can be interpreted as a *revealed* market price, while an OEV is a guess about an *uncertain and hypothetical* market price. However, as acquisition prices are also self-reported they may be affected by similar psychological biases as OEVs. In this sense, we check whether OEVs are *systematically different* to *(self-reported) acquisition prices*.

The three convergent validity tests are set up in the following way: we estimate three different types of *subjective residential property price indices (S-RPPI)* making use of OEVs and compare them to *objective counterparts* based on transaction data *(O-RPPIs)*.

The first index is a *(i) hedonic index* that combines OEVs with house characteristics. The second index is called a *(ii) repeat-observations panel index (ROP)* and makes use of price pairs provided by the same owner for the same dwelling at different points in time. We name the third index a *(iii) repeat-observations acquisition index (ROA)*, which links more objective acquisition prices to subjective OEVs. Whenever a subjective index reproduces trends measured by O-RPPIs, we conclude that the data used to compile the subjective index is internally consistent.

The index construction methodologies are described in detail in Appendix A. The following subsection 2.3 reports the convergent validity test results.

2.3 Test Results

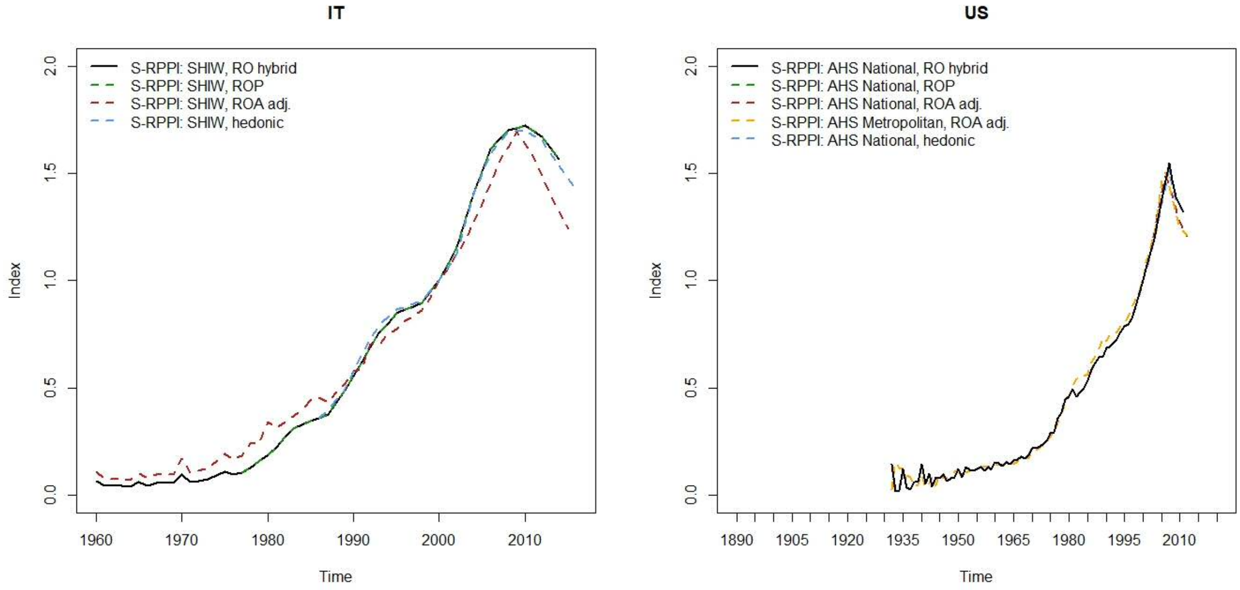
We use two exceptionally comprehensive surveys, which allow us to compile the three main types of subjective indices: hedonic, ROP and ROA (see Appendix A for methodological details). Specifically, we use subjective house price data from the *Italian Survey on Household Income and Wealth (SHIW)* and the *American Housing Survey (AHS)* for this purpose. The data sets are described in detail in Appendix B.

Both surveys have a panel component enabling us to compile ROPs and provide sufficient information on dwelling characteristics and location to estimate hedonic models. As described in detail in Appendix B, there is a national and metropolitan version of the AHS. Only the national version has a panel component and since the mix of geographical areas captured by the metropolitan version changes from survey wave to survey wave, a compilation of hedonic indices is problematic. We hence focus generally on the national sample and compile exclusively ROA indices for the metropolitan sample.

2.3.1 OEVs and Dwelling Characteristics, and Repeated OEVs

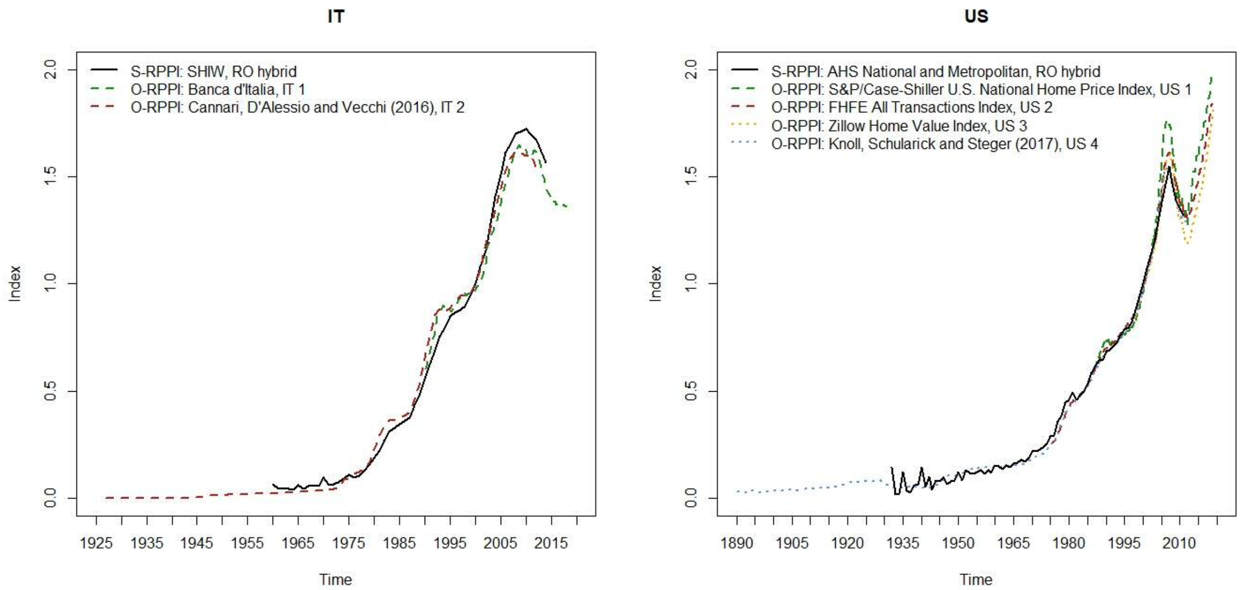
Table 1 reports pairwise Pearson correlation coefficients for annualized changes in S-RPPIs. ROP and hedonic indices show strikingly high correlations, which reassuringly indicates that

Figure 1: S-RPPIs.



Notes: The figures compare several types of S-RPPIs. All indices are normalized to 2000.

Figure 2: S-RPPIs versus O-RPPIs.



Notes: The figures compare the subjective RO hybrid index to O-RPPIs. Details about the O-RPPIs used are documented in Appendix C. All indices are normalized to 2000.

both indices elicit essentially the same house price dynamics (see also Figure 1).

Comparing the subjective ROP and hedonic indices to O-RPPIs, we find again strikingly high correlations ranging between 0.81 and 0.96 as reported in Table 2 (see also Figure 2). This is a remarkable result as the various O-RPPIs we use (which are described in more detail in Appendix C) differ in terms of geographical coverage, data used (sales prices, appraised prices, imputed prices) and index construction methodology applied.

Table 1: Correlations: Changes in S-RPPIs

Italy

	Hedonic	ROP	ROA	ROA adj.	RO hybrid
Hedonic	1	0.925	0.084	0.557	0.925
ROP	–	1	0.128	0.717	1.000
ROA	–	–	1	0.142	0.128
ROA adj.	–	–	–	1	0.707
RO hybrid	–	–	–	–	1

US

	Hedonic	ROP	ROA national	ROA adj. national	ROA metro	ROA adj. metro	RO hybrid
Hedonic	1	0.998	0.498	0.856	0.727	0.823	0.997
ROP	–	1	0.498	0.853	0.730	0.836	0.999
ROA national	–	–	1	0.415	0.858	0.466	0.504
ROA national adj.	–	–	–	1	0.495	0.926	0.852
ROA metro	–	–	–	–	1	0.561	0.765
ROA metro adj.	–	–	–	–	–	1	0.837
RO hybrid	–	–	–	–	–	–	1

Notes: The table reports pairwise Pearson correlation coefficients for annualized changes in S-RPPIs for the time period where all indices are available. Italian results use the SHIW; the time period considered is 1987 to 2014. The US results use the AHS national and metropolitan survey; the time period considered is 1999 to 2013. The metropolitan sample has no panel component and hence no ROP or RO hybrid indices can be computed. The hedonic index uses the national sample as each wave refers to a different set of metropolitan areas, which confounds the two-period hedonic models.

We hence firmly conclude that subjective ROP and hedonic indices survive our convergent validity test. This means, that the relationship between dwelling characteristics and OEVs, and repeated OEVs provided by the same owner are internally consistent: the market dynamics inherent in these data are coherent with objectively measured market trends.

2.3.2 OEVs and Acquisition Prices

In contrast, ROA indices are only weakly related to other subjective or objective indices (see Table 1 and Table 2). For Italy, the Pearson correlation coefficients are not even statistical significantly different from zero.

The source of these weak correlations becomes evident from Figure 3, which plots the ROA index: in survey years, there are spikes in the index, while in non-survey years the index returns to its original price path. Such a pattern – clearly not a realistic feature of the housing market – emerges when the prices reported for survey years, i.e., the OEVs, are systematically higher than the acquisition prices that mainly relate to non-survey years. We confirm this in a simulation exercise calibrated to the US housing market (see Appendix D).

In the US, the spikes are less pronounced than in Italy. Still, in every single survey year, the index jumps up.

[4] Let (o_t, s_t) denote the change in an objective index o_t and an subjective index (s_t) in period t. We obtain bootstrap confidence intervals from the following sampling strategy: we re-sampling 100 pairs with replacement from $(o_t, s_t)_{t=1}^T$ and compute a Pearson correlation coefficient. We repeat this 10,000 times and compute the confidence interval from the resulting distribution of ρ.

Table 2: Correlations: Changes in S-RPPIs vs. Changes in O-RPPIs

Italy

	IT 1 (Banca d'Italia)			IT 2 (Cannari et al., 2016)		
	ρ	95% CI		ρ	95% CI	
Hedonic	0.817	[0.750;	0.871]	0.811	[0.739;	0.867]
ROP	0.816	[0.753;	0.866]	0.837	[0.781;	0.882]
ROA	0.111	[-0.087;	0.296]	0.066	[-0.079;	0.217]
ROA adj.	0.658	[0.495;	0.793]	0.607	[0.478;	0.710]
RO hybrid	0.816	[0.753;	0.866]	0.837	[0.779;	0.882]

US

	US 1 (Case & Shiller)			US 2 (FHFE)			US 3 (Zillow)		
	ρ	95% CI		ρ	95% CI		ρ	95% CI	
Hedonic	0.910	[0.862;	0.950]	0.939	[0.908;	0.963]	0.959	[0.939;	0.974]
ROP	0.916	[0.869;	0.954]	0.945	[0.917;	0.968]	0.962	[0.944;	0.976]
ROA national	0.446	[0.289;	0.581]	0.472	[0.310;	0.609]	0.411	[0.252;	0.555]
ROA national adj.	0.937	[0.926;	0.950]	0.937	[0.918;	0.955]	0.924	[0.902;	0.946]
ROA metro	0.655	[0.519;	0.757]	0.663	[0.543;	0.758]	0.650	[0.540;	0.738]
ROA metro adj.	0.937	[0.923;	0.952]	0.937	[0.918;	0.954]	0.886	[0.849;	0.917]
RO hybrid	0.917	[0.867;	0.956]	0.948	[0.919;	0.971]	0.963	[0.945;	0.978]

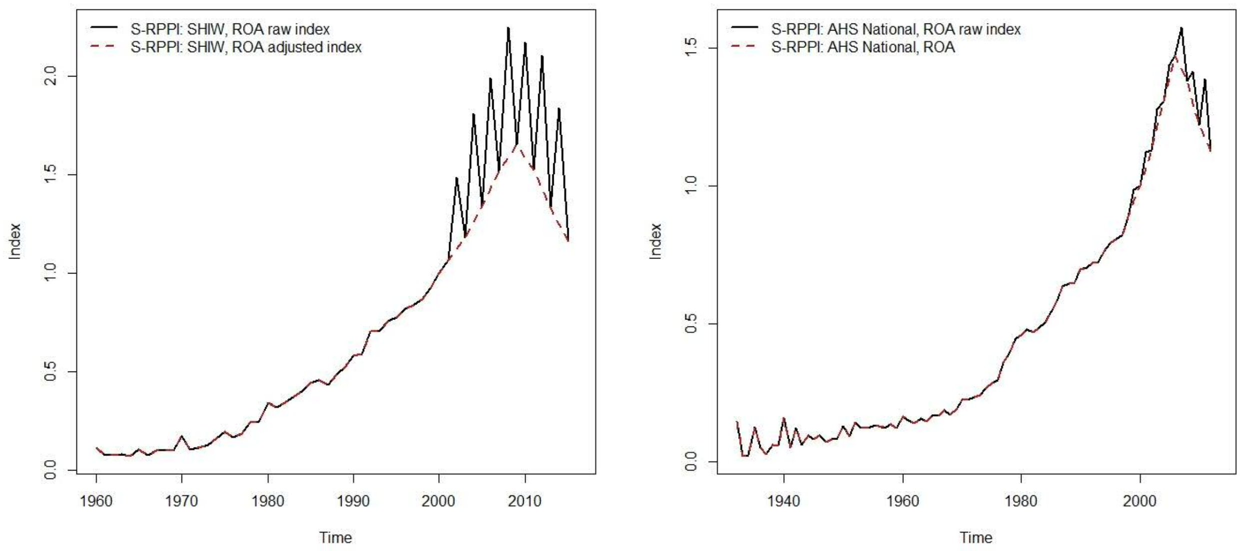
Notes: Pearson correlation coefficients ρ for O-RPPIs and S-RPPIs are reported together with 95% bootstrap confidence intervals.[4] We use data for the time period, where all indices are available. For IT 1, this period is 1991 to 2014, and for IT 2, 1987 to 2012. In the US, this is always 2000 to 2011. Higher frequency indices are annualized by taking averages. As the later years of the index by Knoll et al. (2017) coincides with the FHFE index, we do not report results here explicitly. ROA indices include dwellings that were acquired in survey years, when excluding them Pearson correlation coefficients drop by roughly 0.06 to 0.07 in the US but drop by less than 0.001 in Italy.

This finding is consistent with previous studies that document a tendency of owners to provide overly optimistic price estimates. While our data does not allow us to explicitly test for the *source* of such a bias – e.g., if it was indeed an endowment effect, we would need a measure of loss aversion, which is not available in survey data – our findings are *consistent* with an endowment effect: owners' may put extra value on their home simply because they own it.

Another potential explanation would be that owners collectively believe that their homes appreciated more than those on the market or, put differently, their assets outperformed the market. This belief is different as compared to a belief that the intrinsic value of their home is higher. We can rule out this explanation: if it was a belief about higher appreciation rates, these would also be reflected in ROP indices. ROP indices, however, match objective indices almost perfectly. Also, we simulate this intuition and find that this explanation, in contrast to a one-time premium consistent with an endowment effect, does not reproduce the observed pattern (see again Appendix D).

A third explanation focuses on differences in tastes of owners and "the market." Heston and Nakamura (2009) report that owners tend to provide significantly higher estimates of the hypothetical market rent their home could achieve as compared to observed market rents. One explanation they provide is that owners may place greater value than the market on certain features of their home, i.e., that owners' subjective hedonic shadow prices associated with certain amenities are higher than the "market shadow price" and that the taste of owners is different than the taste of renters. If that was indeed the prime explanation, this must already be re-

Figure 3: ROA Indices: Combining Objective and Subjective Prices.



Notes: The figure compares a raw ROA index to a ROA index, where index numbers for survey years are discarded and replaces by interpolated values. Survey years in Italy (left panel): 2002, 2004, 2006, 2008, 2010, 2012, 2014 and 2016; survey years in the US (right panel): 1999, 2001, 2003, 2005, 2007, 2009, 2011 and 2013. Indices are normalized to the year 2000. Dwellings that were acquired in survey years are discarded.

flected in the acquisition price. Some owners certainly remodel their acquisition after purchase to make it fit their taste. However, if that happened systematically, we would again need to see this in ROP indices, which we do not.

Regardless of the reason, we conclude that OEVs are systematically higher than objective market prices. Hence, our third convergent validity test fails and we decided against using acquisition prices in our later analyses. We highly recommend not to blindly mix OEVs and acquisition prices in empirical research.

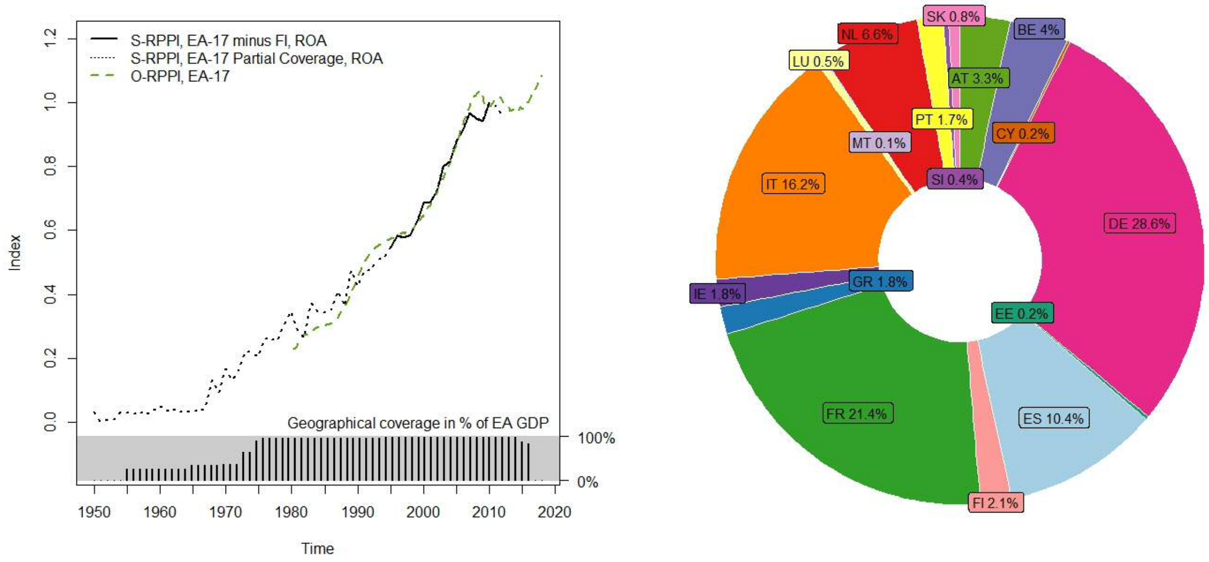
2.4 The Usefulness of S-RPPIs

We compile S-RPPIs to perform convergent validity tests. Yet, they are also of interest by themselves: in the absence of (trustworthy) O-RPPIs or whenever O-RPPIs are only available for short periods of time, S-RPPIs may be a valuable source of information describing housing market trends.

While the spikes observed in ROA indices certainly do not reflect market trends, leaving out these years and interpolating between survey years leads to a less frequent and precise index, but the result is still well capable to reproduce *overall trends*. This partly interpolated index is called the “ROA adjusted” index and shown in Figure 1 and Figure 3. The correlations with other indices is not as high as for the cleaner ROP and hedonic indices, but still over 0.6 for Italy and even around 0.9 in the US (see Table 1 and Table 2).

A “best-off” index hence exploits the ROA index’ advantage of being long *as well as* the ROP (or hedonic) index’ advantage of being more precise in survey years. This *hybrid RO index* equals the ROP index whenever it exists and extends it to the past using the ROA index. Methodological details are again explained in Appendix A.

Figure 4: An S-RPPI for the Euro Area.



Notes: The official O-RPPI for the Euro area is a weighted index of country-specific O-RPPI and compiled by the ECB. Weights are determined by a country's share of Euro area GDP. We replicate a S-RPPI for the Euro area and use the 2013 GDP in current prices, which is composed as shown in the right panel. For Finland, no acquisition prices are collected and hence no S-RPPI is computed and the Euro area S-RPPI misses Finland's contribution. Since S-RPPIs are of different lengths, only between 1995 and 2010 the index is representative for the EA-17 (except Finland). The coverage in terms of Euro area GDP is indicated by the bars at the bottom of the left panel.

To rule out that our results may only be valid for Italy and the US, we extend our analysis here to several European countries. We therefore make use of the pan-European *Household Finance and Consumption Survey (HFCS)*. As we have shown that an adjusted ROA-index correlates with the other subjective indices, we compute such indices for all countries participating in the HFCS and benchmark them against O-RPPIs.

Although the adjusted ROA methodology is not our preferred one, the information available in the HFCS only allows us to construct such indices. On top, the HFCS is a "young survey" with currently only two waves available. These are probably the worst circumstances to estimate a subjective index and, therefore, – if our indices still track objective counterparts – this setting provides a very strong sensitivity analysis.

Again, we find that S-RPPIs – even those relying on very small numbers of observations – reproduce overall market trends strikingly well. The full set of results is reported in Appendix A.

Figure 4 shows an aggregated S-RPPI for the Euro area. We aggregate country-specific S-RPPIs by weighting them with the country's share of Euro area GDP. This methods mimics the aggregation procedure applied by the European Central Bank to obtain a Euro Area O-RPPI. Even on this supra-aggregated level, S-RPPIs and O-RPPIs match strikingly well.

Hence, we conclude that our findings are not particular to Italy or the US, but generalize to a large set of countries. These countries experienced quite different price dynamics in the past, which further increases confidence in the persistence and general nature of our results.

3 The Link between Expectations and House Prices

The remaining of the paper aims at illustrating how OEVs, once linked to dwelling characteristics in an hedonic fashion, can be used at the individual level to explore the relationship between expectations and house prices through the lenses of the *user cost model*.

3.1 Theoretical Considerations

The *user cost model* constitutes a standard model for house prices. It equates the *total cost of owning a house* for one period with the *benefit of owning*. This benefit is empirically usually associated with the rent (Poterba, 1984; Himmelberg et al., 2005). Glaeser and Nathanson (2015) point out that it could also be interpreted as the *benefit of living in a certain area* with all amenities and (income) opportunities this location provides. This interpretation is, for instance, followed by Head et al. (2014) and Glaeser et al. (2014). Our approach is valid for both interpretations, but the way we control for differences in the fundamental value empirically aligns best with the latter one.

From a simple inter-temporal no-arbitrage condition, it follows that in equilibrium the user cost must equal the benefit. With P_t denoting the house price in period t and R_t the benefit, respectively, a stylized version of the user cost formula (Himmelberg et al., 2005) can be written as:

$$P_t \cdot u_t = P_t \cdot (r_t + \delta_t + \omega_t + \gamma_t - g_t) = R_t,$$

where $u_t = r_t + \delta_t + \omega_t + \gamma_t - g_t$ is the per-dollar user cost with r_t denoting the interest rate, δ_t depreciation net of maintenance, ω_t running and average transaction costs (including taxes), γ_t a risk premium, and $g_t = \mathbb{E}_t\left[\frac{P_{t+1}-P_t}{P_t}\right]$ expected capital gains (or losses). An additional term is included when owners can tax-deduct mortgage interest payments.

Denoting $u_t^- = u_t + g_t$ the cost excluding expected capital gains, the user cost formula can be written in the form of a linear asset pricing model with discount factor u_t^-

$$P_t = \frac{R_t}{(1+u_t^-)} + \frac{\mathbb{E}_t\left[P_{t+1}\right]}{(1+u_t^-)}. \tag{1}$$

Hence, the current house price equals the discounted sum of the benefit and the expected future house price, and one concludes that *ceteris paribus* high expected future capital gains justify high prices today relative to the benefit. *Vice versa*, expected falling prices justify lower current prices.

In the extreme case, excessive expectations may even breed a (rational) housing bubble. Assuming a constant discount factor $u_t^- \equiv u^-$ and solving the recursion by forward iteration yields

$$P_t = \sum_{i=0}^{\infty} \frac{\mathbb{E}_t[R_{t+i}]}{(1+u^-)^{i+1}} + \lim_{j\to\infty} \frac{\mathbb{E}_t P_{t+j}}{(1+u^-)^{j+1}}. \tag{2}$$

The first component $\sum_{i=0}^{\infty} \frac{\mathbb{E}_t[R_{t+i}]}{(1+u^-)^{i+1}}$ is the present value of future benefits and hence constitutes the fundamental value of the home. The term $\lim_{j\to\infty} \frac{\mathbb{E}_t P_{t+j}}{(1+u^-)^{j+1}}$ – if not assumed to equal zero – is a rational bubble that purely depends on expected future price appreciation (Blanchard and Watson, 1982).

In the words of Stiglitz (1990), an asset bubble exists "if the reason that the price is high today is only because investors believe that the selling price is high tomorrow – when fundamental factors do not seem to justify such a price." Similarly, Himmelberg et al. (2005) write, "we think of a housing bubble as being driven by homebuyers who are willing to pay inflated prices for houses today because they expect unrealistically high housing appreciation in the future."

While equation (2) relates market prices to market expectations, the latter two definitions of housing bubbles refer to the *demand side* of the housing market and describe how excessive expectations among investors or buyers may drive up prices. Also previous empirical research clearly sets the focus there: Piazzesi and Schneider (2009) find an increasing share of optimists during the built-up of the US housing bubble. Optimism is measured on the demand side: the survey they use asks whether it is a good idea to buy a house given the current market conditions. Similarly, Case and Shiller (1988) surveyed recent home-buyers in four regions in the US. They find that high expectations about future house price growth was an important motivating factor for buying a home and that the share of buyers that believed that house prices in their area will increase was almost 100% in booming markets. Case et al. (2012) document how expectations among recent home-buyers co-move with the local housing cycle. In particular, long-term expectations reached enormous levels at the peak of the US boom around 2005.

These studies provide evidence for a link – albeit not on the micro level – between expectations and the *willingness-to-pay (WTP)*. But what about the supply side? Standard economic theory claims that with rational market participants (at least in the absence of income constraints) the WTP equals the *willingness-to-accept (WTA)*. WTA and WTP coincide at the fair market value (Willig, 1976). However, experiments consistently find WTP-WTA gaps potentially driven by an endowment effect (Thaler, 1980; Kahneman et al., 1990). Our results from subsection 2.3 suggest that such a gap may be large in the secondary housing market where homeowners constitute the supply side. This is also supported by Bao and Gong (2016). In surveys, homeowners estimate the price of their home and hence the OEV can be expected to reflect owners' WTA.

It is important to note that the existence of a potential endowment effect has no implications for the link between prices and expectations as long as an owner's $\rho > 0$ is roughly constant over time.[5] This can be seen by writing the WTA as

$$P_t^{WTA} = (1+\rho) \cdot P_t, \tag{3}$$

with a constant endowment premium $\rho > 0$. Replacing the market price by owners' WTA in the user cost model yields a valuation formula for owners,

$$P_t^{WTA} = \frac{R_t}{(1+u^-)} + \frac{\mathbb{E}_t\left[P_{t+1}^{WTA}\right]}{(1+u^-)} \quad \text{or equivalently} \quad P_t = \frac{(1+\rho)\cdot R_t}{(1+u^-)} + \frac{\mathbb{E}_t\left[P_{t+1}\right]}{(1+u^-)}. \tag{4}$$

Hence, an endowment effect can be interpreted as a larger housing benefit for current owners as compared to buyers. This is consistent with the finding of Heston and Nakamura (2009), who document that home-owners report higher hypothetical rents for their homes as compared to market rents.

In the presence of an endowment effect and owners being unwilling to sell below their inflated price perception, buyers need to have higher expectations about future price appreciation than owners to compensate for the latter. Put differently, those people deciding to buy may be those having particularly high expectations.

To show this, we write buyers' expectations g_t^B as owners' expectations g_t plus a mark-up $\eta > 0$,

$$g_t^B = g_t + \eta,$$

[5] Alternatively, one could interpret ρ also as the hurdle of moving owners anticipate. Again, this interpretation fits our model framework.

and the valuation formula on the demand side as

$$P_t^{WTP}(u^- - g_t^B) = R_t.$$

Transactions take place, when $P_t^{WTA} = P_t^{WTP}$. Plugging in owners' and buyers' valuation formulas and solving for η yields

$$\eta = \frac{R_t}{P_t} \cdot \frac{\rho \cdot (2+\rho)}{1+\rho}.$$

There are three implications: first, $\rho = 0$ implies $\eta = 0$, i.e., in the absence of an endowment effect, no inflated expectations on the buyers' side are needed to seal a deal. Second, the larger the endowment effect, the larger the mark-up η needs to be. Third, the mark-up is proportional to the rental yield, $\frac{R_t}{P_t}$. Thus, when the return on housing increases, the mark-up needed to enable transactions rises too.

Hence, there are good reasons to belief that sentiments on the supply side may differ to sentiments on the demand side, and, thus, need to be studied separately.

3.2 Econometric strategy

We test for a causal relationship of expectations and OEVs. We do this on an individual level: a particular OEV is linked to the owners' personal expectations.

Our econometric set-up directly derives from the *user cost model*. As every house is different, it would be quite shortsighted to relate the OEV to stated expectations only. We rather also control for dwelling and owner characteristics that justify differences in prices. In other words, we include the variable of interest into a hedonic pricing model (Rosen, 1974):

$$\log(OEV_{it}) = \alpha_0 + \alpha_1 Optimism_{it} + \alpha_2 Pessimism_{it} + \alpha_3 House_{it} + \alpha_4 Indiv_{it} + \varepsilon_{it}, \qquad (5)$$

where $Optimism_{it}$ and $Pessimism_{it}$ are dummies that equal 1 if the respondent is respectively considered as optimist or pessimist (see subsection 4.1.1 and 4.1.2 and Appendix B for more details on the way we measure $Optimism_{it}$ and $Pessimism_{it}$ empirically). We use *neutral* individuals as the reference category.

The user cost model described in subsection 3.1 predicts that $\alpha_1 > 0$ and $\alpha_2 < 0$.

Importantly, Equation 5 includes hedonic dwelling characteristics $House_{it}$ describing the structure as well as the location. These aim to capture fundamentally justified differences in house prices. Additionally, we include socio-economic and demographic characteristics of the owner $Indiv_{it}$. These characteristics – and in particular income – are expected to proxy aspects of unobserved differences in the quality of the structure. For instance, the higher the income the more an owner can and may invest in maintenance, which creates differences in the quality not related to the age of the structure that might in return influence expectations.

Thereby, it is important to note that we do not aim at describing the fundamental value comprehensively and – as we estimate shadow prices from subjective data – we do not pretend to get the level right. In fact, due to the results in subsection 2.3 we expect that OEVs are systematically higher than the market value. Our mere goal is to test whether differences in expectations translate into differences in OEVs. As shown in subsection 3.1, an inflated level among OEVs as compared to market prices has no impact on the theoretical relationship between expectations and OEVs. From an empirical perspective, differences between OEVs and market prices should not influence the estimation of α_1 and α_2 under the assumption that they are orthogonal to the vector of individual expectations, conditionally on the vectors $House_{it}$ and $Indiv_{it}$. However, to control for fundamentally justified variation in house prices, it is crucial that OEVs and dwelling characteristics are linked in a hedonic way. This is confirmed by our first convergent validity test.

Additionally, we need to minimize the potential that our measure of expectations could be correlated with unobserved fundamental factors driving differences in prices. With this regard, the inclusion of locational characteristics serves a second function: if there are justified reasons to assume that a certain area will become more attractive – as, for instance, a new company announced to move into the area, or a large infrastructure project is about to improve public transportation – such an expected price increase is shared in the same area and is captured by locational characteristics. With this regard, the individual characteristics too serve a second function: certain groups of people – e.g., certain age groups – may report systematically different prices than others. We do not want our results to be driven by these factors and the individual controls should capture any such effects. As a robustness check (see subsection 4.4), we also include more subjective characteristics for the same reason.

4 Empirical Analyses

4.1 Data

We searched for survey data including an OEV, a reasonable set of individual and hedonic house characteristics – in particular information on the location of the dwelling –, and a question eliciting expected future changes in the value of the home.[6]

The US Federal Reserve Bank of New York's *Survey of Consumer Expectations (SCE)* and the Italian *Survey on Household Income and Wealth (SHIW)* fulfill these criteria and are presented in the following (for even more details see Appendix B).[7]

4.1.1 Federal Reserve Bank of New York's Survey of Consumer Expectations – Housing Module

We base our empirical analyses of the U.S. housing market on data coming from the housing module of the SCE. Active members who had participated in a SCE monthly survey in the prior eleven months were invited to participate in the housing module. The survey has been conducted every year since 2014 and 4,641 respondents participated between 2014 and 2017.

The housing module does not have a longitudinal component but it contains rich blocks of questions. The respondents are asked, among other things, about their expectations regarding future home price changes, financing conditions, past housing-related behaviour (such as buying a home and debts taken out to purchase real estate), and the likelihood of buying a home in the future. Respondents also provide information about house characteristics, their household income, and many other demographic variables. When appropriate, questions had built-in logical checks (for instance, percent chances of an exhaustive set of events had to sum to 100). Item non-response is extremely rare, and almost never exceeds one percent for any question.

Owners in the housing module are also asked to report the current and the expected values of their home via the following questions: *What do you think your home would sell for today? And in one year from today?*. We use the response to the first question to measure OEVs. We

[6]As we consistently follow a micro approach here, surveys with questions about house prices expectations in general are discarded.

[7]Unfortunately, the AHS used for convergent validity tests does not include questions on expectations. In contrast, the SCE lacks a panel component and an acquisition price. Hence, we decided to use the two surveys complementarily. The HFCS core questionnaire does not include a question on expectations and generally lacks geographical information.

assume that an owner has optimistic expectations if the future value of the house is at least 2% higher than its current value. Conversely, an owner has pessimistic expectations if the future value of the house is at least 2% lower than its current value. We show in Appendix E that our conclusions are not sensitive to this specific threshold.

4.1.2 Survey on Household Income and Wealth

We use the SHIW as data source for Italy. Albeit all waves from 2010 onward include a question on expected future house prices, the phrasing of the question, the proposed answers and the type of house prices referred to differ in every single survey wave. There are strongly framed questions formulated in terms of price decreases or increases and questions phrased in neutral language. There are also simple questions asking whether prices are likely to go up, stay the same or fall, and very complex versions asking for a full distribution of beliefs. Questions sometimes refer to house prices in general, to the price in the household's neighbourhood or to the value of the household's home directly.

We make use of the surveys fielded in 2014 and 2016 because they share a set of questions measuring the house value at the time of the interview as well as the anticipated value in 12 months. In both waves, the current value of the house is elicited using the following question: *In your opinion, how much is your house/flat worth (unoccupied)? In other words, what price could you ask for it today (including any cellar, garage or attic)?.*

Expectations regarding the future price of the main residence were elicited differently in 2014 and 2016. In 2014, respondents are asked the following questions: *In your opinion, at the end of 2015, the value of your main dwelling will be*

- *higher,*
- *the same,*
- *lower.*

We consider here that an owner is an optimist (pessimist) if she thinks that that the value of the main residence will be *"higher"* (*"lower"*).

In 2016, half of the respondents were asked to respond to the following question: *In your opinion, how will the value of your principal residence change? Assign a total of 100 points among the options below according to which you think is more likely: give high points to the most likely and low points to the least likely. Over the next 12 months the value of your residence will:*

- *Decrease by over 5%*
- *Decrease between 5 and 2%*
- *Fluctuate between -2% and +2%*
- *Increase between 2 and 5%*
- *Increase by over 5%*

For the other half, the amount in Euro of a, say, 5% increase were computed and presented to the respondent. The amounts use the self-reported value of the property as basis.

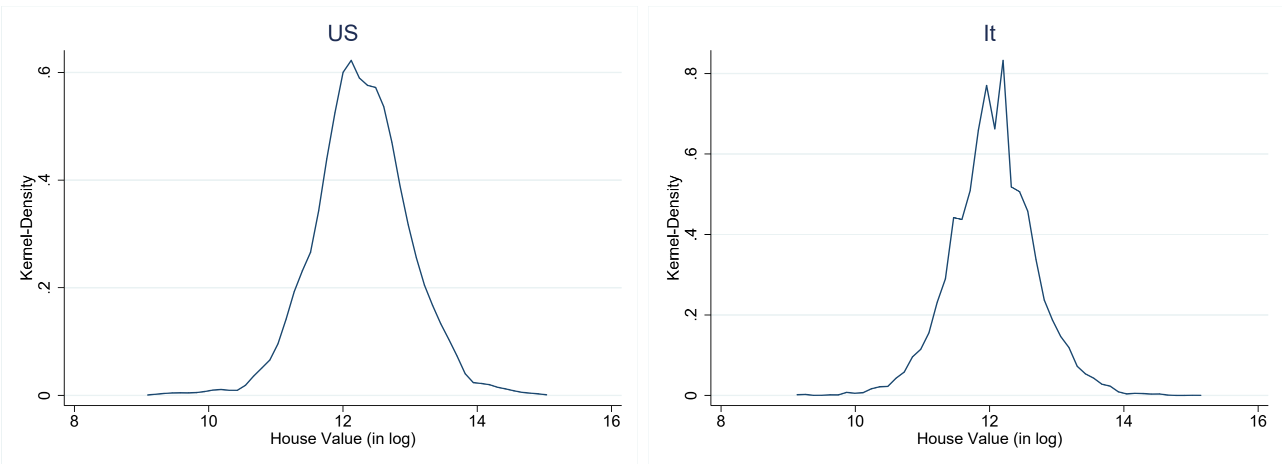
To remain conssitent with our analysis using the US data, an owner is called an optimist (pessimist) if she assigns strictly more points to the options *"Increase between 2 and 5%"* and *"Increase by over 5%"* (*"Decrease by over 5%"* and *"Decrease between 5 and 2%"*) than to the other options.

4.1.3 Estimation Samples

We restrict our analysis in both datasets to homeowners between age 20 and 90 with valid information on house value, expectations and control variables. It produces estimation samples of 2,661 and 10,646 observations respectively for the U.S. and for Italy.

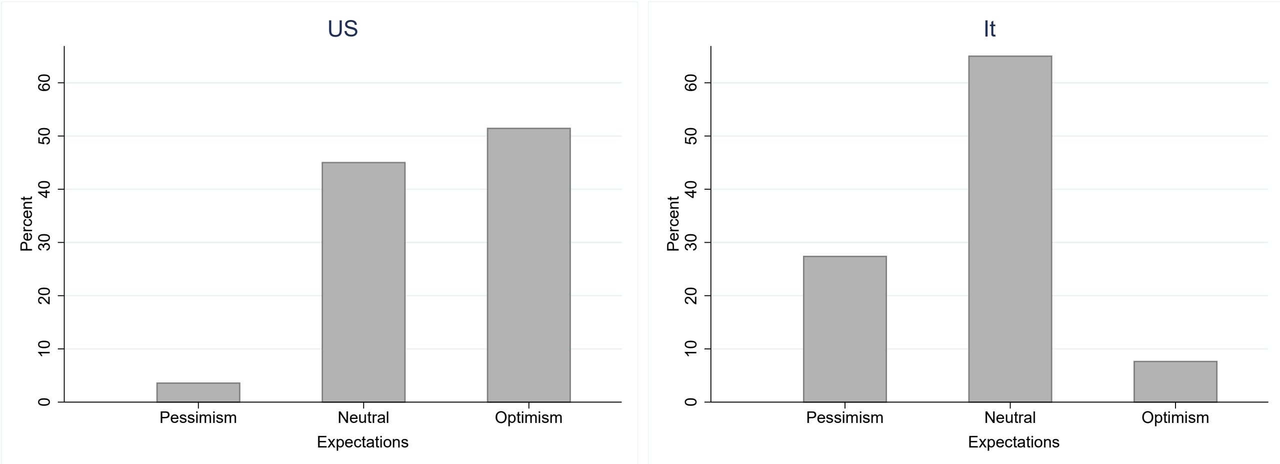
Figure 5 and Figure 6 show the kernel density of the self-reported house value (in log) and the distribution of expectations in the US and Italy. While the self-reported house values are log-normally distributed in both estimation samples the distributions of expectations are remarkably different across datasets. On one hand, almost half of the respondents of the U.S. sample report optimistic expectations and more than 40% report neutral expectations. On the other hand, more than 60% of the Italian homeowners have neutral expectations and around 30% of them are pessimistic. Such differences in expectations are actually not surprising. In Figure 7, we show the housing and business cycles respectively for the US and in Italy. The grey area indicates when the data of the Housing module of SCE and the SHIW have been collected. As revealed in Figure 7, the interviews for the Housing module of SCE took place in a booming period (2014-2017) both in terms of GDP and housing prices. On the contrary, Figure 7 shows that the 2014-SHIW data has been collected during a period of recession on the housing market. Note also that the GDP in 2014 was equal to its 2000-level. The 2016-SHIW was fielded in a slightly different period: the housing market was relatively stable while GDP was growing. These differences are reflected in the distributions of expectations: in Italy, the share of optimists increases from 4.11% to 11.37% between 2014 and 2016. Conversely, the share of pessimistic owners in our Italian estimation sample decreases by almost 15 percentage points over the same period (from 34.08% to 20.66%). The remarkable differences in the distributions of expectations across the US and Italy seem to reflect the heterogeneity of their respective domestic business and housing cycles. Complete descriptive statistics regarding the estimation samples can be found in in Table E.1 and Table E.2 of Appendix E.

Figure 5: Kernel Density of OEV (in log) - U.S. and Italy Samples



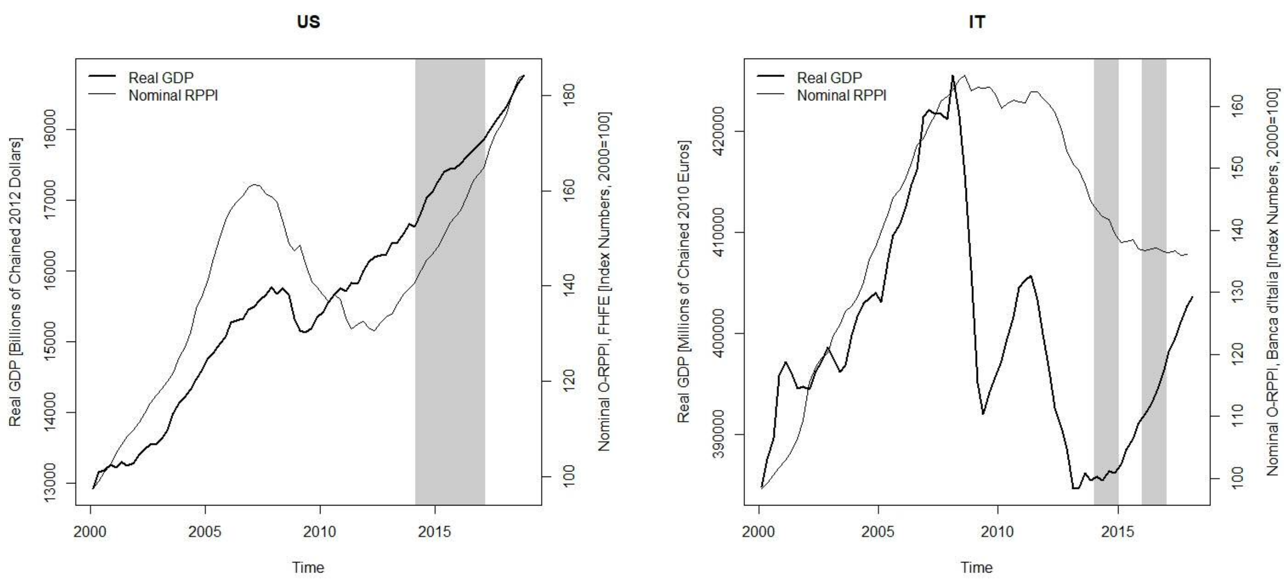
Notes: These figure refers to the estimation samples described in subsubsection 4.1.3.

Figure 6: Distribution of Expectations regarding future OEV: U.S. and Italy



Notes: These figure refers to the estimation samples described in subsubsection 4.1.3.

Figure 7: Business Cycles and Timing of the Interviews: U.S. and Italy



Notes: The grey area indicate when the interviews for the Housing module of SCE and SHIW took place. The thick line represents real GDP and the thin line a nominal O-RPPI. Sources: EUROSTAT (IT GDP), Bank for International Settlements (IT O-RPPI), U.S. Bureau of Economic Analysis (US GDP), U.S. Federal Housing Finance Agency (US O-RPPI).

4.2 Main Results

Panels A and B of Table 3 display the estimation results for Equation 5 respectively for the U.S. and Italian estimation samples. In column (1), we only control for *Optimism* and *Pessimism*. Estimates are highly significant and positive for *Optimism* and negative for *Pessimism* suggesting that optimists report significantly larger OEVs than individuals with neutral expectations while pessimists report significantly lower OEVs. These results are consistent with the predictions of the user cost formula.

However, as stated above, these estimates may confound the effects of various omitted variables. This is why we control for a vector of house characteristics *House* in columns (2) and for a set of individual characteristics *Indiv* in columns (3). As expected, *Optimism* and *Pessimism* attract lower estimates. All the estimates associated with *Optimism* remain significantly different from zero at conventional levels while the ones associated with *Pessimism* loose their significance in some specifications. A Vuong-closeness test for non-nested models (see Vuong, 1989) reveals that the quality of the fit, as measured by the adjusted R^2, is significantly better in columns (2) than in columns (3).

Consistent with our arguments in subsection 3.2, we finally control for both *House* and *Indiv* in columns (4) and the estimated parameters are qualitatively in line with the previous results. Having optimistic expectations (as compared to neutral expectations) significantly increases the OEV by 6.1 percents in the U.S. and by 6.4 percents in Italy. On the contrary, being pessimistic about the future value of the house is still associated with negative coefficients but they are not significantly different from zero anymore. This means that the effects of *Pessimism* we were capturing in columns (1) were confounding the influence of the variables included in the vectors *House* and *Indiv*. While the estimates of *Optimism* and *Pessimism* are not significantly different in absolute terms in the US, a Wald test confirms that the effect of *Optimism* is significantly higher at 5% level than the effect of *Pessimism* in Italy.

This set of results is in line with the predictions of the user cost formula. At the equilibrium, high expectations (optimistic owners here) are associated with significantly larger OEVs as compared to individuals with neutral expectations. The effects of positive expectations are remarkably similar in terms of magnitude across countries while negative expectations are associated with lower OEVs but not significantly different from the OEVs of neutral owners. Note that these conclusions hold in countries with dramatically different business cycles (as revealed in Figure 7).

Our conclusions are not driven by the way we harmonized the measure of *Optimism* and *Pessimism* across datasets. We re-estimated Equation 5 separately for each version of the questionnaire used by the interviewers and used the original measures of owners' expectations as independent variables. The results are shown in Table E.3 in Appendix E and confirm our conclusions.

The estimated parameters for physical and locational characteristics are reported in Table E.4 and Table E.5 in Appendix E. They fulfill our expectations: prices increase with additional living area, additional bathrooms and the rating of the dwelling. Prices are higher in larger municipalities. When significant, the signs and order of magnitudes of the various categories are as expected.

Dwellings' age is a delicate characteristic due to its complicated functional form. Two potentially opposing effects drive the impact of the age of the structure on prices: depreciation reduces the value as dwellings age, while structures built in certain time periods may be more valuable due to vintage effects. Hence, age is often included as a quadratic polynomial. We also re-estimate our models including age as a smooth function with an *a priori* unspecified functional form, $f(age)$. The functional form is purely data driven and estimated via penalized least squares (see Wood, 2006, for details on *Generalized Additive Models*). We stick to the parametric estimation as the signs and magnitudes of the effects most interesting here remain virtually unchanged.

The estimated parameters for respondents' characteristics are reported in Table E.4 and Table E.5 in Appendix E. In both estimation samples, we find higher self-reported house values for higher levels of household income and education. These correlations are expected in that they may also partly reflect the unobserved quality of the dwelling. While house prices are significantly higher when they are reported by respondents who are married, older and women

Table 3: Owner Estimated Value (OEV) and Price Expectations: OLS results for U.S. and Italy

Panel A: U.S.	OEV (log)			
	(1)	(2)	(3)	(4)
Optimism	0.114***	0.107***	0.061*	0.061**
	(0.027)	(0.023)	(0.024)	(0.021)
Pessimism	-0.204**	-0.094	-0.169**	-0.088
	(0.074)	(0.063)	(0.064)	(0.057)
Observations	2661	2661	2661	2661
Adjusted R^2	0.011	0.297	0.340	0.492
House characteristics	.	✓	.	✓
Individual Controls	.	.	✓	✓
Panel B: Italy	OEV (log)			
	(1)	(2)	(3)	(4)
Optimism	0.142***	0.067***	0.116***	0.064***
	(0.023)	(0.015)	(0.020)	(0.015)
Pessimism	-0.068***	-0.025**	-0.011	-0.013
	(0.014)	(0.009)	(0.012)	(0.009)
Observations	10646	10646	10646	10646
Adjusted R^2	0.013	0.599	0.286	0.617
House characteristics	.	✓	.	✓
Individual Controls	.	.	✓	✓

Notes: Standard errors are clustered at the individual level. All regressions include year dummies. In the U.S. panel, the house characteristics includes dummies for the size of the the municipality, dummies for the type of residence, dummies for the age of the residence and state*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies, an employment dummy and household income categories. In the Italian panel, the house characteristics includes the surface of the dwelling (in log), the number of bathrooms, dummies for the rating of the property, dummies for the size of the municipality, the age and the age squared of the residence and large city*region*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies and the household income (in log). Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^+$ if the p-value is lower than 0.1.

in the US, the marital status, age and gender make no difference in Italy.[8]

4.3 Reverse Causality

One may worry that we misinterpret the sense of causality. For example, an owner of a relatively expensive house may be more likely to be optimistic about the future. We make use of the panel component of the Italian dataset (2,365 individuals) to address the issue of reverse causality. For that particular part of the analysis, it is crucial that repeated OEVs provided by the same owner at different points in time reflect indeed changes in market prices – a feature confirmed by our convergent validity tests.

Table 4: Changes in Types - Italy

	$\text{Pessimistic}_{2016}$	Neutral_{2016}	Optimistic_{2016}	Total
$\text{Pessimistic}_{2014}$	176 [36.74]	263 [54.91]	40 [8.35]	479
Neutral_{2014}	568 [32.89]	972 [56.28]	187 [10.83]	1727
Optimistic_{2014}	34 [34.00]	52 [52.00]	14 [14.00]	100
Total	778	1287	241	2306

Notes: The table reports the number and mobility of types in the Italian SHIW. Figures in squared brackets are horizontal percentages.

Table 4 reports the number of transitions from one type to another between 2014 and 2016. The diagonal of the transition matrix does *not* attract the vast majority of observations: this indicates a certain degree of mobility between types which offers an opportunity to test for reverse causality straightforwardly.

We prove that *(i)* past expectations indeed predict current OEVs by substituting the current beliefs by those two years earlier:

$$\log(OEV_{it}) = \beta_0 + \beta_1 Optimism_{it-2} + \beta_2 Pessimism_{it-2} + \beta_3 House_{it} + \beta_4 Indiv_{it} + \varepsilon_{it}, \quad (6)$$

while *(ii)* past OEVs do *not* predict current expectations by estimating

$$\text{logit}[(Optimism, Pessimism, Neutral)_{it}] = \beta_5 + \beta_6 OEV_{it-2} + \beta_7 House_{it} + \beta_8 Indiv_{it} + \varepsilon_{it}. \quad (7)$$

The three-dimensional dummy-vector $(Optimism, Pessimism, Neutral)_{it}$ describes individual i's beliefs at time t, thus distinguishing between optimists $(1, 0, 0)$, pessimists $(0, 1, 0)$ and individuals with neutral expectations $(0, 0, 1)$. Model 6 and model 7 include the same sets of controls as before model 5.

We report results in Table 5. Column (1) confirms that being optimistic in 2014 predicts a higher OEV in 2016. Pessimism still has no predictive power. The last three columns in Table 5 display the marginal effects of OEVs in 2014 on expectations in 2016. None of the effects is significantly different from zero.

[8] One may also wonder whether our results differ across various sub-samples. We interacted the effect of *Optimism* and *Pessimism* with age, gender, marital status, education, income and various house characteristics and failed to find significant estimates. We also run quantile regressions but once again failed to find any statistical differences in the effects of *Optimism* and *Pessimism* across the OEVs' distribution.

Summing up, reverse causality does not seem to be a threat to our identification: past OEVs do not affect current expectations while we still identify a significant relationship between past optimism and current OEVs.

Table 5: Owner Estimated Value (OEV) and Price Expectations: Reverse Causality - Italy

	OLS	Multinomial Logit		
	OEV_{2016} (log)	$Optimism_{2016}$	$Pessimism_{2016}$	$Neutral_{2016}$
	(1)	(2)	(3)	(4)
$Optimism_{2014}$	0.130**			
	(0.043)			
$Pessimism_{2014}$	0.025			
	(0.023)			
OEV_{2014} (log)		0.010	0.005	-0.015
		(0.019)	(0.020)	(0.020)
Observations	2306	2306	2306	2306
Adjusted R^2	0.541	.	.	.
House Characteristics	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓

Notes: Standard errors are clustered at the individual level. Estimates in columns (3), (4) and (5) are marginal effects. All regressions include year dummies. The house characteristics includes the surface of the dwelling (in log), the number of bathrooms, dummies for the rating of the property, dummies for the size of the municipality, the age and the age squared of the residence and large city*region*year fixed effects. Individual controls include the age, the age squared, the gender, the marital status, education dummies and the household income (in log). Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^+$ if the p-value is lower than 0.1.

4.4 Robustness Checks

Although concerns about reverse causality have been alleviated, a causal interpretation of our estimates supposes to additionally rule out the influence of all confounding factors. While we already reduced this concern by neutralizing the influence of a large set of *objective* characteristics ($House_{it}$ and $Indiv_{it}$), one may worry that our estimates may still be biased because of *subjective* omitted variables. We address this issue below.

4.4.1 Hedonic Recall Bias

Prati (2017) shows that individuals are affected by a *hedonic recall bias*: happy individuals tend to over-estimate their actual wage. In our setting, such a hedonic recall bias may influence both, owners' stated expectations and the OEV.

We address this issue by adding a measure of happiness as an additional control while otherwise relying on the same structure as before (see Table 3, Panel B). While there is no measure of happiness in the American dataset, a measure of life satisfaction is available in the Italian one. Life satisfaction is known to be a cognitive measure of happiness that has been extensively used in the literature as a proxy for utility and welfare (see Clark, 2018). The Italian survey elicits life satisfaction as follows: *"Considering all the aspects of your life, how happy would you say you are? Please score on a scale from 1 to 10, where 1 means "extremely unhappy" and 10 "extremely happy", and the intermediate numbers serve to graduate the response."*

The results appear in column (1) of Table 6. While life satisfaction is unsurprisingly positively correlated with logged OEVs as predicted by the *hedonic recall bias* (not reported), the estimates of *Optimism* and *Pessimism* remain very similar to the baseline coefficients.

4.4.2 Dispositional Optimism and Competing Types of Expectations

One could be concerned that the way we define *Optimism* and *Pessimism* may not only reflect the effect of expectations about future prices: they may just capture expectations regarding different domains.

The literature refers to this phenomena as *dispositional optimism* which is described as *generalized positive expectations about future events.* A *setting-specific optimistic bias* (which we try to measure here) is different to dispositional optimism, in that the latter is a psychological trait *lying at the heart of an individual's outlook on life in general* (Weinstein and Klein, 1996). Optimistic people work harder, expect to retire later, are more likely to remarry, invest more in individual stocks, and save more (Puri and Robinson, 2007). So, is our measure of optimism potentially just a proxy for dispositional optimism not necessarily related to house prices?

We address this problem by re-estimating model (5) including standard proxies for dispositional optimism (*real income growth* and *overall economic conditions*).

We control for expectations regarding the change in real income growth in Italy in the column (2) of Table 6 and for expectations concerning future overall economic conditions in the U.S. sample in column (3) of the same Table. Overall, our results remain unchanged and our baseline estimates do not reflect competing types of expectations.

4.4.3 Are Expectations capturing the Effect of Unobserved Fundamental Information?

Above, we showed that our measures of expectations are neither influenced by a hedonic recall bias nor by competing types of expectations. One may still wonder whether the effect of optimism (and pessimism to a lower extent) still reflects some objective fundamental information that we do not observe.

We address this by testing whether expectations are *ex post* justified, meaning that individuals with optimistic expectations in year t indeed report significantly higher price in year $t+2$ as compared to individuals that had neutral or pessimistic expectations. We exploit again the panel component of the Italian survey taking place every two years and estimate the following regression:

$$\frac{\text{OEV}_{t+2}}{\text{OEV}_t} = \gamma_0 + \gamma_1 Optimism_{it} + \gamma_2 Pessimism_{it} + \gamma_3 House_{it} + \gamma_4 Indiv_{it} + \varepsilon_{it}. \tag{8}$$

$\text{OEV}_{t+2}/\text{OEV}_t$ is simply the ratio between the OEV in $t+2$ and t. If $Optimism_t$ carried some sort of systematic objective but unobserved information translating into a higher value of the main dwelling in the future, γ_1 should be significantly different from zero and positive. Conversely, if $Pessimism_t$ carried a systematic objective unobserved "negative" information, γ_2 should be negative.

We report estimation results for γ_1 and γ_2 in the last column of Table 6. None of the estimates is statistically significant. This absence of significance supports that our measures of *Optimism* and *Pessimism* do not capture objective but unobserved fundamental information that are systematically affecting the evolution of prices. As such, the effects of expectations that we capture in our different specifications are likely to only reflect the influence of subjective beliefs and fit our interpretation.

4.4.4 Long-Term Expectations

Both, the Italian and U.S. questionnaires aim to measure owners' expectations about the evolution of the value of their dwelling over the 12 months following the interview. One may wonder whether our conclusion are specific to this time horizon. We can address this concern using the U.S. sample since respondents were also asked to report the value of their dwelling five years ahead.

Consistently with our previous analyses, we declare an owner as optimistic (pessimistic) when she expects a price increase (decrease) by 2% or more over the following five years.

Results are reported in column (4) of Table 6. The estimates are qualitatively similar to our main coefficients: the effect of *Optimism* is positive and statistically significant while *Pessimism* tends to be negative but failed to be significantly different from zero at conventional level. Note that the magnitude of *Optimism* is higher than in the other specifications. This may well be explained by a compounding effect due to a longer time horizon.

Table 6: Robustness Checks: Results

	OEV (log)				$\frac{OEV_{t+2}}{OEV_t}$
	Italy		U.S.		Italy
	(1)	(2)	(3)	(4)	(5)
Optimism	0.064***	0.063***	0.120*		0.024
	(0.015)	(0.015)	(0.021)		(0.051)
Pessimism	-0.013	-0.013	-0.057		0.029
	(0.009)	(0.009)	(0.057)		(0.026)
Long-Term Optimism				0.050*	
				(0.047)	
Long-Term Pessimism				-0.043	
				(0.062)	
Observations	10646	10646	2661	2661	2306
Adjusted R^2	0.617	0.618	0.496	0.493	0.031
House Characteristics	✓	✓	✓	✓	✓
Individual Controls	✓	✓	✓	✓	✓
Life Satisfaction	✓	.	.	.	.
Future Real Income Growth	.	✓	.	.	.
Future Economic Conditions	.	.	✓	.	.

Notes: Standard errors are clustered at the individual level. "Long-Term Optimism" and "Long-Term Pessimism" both refer to owners' expectations about the evolution of the value of their dwelling in 5 years. All regressions include year dummies. In the U.S. panel, the house characteristics includes dummies for the size of the the municipality, dummies for the type of residence, dummies for the age of the residence and state*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies, an employment dummy and household income categories. In the Italian panel, the house characteristics includes the surface of the dwelling (in log), the number of bathrooms, dummies for the rating of the property, dummies for the size of the municipality, the age and the age squared of the residence and large city*region*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies and the household income (in log). Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^+$ if the p-value is lower than 0.1.

5 Summary and Conclusions

We argue that subjective data, specifically owners' estimated value (OEV) of the main residence, is a valuable source to study housing market dynamics. To support our arguments, we design and apply three convergent validity tests that assess the internal consistency of three dimension of OEVs. We make use of three different datasets (the American Housing Survey, the Italian SHIW and the European HFCS) to construct subjective residential property price indices.

While we confirm the previously reported tendency of owners to overestimate the value of their dwelling, we also find that OEVs combined with house characteristics as well as repeatedly provided OEVs over time by the same owner reproduce the evolution of objectively observed transaction data strikingly well.

The user cost model predicts that *ceteris paribus* optimistic expectations about future capital gains justify higher prices today, and analogically for pessimistic expectations. We use OEV and expectations regarding future house price growth reported in different waves of the American Survey of Consumer Expectations produced by the Federal Reserve Bank and the Italian SHIW to test the predictions of the user cost model. We control for fundamental differences in house prices in a hedonic fashion.

We find similar results in both datasets. Optimistic owners report home values that are on average about 6–7% higher than owners with neutral expectations regarding the future value of their dwelling. However, we do not find a significant effect of pessimistic expectations. We show that our estimates survive an ample set of robustness checks and our discussion regarding the direction of causality favours the one of expectations affecting prices.

Our results are important for several reasons. From a methodological point of view, we demonstrate that subjective residential property price indices are a credible alternative when one wants to describe the evolution of the housing market in the absence of objective information. We delineate cases where OEVs are a valid source to study housing markets from cases where they are not. The convergent validity tests are carried out for a large set of countries and hence we are confident that our results are fundamental. Second, we test the theoretical predictions of the user cost model on micro data that describes the supply side of the private secondary housing market. We document a highly persistent link: it is present in both the US and Italy as well as during booming and gloomy years. Our results also confirm that stated expectations elicited in a survey carry economically meaningful information. Last, our simple but robust empirical analysis produces evidence of an asymmetric effect of expectations that is neglected by the traditional user cost model.

We hope that our paper will motivate future research in this domain. We think that it is essential to know more about the validity of OEVs. To what extent are OEVs influenced by other types of reporting biases not assessed here? Is the endowment effect truly the prime explanation of inflated OEVs as compared to objective market prices? Do OEVs predict individual behaviour and decision-making that objective characteristics do not? According to us, these are relevant and promising research questions. More has also to be done regarding the relationship between expectations and house prices. While we believe that our results can be causally interpreted, we think that estimating the effect of future expectations on house prices in the setting of a (quasi-natural) experiment would be a valuable complementary contribution.

References

Agarwal, S. (2007). The impact of homeowners' housing wealth misestimation on consumption and saving decisions. *Real Estate Economics*, 35:135–154.

Anenberg, E. (2011). Loss aversion, equity constraints and seller behavior in the real estate market. *Regional Science and Urban Economics*, 41:67–76.

Baffigi, A., Cannari, L., and D'Alessio, G. (2016). Fifty years of household income and wealth surveys: history, methods and future prospects. Technical report, Bank of Italy, Economic Research and International Relations Area.

Bailey, M. J., Muth, R. F., and Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58:933–942.

Bao, H. X. and Gong, C. M. (2016). Endowment effect and housing decisions. *International Journal of Strategic Property Management*, 20(4):341–353.

Benítez-Silva, H., Eren, S., Heiland, F., and Jiménez-Martín, S. (2015). How well do individuals predict the selling prices of their homes? *Journal of Housing Economics*, 29:12–25.

Blanchard, O. J. and Watson, M. W. (1982). Bubbles, rational expectations and financial markets. *NBER Working Papers*, No. 945.

Cannari, L., D'Alessio, G., and Vecchi, G. (2016). House prices in Italy, 1927 – 2012. *Banca D'Italia Questioni di Economia e Finanza (Occational Papers)*, No. 333.

Case, K. E. and Shiller, R. J. (1987). Prices of single family homes since 1970: New indexes for four cities. *New England Economic Review*, Sept/Oct:45–56.

Case, K. E. and Shiller, R. J. (1988). The behavior of home buyers in boom and post-boom markets. *New England Economic Review*, November/December(80):29–46.

Case, K. E. and Shiller, R. J. (1989). The efficiency of the market for single-family homes. *American Economic Review*, 79:125–137.

Case, K. E., Shiller, R. J., and Thompson, A. K. (2012). What have they been thinking? homebuyer behavior in hot and cold markets. *Brookings Papers on Economic Activity*, (2).

Chan, S., Dastrup, S., and Ellen, I. G. (2016). Do homeowners mark to market? a comparison of self-reported and estimated market home values during the housing boom and bust. *Real Estate Economics*, 44:627–657.

Clark, A. E. (2018). Four decades of the economics of happiness: Where next? *Review of Income and Wealth*, 64:245–269.

de Haan, J. and Diewert, W. E., editors (2013). *Handbook on Residential Property Prices Indices (RPPIs).* Methodologies and Working Papers. Eurostat, Luxembourg.

Einiö, M., Kaustia, M., and Puttonen, V. (2008). Price setting and the reluctance to realize losses in apartment markets. *Journal of Economic Psychology*, 29:19–34.

Gallin, J., Molloy, R., Nielsen, E. R., Smith, P. A., and Sommer, K. (2018). Measuring aggregate housing wealth: New insights from an automated valuation model.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics*, 116:1233–1260.

Glaeser, E. L., Gyourko, J., Morales, E., and Nathanson, C. G. (2014). Housing dynamics: An urban approach. *Journal of Urban Economics*, 81:45–56.

Glaeser, E. L. and Nathanson, C. G. (2015). Housing bubbles. In *Handbook of regional and urban economics*, volume 5, pages 701–751. Elsevier.

Goodman Jr, J. L. and Ittner, J. B. (1992). The accuracy of home owners' estimates of house value. *Journal of Housing Economics*, 2:339–357.

Head, A., Lloyd-Ellis, H., and Sun, H. (2014). Search, liquidity, and the dynamics of house prices and construction. *American Economic Review*, 104:1172–1210.

Heston, A. and Nakamura, A. O. (2009). Questions about the equivalence of market rents and user costs for owner occupied housing. *Journal of Housing Economics*, 18:273–279.

Hill, R. J. (2013). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy. *Journal of Economic Surveys*, 27:879–914.

Hill, R. J. and Scholz, M. (2018). Can geospatial data improve house price indexes? a hedonic imputation approach with splines. *Review of Income and Wealth*, 64(4):737–756.

Himmelberg, C., Mayer, C., and Sinai, T. (2005). Assessing high house prices: Bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4):67–92.

Jordà, Ò., Schularick, M., and Taylor, A. M. (2017). Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual*, 31(1):213–263.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy*, 98(6):1325–1348.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *Journal of Economic perspectives*, 5(1):193–206.

Keely, R. and Lyons, R. C. (2019). Debt and taxes: The sale-rent housing price ratio in Dublin since 1945. *Trinity Economics Papers*, 0419.

Kennedy, P. E. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations. *The American Economic Review*, 71(4):801.

Kiel, K. A. and Zabel, J. E. (1999). The accuracy of owner-provided house values: The 1978–1991 American Housing Survey. *Real Estate Economics*, 27(2):263–298.

Kish, L. and Lansing, J. B. (1954). Response errors in estimating the value of homes. *Journal of the American Statistical Association*, 49(267):520–538.

Knoll, K., Schularick, M., and Steger, T. (2017). No price like home: Global house prices, 1870–2012. *American Economic Review*, 107(2):331–353.

Leamer, E. E. (2007). Housing is the business cycle. Paper presented at Housing, Housing Finance, and Monetary Policy Symposium, sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, WY, 30 August – 1 September 2007.

Leamer, E. E. (2015). Housing really is the business cycle: what survives the lessons of 2008–09? *Journal of Money, Credit and Banking*, 47(S1):43–50.

Molloy, R. and Nielsen, E. R. (2018). How can we measure the value of a home? comparing model-based estimates with owner-occupant estimates. *FED Notes*, (2018-10):11.

Pace, K. (1998). Appraisal using generalized additive models. *Journal of Real Estate Research*, 15(1):77–99.

Piazzesi, M. and Schneider, M. (2009). Momentum traders in the housing market: survey evidence and a search model. *American Economic Review*, 99(2):406–11.

Poterba, J. M. (1984). Tax subsidies to owner-occupied housing: An asset-market approach. *The Quarterly Journal of Economics*, 99(4):729–752.

Prati, A. (2017). Hedonic recall bias. Why you should not ask people how much they earn. *Journal of Economic Behavior & Organization*, 143:78–97.

Puri, M. and Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, 86(1):71–99.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1):34.

Scott, P. J. and Lizieri, C. (2012). Consumer house price judgements: new evidence of anchoring and arbitrary coherence. *Journal of Property Research*, 29(1):49–68.

Senik, C. (2014). Wealth and happiness. *Oxford review of economic policy*, 30(1):92–108.

Shiller, R. J. (2015). *Irrational Exuberance*. Princeton University Press, 3rd edition.

Stephens, T. A. and Tyran, J.-R. (2012). 'At least i didn't lose money'-nominal loss aversion shapes evaluations of housing transactions. *University of Copenhagen Dept. of Economics Discussion Paper*, (12-14).

Stiglitz, J. E. (1990). Symposium on bubbles. *Journal of Economic Perspectives*, 4(2):13–18.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, 1(1):39–60.

Van der Cruijsen, C., Jansen, D.-J., and van Rooij, M. (2014). The rose-colored glasses of homeowners. *De Nederlandsche Bank Working Paper*, 421.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333.

Wallace, N. E. and Meese, R. A. (1997). The construction of residential housing price indices: A comparison of repeat-sales, hedonic-regression, and hybrid approaches. *The Journal of Real Estate Finance and Economics*, 14(1/2):51–73.

Waltl, S. R. (2016). A hedonic house price index in continuous time. *International Journal of Housing Markets and Analysis*, 9(4):648–670.

Waltl, S. R. (2019). Variation across price segments and locations: A comprehensive quantile regression analysis of the Sydney housing market. *Real Estate Economics*, 47(33):723–756.

Waltl, S. R. (2020). Multidimensional wealth inequality: A hybrid approach toward distributional national accounts in europe. *INEQ Working Paper Series*, 16.

Weinstein, N. D. and Klein, W. M. (1996). Unrealistic optimism: Present and future. *Journal of Social and Clinical Psychology*, 15(1):1–8.

Willig, R. D. (1976). Consumer's surplus without apology. *The American Economic Review*, 66(4):589–597.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R.* Chapman and Hall / CRC, Boca Raton, Florida.

Appendix

A Subjective House Price Indices: a Tool to Test for Convergent Validity

Wealth and housing surveys usually ask for an OEV. Additionally, surveys sometimes collect the original acquisition price and the year of purchase as well as some dwelling characteristics. Regarding the latter, surveys often provide substantial information on physical characteristics but, unfortunately, regularly lack information on the location of the dwelling.

We use these frequently available pieces of information to construct residential property price indices (RPPI). We provide a methodology for each common set of information available in surveys.

Just like in the case of an objective residential property price index (O-RPPI), also for subjective indices (S-RPPIs) some kind of quality-adjustment needs to be performed to guarantee that like is compared with like. This is important as a large fraction of variation of house prices cross-sectionally and over time is due to differences in locational and physical characteristics of the properties.

We adjust objective techniques in a way suitable for subjective survey data. We propose three main and a hybrid technique, which are inspired by the two major classes used for objective data: repeat-sales and hedonics.

Hence, the S-RPPIs filter out price dynamics in a comparable way as done when compiling O-RPPIs. As the quality-adjustment procedures are comparable, we can conclude that whenever a S-RPPI matches an O-RPPI the data used to construct the S-RPPI contains internally coherent information on housing market dynamics. Our convergent validity tests built upon this conclusion.

Table A.1 and Table A.2 report all S-RPPIs constructed in this article. Furthermore, Figure A.1 compares S-RPPIs for several European countries to objective counterparts.

A.1 A Hedonic Approach

Hedonic valuation assumes that differentiated products are completely described by a vector of objectively measured characteristics and that observed prices of these goods are composed of implicit or "hedonic" prices associated with these characteristics (Rosen, 1974). Due to the large degree of heterogeneity of house, hedonic valuation approaches are widely used in this context. Thereby, the relevant characteristics describe both, the major features of the structure and the attractiveness of the dwelling's location.

Hence, a sufficiently comprehensive list of physical and locational characteristics needs to be available to filter out price changes over time net of differences in characteristics among dwellings. While surveys usually collect the most important physical characteristic (such as the age of the structure and the living surface), they often lack information on location. The Italian SHIW and the US AHS do contain geographical information.

Hedonic indices are criticized for the high data demand in terms of dwelling characteristics, and a potential subjective specification of the hedonic equation. Neither is there consensus on the functional form in which characteristics should enter the equation nor on the ideal selection of the same.[9] While the list of potentially important characteristics is endless, many important

[9] A growing body of literature thus suggests modelling the functional form of characteristics non-parametrically. See Hill and Scholz (2018) and Waltl (2019) for such approaches for location and Pace (1998)

features are rarely available for transaction data and even more are simply not measurable in an objective way. For instance, the exact layout of the rooms is an important price-determining factor, but it is unclear how such information could be classified. Usually, the selection of characteristics is predominately driven by data availability.

In general, all types of hedonic index construction techniques can be applied to subjective data just as to transaction data (see Hill, 2013; de Haan and Diewert, 2013, for a survey on techniques). We suggest a chained hedonic adjacent-period time-dummy index as it is ideally suited for survey data as described after presenting the methodology.

We estimate hedonic models that regress the logged price of the dwelling on a matrix of physical X and locational characteristics L with associated shadow prices β and λ, respectively. We pool data from two adjacent survey waves, and hence time dummies D_t and D_{t-1} enter the model. In the classical time dummy approach, time dummies refer to the period of sale. In the subjective version, the time dummies indicate the period of valuation – the survey year.

The associated parameters δ_s and δ_{s-1} filter out pure price changes over time net of differences in dwelling characteristics:

$$\log(P) = \delta_s \cdot D_s + \delta_{s-1} \cdot D_{s-1} + X\beta + L\lambda + \varepsilon, \tag{9}$$

thereby, ε denotes a vector of independent and identically distributed error terms.

From each model, we extract the price change from period $s-1$ to period s[10] via

$$\exp(\hat{\delta}_s - \hat{\delta}_{s-1}).$$

An index number stating the price change between the base period $t^* = 1$ and period t is obtained by chaining all period-to-period changes up to period t

$$I_{hedonic}(t) = \prod_{s=1}^{t} \exp(\hat{\delta}_s - \hat{\delta}_{s-1}).$$

In the case of bi- or triennial surveys, index numbers for years between survey waves need to be constructed by interpolation.

This chained time-dummy approach has two distinct advantages when dealing with survey data: first, the variables collected in surveys change over time. Often it is the case that a variable is collected for some years, and at some point dropped or replaced by a similar but not identical question. For instance, in the SHIW the categories describing the location of a dwelling changed twice. The chained approach requires only that the same characteristics are available in two adjacent periods but not throughout the entire time span. The flexibility of the chained approach allows one to exploit (almost) all available characteristics.

Second, shadow prices associated with structural and, even more importantly locational, characteristics may too change over time. This is particularly relevant when dealing with long time periods spanning several decades. The chained approach automatically allows shadow prices to change over time.

for physical characteristics.

[10]This back-transformation taking the exponential function yields a biased estimate for the mean (see also Kennedy, 1981). The house price index literature usually does not perform a bias correction since the differences in magnitude are small (see also Hill, 2013). Additionally, Waltl (2016) shows that the resulting estimator is an unbiased for the median and hence the resulting index is conceptually comparable to standard median indices.

A.2 Repeat-observations approaches

A repeat-sales index (see Bailey et al., 1963; Case and Shiller, 1987, 1989) exploits price information from repeatedly sold dwellings. Only price pairs of the same dwelling are considered and aggregated to an overall index. As the price of the same dwellings is compared over time, an almost perfect quality-adjustment is obtained. The standard repeat-sales index is criticized for ignoring changes in the age (and hence often the quality) of the structure, measurement errors from unobserved renovations, the exclusion of first sales, and a potential lemons' bias, as low-quality homes tend to sell more frequently (see Wallace and Meese, 1997).

We propose two subjective indices inspired by the repeat-sales methodology: the *repeat-observations-panel* (ROP) and the *repeat-observations-acquisitions* (ROA). We also compile a hybrid version of the two labelled the *hybrid RO* approach.

A.2.1 The ROP approach

For surveys including a panel component, repeated OEVs referring to the same dwelling can be linked and used to compile a repeat-sales-type index. In fact, the major points of critique of the repeat-sales technique do not apply to survey panel data: there is no reason to believe that surveys over-represent owners of low-quality homes and hence a lemons' bias is ruled out.[11]

Just like for objective repeat-sales indices, homes that were just bought are excluded. The prices of these dwellings will only be included into the index as soon as the household was contacted again in the subsequent survey wave. This is, however, much less of a concern as for transaction data, since surveys take place much more frequently than re-sells.

Due to short periods of time between survey waves (usually wealth surveys take place every two to three years), the aging of the dwelling is less important. Additionally, some surveys contain information on major renovations, which can be made use of.

To construct a standard repeat-sales index, the logged ratio of prices related to subsequent sales is regressed on dummy variables indicating the period of first and repeated sale. When using survey data, we take the ratio of OEVs obtained from subsequent survey waves:

$$\log\left(\frac{P_{i,t}}{P_{i,s}}\right) = \sum_{m=1}^{T_S} \tau_{ROP}^m D_i^m + \varepsilon_i, \tag{10}$$

where $P_{i,t}$ and $P_{i,s}$ denote dwelling i's OEV at times t and s respectively with $1 \leq s < t \leq T_S$ and T_S the number of survey waves. In a standard repeat-sales framework, the dummy variables D_i^m are $D_i^{m=t} = 1$ for the period of re-sale, $D_i^{m=s} = -1$ for the period of the first sale and 0 otherwise. For survey data, the periods of sale and re-sale are identified with subsequent survey years. Finally, ε_i denotes an independent and identically distributed error term.

The estimated coefficients $\hat{\tau}^t$ are used to construct the price index: the change in prices between the base period $t^* = 1$ and period t is given by

$$I_{ROP}(t) = \exp\left(\hat{\tau}_{ROP}^t\right).$$

Note that in the common case of bi- or triennial surveys, index numbers for years between survey waves – just as in the case of hedonic indices – are obtained by interpolation.

[11] However, wealth surveys are known to under-represent the wealthiest of the wealthy households, which also tend to be the owners of the most valuable properties in a country (for differences in average housing wealth across the distribution, see Waltl, 2020).

A.2.2 The ROA approach

The ROA approach differs from the ROP approach in the way price pairs are constructed. In the absence of a panel structure, the original acquisition price can be matched to an OEV to identify price pairs entering a repeat-sales model.

Let $t_{S(i)}$ for $i \in \{1, \ldots, T_S\}$ denote survey years. The earliest acquisition year is set as base period 1, and the number of years elapsed between the earliest acquisition year and the latest survey wave is T. The function $S : \{1, \ldots, T_S\} \longrightarrow \{1, \ldots, T\}$ maps survey years into the period of observation.

We denote the year of acquisition by s, the acquisition price by $P_{i,s}$ and the OEV provided in a survey year $t \in \{t_{S(i)}\}_{i=1}^{T_S}$ by $P_{i,t}$. Consequently, $D_i^{m=t} = 1$, $D_i^{m=s} = -1$ and $D_i^m = 0$ for $m \notin \{s, t\}$. *Per construction,* $1 \leq s < t \leq T$ and $T_S \leq T$. Then, model (10) turns into

$$\log\left(\frac{P_{i,t}}{P_{i,s}}\right) = \sum_{m=1}^{T} \tau_{ROA}^m D_i^m + \varepsilon_i,$$

and the estimated parameters $\hat{\tau}_{ROA}$ are used to construct the index

$$I_{ROA}(t) = \exp\left(\hat{\tau}_{ROA}^t\right).$$

A major advantage is that the approach can be applied to cross-sectional surveys. This largely increases the number of countries for which such an index can be estimated. As long as some homes were acquired in years between survey waves, the ROA approach yields index numbers also for these intermediate years. Additionally, the start of the index is not the first survey year but, at least in theory, the year of the earliest acquisition. I.e., the index spans the period between the earliest acquisition year and the last survey year, and provides index numbers not just for the subset of survey years $\{S(1), \ldots, S(T_S)\} \subset \{1, \ldots, T\}$.

In practise, one would disregard years with only very few acquisitions and let the index start once there are sufficient observations to guarantee a stable index. Still, the ROA approach may produce index numbers for several decades and may even lead to substantially longer time series than currently available objective indices.

The length of the index comes at the cost of potentially long time spans between the two observed prices for the same dwelling. In the meantime, the dwelling may have been refurbished, renovated and/or depreciated, which makes the dwelling at time of acquisition less comparable to the dwelling at the time of the survey interview. In particular, the index is not expected to be very precise for years very distant in time from the survey, but precision increases when approaching the first survey year. Additionally, the precision of the index will increase with every additional survey wave as new price pairs become available.

The problem arising when mixing objective acquisition prices with subjective OEVs is discussed in section 2. We there also show that for the time period overlapping the survey years, the index is affected by unrealistic spikes stemming from systematically higher OEVs than objective prices. However, index number preceding the first survey year are reliable.

As shown in Appendix D, in the case of bi- or triennial surveys, also the ROA index numbers between survey years tend to be reliable. Hence, the *adjusted ROA index* leaves out index numbers referring to survey years and reconstructs the missing index numbers via linear interpolation, i.e., for a biennial survey, the adjusted ROA index is given by

$$I_{ROA_{adj}}(t) = \begin{cases} I_{ROA}(t) & \text{if } t \notin \{t_{S(i)}\}_{i=1}^{T_S}, \\ \frac{I_{ROA}(t+1) - I_{ROA}(t-1)}{2} & \text{if } t \in \{t_{S(i)}\}_{i=1}^{T_S}. \end{cases}$$

A.2.3 The RO hybrid approach

The ROP approach links repeatedly reported prices for the same dwelling, while the ROA approach links acquisition and current prices.

While the ROP is expected to precisely measure price dynamics during survey years, i.e., for $\{S(1), \ldots, S(T_S)\}$, the ROA approach is very useful due to the length of the resulting series. However, the ROA index is expected to be affected by systematic over-reporting among home-owners in the period overlapping survey years.

To benefit from the advantages of either approach, a hybrid approach splices the ROP and ROA indices together: prior to the year of the first survey wave, an ROA index yields reliable results, whereas thereafter a ROP index should be preferred.

To compute the hybrid index, the ROA index is re-estimated leaving out data from the first survey year $t_{S(1)}$. The resulting index I_{ROA} as well as the ROP index I_{ROP} are normalized to $t_{S(1)}$, i.e.,

$$\tilde{I}_{ROA}(t) = \frac{I_{ROA}(t)}{I_{ROA}(t_{S(1)})} \quad \text{and} \quad \tilde{I}_{ROP}(t) = \frac{I_{ROP}(t)}{I_{ROP}(t_{S(1)})}.$$

Then, the hybrid index is obtained via

$$I_{RO_{hybrid}}(t) = \begin{cases} \tilde{I}_{ROA}(t) & \text{if } t \leq t_{S(1)}, \\ \tilde{I}_{ROP}(t) & \text{if } t > t_{S(1)}. \end{cases}$$

Per construction, $I_{RO_{hybrid}}(t_{S(1)}) = 1$, but it can be normalized to any preferred base period.

In the absence of a panel structure, the ROP index can be substituted by a hedonic index.

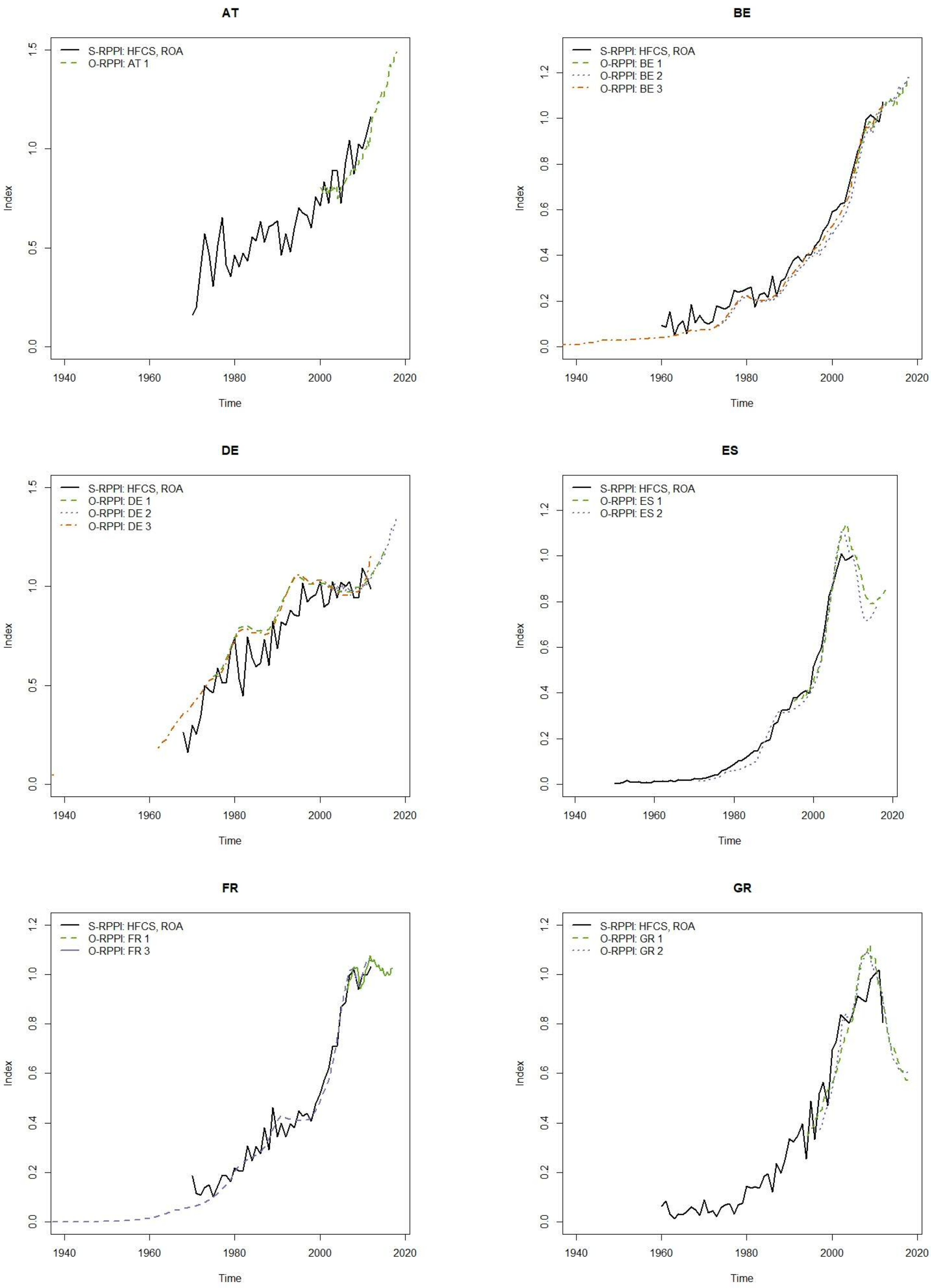
Table A.1: S-RPPIs for the US

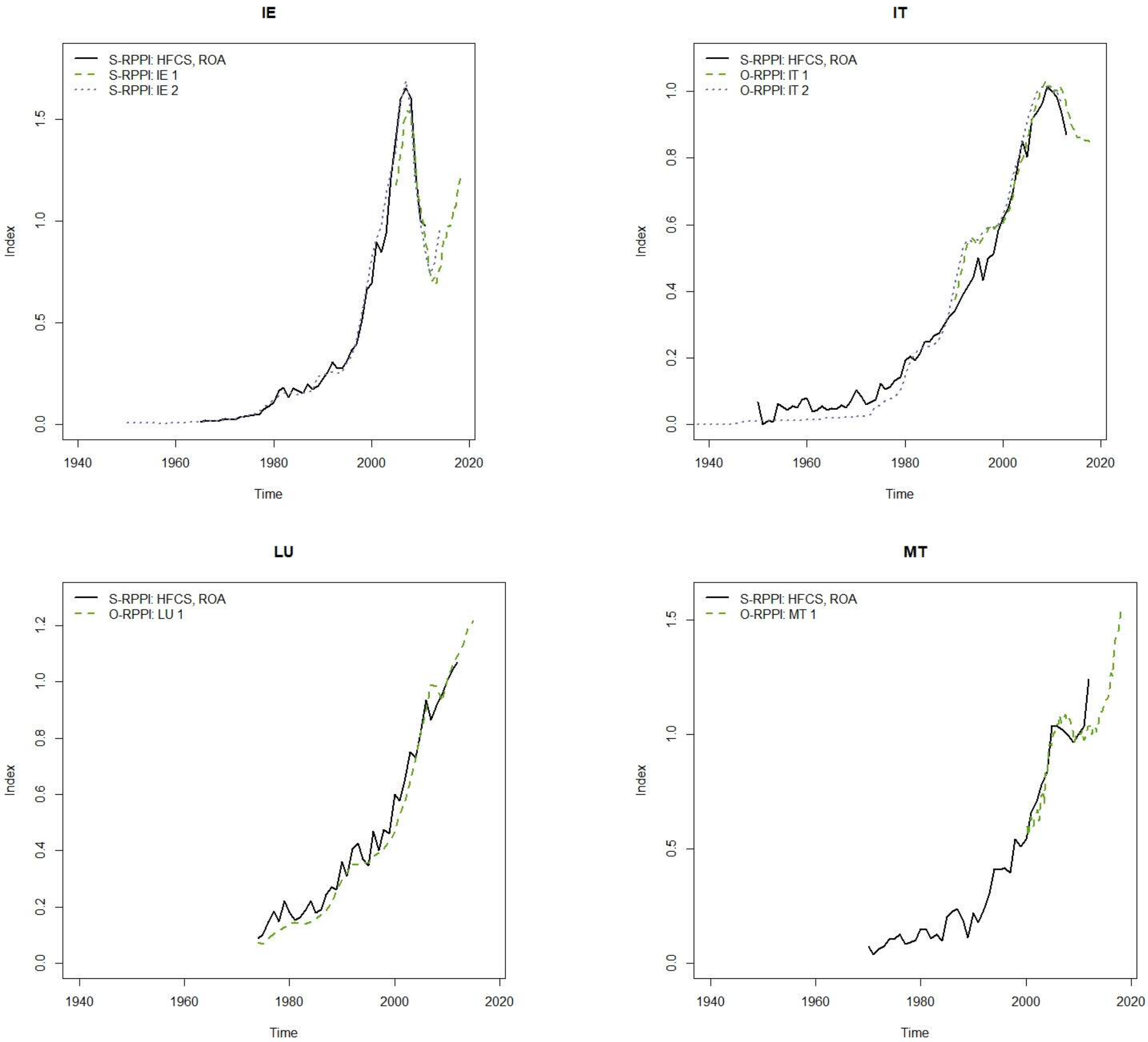
Year	ROP	ROA (national)	ROA adj	ROA (metropolitan)	ROA adj. (metropolitan)	RO hybrid	Hedonic
1932		0.144	0.144	0.027	0.027	0.144	
1933		0.019	0.019	0.153	0.153	0.019	
1934		0.022	0.022	0.130	0.130	0.022	
1935		0.124	0.124	0.124	0.124	0.124	
1936		0.035	0.035	0.091	0.091	0.035	
1937		0.026	0.026	0.084	0.084	0.026	
1938		0.062	0.062	0.040	0.040	0.062	
1939		0.061	0.061	0.051	0.051	0.061	
1940		0.146	0.146	0.097	0.097	0.146	
1941		0.053	0.053	0.040	0.040	0.053	
1942		0.103	0.103	0.055	0.055	0.103	
1943		0.041	0.041	0.086	0.086	0.041	
1944		0.083	0.083	0.043	0.043	0.083	
1945		0.078	0.078	0.098	0.098	0.078	
1946		0.096	0.096	0.081	0.081	0.096	
1947		0.065	0.065	0.083	0.083	0.065	
1948		0.080	0.080	0.093	0.093	0.080	
1949		0.083	0.083	0.112	0.112	0.083	
1950		0.123	0.123	0.118	0.118	0.123	
1951		0.084	0.084	0.101	0.101	0.084	
1952		0.134	0.134	0.107	0.107	0.134	
1953		0.118	0.118	0.115	0.115	0.118	
1954		0.115	0.115	0.120	0.120	0.115	
1955		0.123	0.123	0.118	0.118	0.123	
1956		0.132	0.132	0.134	0.134	0.132	
1957		0.116	0.116	0.132	0.132	0.116	
1958		0.131	0.131	0.132	0.132	0.131	
1959		0.121	0.121	0.133	0.133	0.121	
1960		0.153	0.153	0.147	0.147	0.153	
1961		0.149	0.149	0.140	0.140	0.149	
1962		0.138	0.138	0.145	0.145	0.138	
1963		0.154	0.154	0.139	0.139	0.154	
1964		0.144	0.144	0.146	0.146	0.144	
1965		0.162	0.162	0.147	0.147	0.162	
1966		0.167	0.167	0.160	0.160	0.167	
1967		0.179	0.179	0.165	0.165	0.179	
1968		0.170	0.170	0.172	0.172	0.170	
1969		0.187	0.187	0.171	0.171	0.187	
1970		0.220	0.220	0.200	0.200	0.220	
1971		0.220	0.220	0.211	0.211	0.220	
1972		0.227	0.227	0.224	0.224	0.227	
1973		0.242	0.242	0.238	0.238	0.242	
1974		0.258	0.258	0.256	0.256	0.258	
1975		0.289	0.289	0.273	0.273	0.289	
1976		0.295	0.295	0.300	0.300	0.295	
1977		0.360	0.360	0.339	0.339	0.360	
1978		0.385	0.385	0.383	0.383	0.385	
1979		0.446	0.446	0.430	0.430	0.446	
1980		0.459	0.459	0.462	0.462	0.459	
1981		0.493	0.493	0.509	0.509	0.493	
1982		0.458	0.458	0.543	0.543	0.458	
1983		0.482	0.482	0.546	0.546	0.482	
1984		0.498	0.498	0.557	0.557	0.498	
1985		0.534	0.534	0.566	0.566	0.534	
1986		0.588	0.588	0.620	0.620	0.588	
1987		0.618	0.618	0.648	0.648	0.618	
1988		0.642	0.642	0.684	0.684	0.642	
1989		0.645	0.645	0.734	0.734	0.645	
1990		0.686	0.686	0.720	0.720	0.686	
1991		0.690	0.690	0.746	0.746	0.690	
1992		0.707	0.707	0.741	0.741	0.707	
1993		0.726	0.726	0.762	0.762	0.726	
1994		0.763	0.763	0.789	0.789	0.763	
1995		0.786	0.786	0.803	0.803	0.786	
1996		0.795	0.795	0.829	0.829	0.795	
1997		0.822	0.822	0.872	0.872	0.822	
1998		0.878	0.878	0.934	0.909	0.878	
1999	0.931	0.977	0.939	0.945	0.945	0.939	0.936
2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2001	1.069	1.120	1.062	1.097	1.097	1.069	1.064
2002	1.132	1.124	1.124	1.110	1.154	1.132	1.128
2003	1.194	1.287	1.222	1.211	1.211	1.194	1.193
2004	1.288	1.320	1.320	1.238	1.337	1.288	1.276
2005	1.383	1.502	1.405	1.463	1.463	1.383	1.360
2006	1.465	1.489	1.489	1.503	1.503	1.465	1.437
2007	1.547	1.636	1.444	1.698	1.447	1.547	1.515
2008	1.468	1.400	1.400	1.390	1.390	1.468	1.446
2009	1.389	1.461	1.338	1.308	1.308	1.389	1.377
2010	1.356	1.277	1.277	1.243	1.243	1.356	1.348
2011	1.323	1.418	1.241	1.322	1.228	1.323	1.320
2012		1.206	1.206	1.213	1.213		1.323
2013							1.327

Table A.2: S-RPPIs for Italy

Year	ROP	ROA	ROA adj	RO	Hedonic
1960		0.112	0.112	0.065	
1961		0.078	0.078	0.045	
1962		0.075	0.075	0.043	
1963		0.080	0.080	0.046	
1964		0.070	0.070	0.040	
1965		0.105	0.105	0.061	
1966		0.079	0.079	0.046	
1967		0.099	0.099	0.057	
1968		0.098	0.098	0.057	
1969		0.098	0.098	0.057	
1970		0.170	0.170	0.098	
1971		0.107	0.107	0.062	
1972		0.115	0.115	0.066	
1973		0.129	0.129	0.074	
1974		0.159	0.159	0.092	
1975		0.193	0.193	0.112	
1976		0.167	0.167	0.097	
1977	0.106	0.183	0.183	0.106	
1978	0.126	0.240	0.240	0.126	
1979	0.157	0.242	0.242	0.157	
1980	0.191	0.343	0.343	0.191	
1981	0.222	0.317	0.317	0.222	
1982	0.270	0.343	0.343	0.270	
1983	0.314	0.369	0.369	0.314	
1984	0.331	0.400	0.400	0.331	
1985	0.346	0.441	0.441	0.346	
1986	0.362	0.456	0.456	0.362	0.359
1987	0.375	0.432	0.432	0.375	0.401
1988	0.432	0.481	0.481	0.432	0.448
1989	0.489	0.519	0.519	0.489	0.495
1990	0.555	0.580	0.580	0.555	0.579
1991	0.622	0.591	0.591	0.622	0.663
1992	0.690	0.704	0.704	0.690	0.727
1993	0.757	0.702	0.702	0.757	0.791
1994	0.803	0.752	0.752	0.803	0.829
1995	0.848	0.774	0.774	0.848	0.868
1996	0.865	0.815	0.815	0.865	0.880
1997	0.881	0.834	0.834	0.881	0.892
1998	0.898	0.863	0.863	0.898	0.904
1999	0.949	0.917	0.917	0.949	0.952
2000	1.000	1.000	1.000	1.000	1.000
2001	1.078	1.060	1.060	1.078	1.063
2002	1.157	1.407	1.124	1.157	1.126
2003	1.282	1.187	1.187	1.282	1.263
2004	1.406	1.730	1.271	1.406	1.400
2005	1.510	1.355	1.355	1.510	1.496
2006	1.614	1.954	1.453	1.614	1.593
2007	1.657	1.552	1.552	1.657	1.642
2008	1.700	2.216	1.622	1.700	1.692
2009	1.710	1.691	1.691	1.710	1.696
2010	1.721	2.181	1.633	1.721	1.700
2011	1.695	1.575	1.575	1.695	1.675
2012	1.670	2.141	1.487	1.670	1.651
2013	1.617	1.399	1.399	1.617	1.593
2014	1.564	1.893	1.321	1.564	1.536
2015		1.243	1.243		1.476
2016					1.415

Figure A.1: ROA Indices for European Countries – HFCS data





B Data

B.1 American Housing Survey

B.1.1 Description of the data

The American Housing Survey (AHS), sponsored by the US Department of Housing and Urban Development and carried out by the US Census Bureau, is a national housing survey collecting various housing and demographic characteristics.

The sample unit is the dwelling. There are two samples: a national and a metropolitan sample. The national sample is longitudinal and conducted in odd-numbered years. The metropolitan sample is cross-sectional and usually conducted in even-numbered years. The survey documentation can be found online.[12]

We only use data for owner-occupied dwellings excluding mobile homes. The necessary identifiers are available from 1999 onward. Hence, we use survey data from 1999 onward. The latest

[12] `https://www.census.gov/programs-surveys/ahs/about.html`, last accessed on May 28, 2019.

available wave is from 2013. Table B.1 reports summary statistics for the national sample.

Table B.1: Descriptive Statistics - AHS national sample

	Mean	S.D.	Min	Max
OEV (log)	12.02	0.82	7.00	15.30
Acquisition price (log)	11.33	1.36	0	14.18
House Characteristics:				
Surface (square feet, log)	7.53	0.55	4.60	10.12
Age of the Dwelling (years)	40.82	25.85	0.00	94.00
Number of bedrooms	3.15	0.88	0.00	10.00
Number of full bathrooms	1.77	0.75	0.00	10.00
Number of half bathrooms	0.44	0.57	0.00	10.00
Central air conditioning	0.70		0	1
Complete kitchen	0.99		0	1
Garage	0.82		0	1
Type of dwelling:				
Detached House	0.90		0	1
Attached House	0.06		0	1
Multi-units structure	0.05		0	1
Geographical information:				
Region:				
Northeast	0.19		0	1
Midwest	0.27		0	1
South	0.34		0	1
West	0.20		0	1
Location of the dwelling:				
Central cities of metropolitan areas	0.23		0	1
Metropolitan area, not in central city, urban	0.39		0	1
Metropolitan area, not in central city, rural	0.17		0	1
Outside metropolitan areas, urban	0.07		0	1
Outside metropolitan areas, rural	0.15		0	1
Number of observations:				212,932
Survey waves:	biannual between 1999 and 2013			

Notes: The table reports summary statistics for the AHS national sample. We consider only owner-occupied dwellings for which the characteristics reported in the table are available. Mobile homes and dwellings with an OEV of less than USD 1,000 are discarded. The OEV, the acquisition price and several characteristics are top-coded.

B.1.2 Elicitation of the OEV

The survey participant is asked the following question: *How much do you think house/ lot/ apartment/ mobile home/ property would sell for on today's market?*

B.1.3 Methodological Notes on the S-RPPI

The hedonic index. Next to year dummies, we include the number of bedrooms (five categories: 1, 2, 3, 4, more than 4), the number of bathrooms including half-bathrooms (ten categories: 0, 1, 1.5, ...,4, 4.5, more than 4.5), the age of the dwelling (with unspecified functional form $f(age)$), the living surface (in log), as well as dummies indicating the existence of a complete kitchen, central air conditioning and the type of dwelling (detached house, attached house or multi-units structure). Additionally, we include the location of the dwelling via region dummies and

As described in Appendix A, hedonic models are estimated by pooling two consecutive survey waves. Table B.2 reports regression results when pooling all waves and including time dummies.

Table B.2: Hedonic regression: US

	OEV (log)	
Surface (in log)	0.193***	(0.000)
Age (in years)	-0.001***	(0.000)
Age squared (in years)	0.000***	(0.000)
1 bedroom (ref. category)	.	(.)
2 bedrooms	-0.023*	(0.021)
3 bedrooms	0.056***	(0.000)
4 bedrooms	0.156***	(0.000)
5+ bedrooms	0.138***	(0.000)
0 bathrooms (ref. category)	.	(.)
1 bathroom	0.223***	(0.000)
1.5 bathrooms	0.407***	(0.000)
2 bathrooms	0.543***	(0.000)
2.5 bathrooms	0.783***	(0.000)
3 bathrooms	0.874***	(0.000)
3.5 bathrooms	1.100***	(0.000)
4 bathrooms	1.204***	(0.000)
4.5 bathrooms	1.387***	(0.000)
5+ bathrooms	1.513***	(0.000)
Air conditioning	0.042***	(0.000)
Complete kitchen	0.082***	(0.000)
Garage	0.097***	(0.000)
Detached house (ref. category)	.	(.)
Attached house	-0.030***	(0.000)
Multi-units structure	0.174***	(0.000)
Constant	10.04***	(0.000)
Observations		212,932
Adjusted R^2		0.4303

Notes: The regression also contains year, region and location dummies. The region and location dummies are interacted.

The ROP index. The index links OEVs reported for the same dwelling over time. There are 202,136 observations that appear at least twice in the sample. These observations refer to 39,794 unique dwellings. We construct 162,342 price pairs, which enter the repeat-sales model.

The ROA index. The index links a dwellings acquisition price to the OEV. The earliest acquisition price is reported for 1991, however, due to low numbers of observations at the beginning of the century, we only consider acquisition prices from 1932 onward. We disregard observations with an acquisition price of less than USD 100 or where the acquisition price is missing.

We use 174,641 price pairs to estimate the ROA index. We also estimate a ROA index from the metropolitan sample, which yields 157,381 price pairs.

B.2 Italian Survey on Household Income and Wealth

B.2.1 Description of the data

The Survey on Household Income and Wealth (SHIW), sponsored and carried out by the Bank of Italy, is a national survey collecting data on the incomes, wealth and savings of Italian households as well as demographic characteristics (see Baffigi et al., 2016, as well as the online documentation (Table 13) for more details on the evolution of the dataset).

The sample unit is the household. Part of the survey is longitudinal.

B.2.2 Elicitation of the OEV

In 2014 and 2016, the current value of the house is measured using the same following question: *In your opinion, how much is your house/flat worth (unoccupied)? In other words, what price could you ask for it today (including any cellar, garage or attic)?*.

B.2.3 Measure of Expectations

The 2016 wave. In 2016, the question referred to expected price changes of the household's main residence. Respondents are asked to specify their full distribution of beliefs. A survey experiment was run to test for framing effects.

Half of the respondents were asked the following question:

> In your opinion, how will the value of your principal residence change? Assign a total of 100 points among the options below according to which you think is more likely: give high points to the most likely and low points to the least likely. Over the next 12 months the value of your residence will:
>
> - Decrease by over 5%
> - Decrease between 5 and 2%
> - Fluctuate between -2% and +2%
> - Increase between 2 and 5%
> - Increase by over 5%

For the other half, the exact amounts of a, say, 5% increase were computed and presented to the respondent. The amounts use the self-reported value of the property as basis.

The 2014 wave. In 2014, the question also refers to price changes of the household's main residence. Respondents have three options to choose from:

> In your opinion, at the end of 2015, the value of your main dwelling will be
>
> - higher,
> - the same,
> - lower.

Table B.3: Descriptive Statistics - SHIW

	Mean	S.D.	Min	Max
OEV (log)	11.40	1.04	6.94	15.42
Acquisition price (log)	10.706	1.21	6.932	14.691
House Characteristics:				
Surface (square meter, log)	4.63	0.39	1.95	6.96
Age of the Dwelling (years)	49.79	65.57	0.00	1014.00
Number of bathrooms > 2	0.46		0	1
Rating of the Property 1 (1986 – 2006):				
Luxury	0.02		0	1
Highly desirable	0.12		0	1
Mid-Range	0.63		0	1
Modest	0.17		0	1
Low-Income	0.05		0	1
Very Low-Income	0.01		0	1
Rural	0.01		0	1
Other	0.00		0	1
Rating of the Property 2 (2012 – 2016):				
Luxury	0.01		0	1
Highly desirable	0.16		0	1
Mid-Range	0.67		0	1
Modest	0.11		0	1
Low-Income	0.04		0	1
Very Low-Income	0.01		0	1
Geographical information:				
Region (20 regions): *Not reported*				
Location of the dwelling 1 (1986 – 1993):				
Countryside	0.04		0	1
Town outskirts	0.41		0	1
Between outskirts and town centre	0.29		0	1
Town centre	0.22		0	1
Highly desirable	0.03		0	1
Run-down	0.01		0	1
Location of the dwelling 2 (1993 – 2016, except 2008):				
Isolated area, countryside	0.06		0	1
Town outskirts	0.30		0	1
Between outskirts and town centre	0.32		0	1
Town centre	0.23		0	1
Other	0.00		0	1
Hamlet	0.08		0	1
Number of observations:				76795
Survey waves:				annual between 1977 and 1984; 1986; biannual between 1987 and 1995; biannual between 1998 and 2016

Notes: The table reports summary statistics for the SHIW. We consider only owner-occupied dwellings. Dwellings with a missing OEV or an OEV of less than EUR 1,000 are discarded. Also, acquisition prices lower than EUR 1,000, and where the ratio between OEV and acquisition price was less than 0.5 were deleted. Several variables were not collected from 1977 on. The summary statistics hence refer to those years, where the information was collected. The full documentation of the survey is found online.[13]

Table B.4: Hedonic specification – IT

	Specification						Model details	
	I	II	III	IV	V	VI	n	Adj. R^2
Structural characteristics								
Age (in years, non-parametric)	✓	✓	✓	✓	✓	✓		
Surface (in square meter, log)	✓	✓	✓	✓	✓	✓		
Number of bathrooms ($>$ or ≤ 2)	✓	✓			✓	✓		
Rating of the Property 1			✓	✓	✓			
Rating of the Property 2						✓		
Geographical characteristics								
Region	✓	✓	✓	✓	✓	✓		
Size of the municipality	✓	✓	✓	✓	✓	✓		
Interaction: Region and Size	✓	✓	✓	✓	✓	✓		
Location of the dwelling 1				✓				
Location of the dwelling 2		✓			✓	✓		
Two-period hedonic models								
1986 – 1987				✓			4420	0.633
1987 – 1989				✓			4255	0.669
1989 – 1991				✓			6458	0.633
1991 – 1993			✓				7491	0.617
1993 – 1995					✓		7772	0.657
1995 – 1998					✓		7948	0.665
1998 – 2000					✓		7777	0.658
2000 – 2002					✓		8142	0.637
2002 – 2004					✓		8498	0.628
2004 – 2006					✓		8208	0.567
2006 – 2008	✓						8194	0.520
2008 – 2010	✓						8813	0.559
2010 – 2012		✓					9078	0.576
2012 – 2014						✓	9310	0.639
2014 – 2016						✓	9199	0.640
Average							7704	0.620
Total							115563	

Notes: The table reports the hedonic specifications used to construct the hedonic index as well as the number of observations n entering each model and the adjusted R^2. Each model pools data from two consecutive survey waves, includes time dummies and is estimated via OLS. The definitions of the variables *Location of the dwelling 1/2* and *Rating of the Property 1/2* are provided in Table B.3.

B.2.4 Methodological Notes on the S-RPPI

The hedonic index. From 1986, the survey contains a sufficient number of dwelling characteristics that can be used to estimate hedonic models. Due to the frequent changes in variables and definitions, the hedonic equation is not fixed but allowed to flexibly change over time. We make use of six specifications that comprehensively exploit the information collected in the survey (see Table B.4).

Due to the change in the specifications, we do not report detailed regression results here (see Table B.4 for some general results). However, coefficients are as expected: the price increases with increasing living surface, when there are more bathrooms, and the rating of the property is higher. The functional form of age is roughly quadratic. The locational characteristics including the interactions are highly significant.

The ROP index. The index links OEVs reported for the same dwelling over time. There are 49,389 observations that appear at least twice in the sample. These observations refer to 13,844 unique dwellings. We construct 35,545 price pairs, which enter the repeat-sales model.

The ROA index. The index links a dwellings acquisition price to the OEV. The earliest acquisition price is reported for 1900, however, due to low numbers of observations, we only consider acquisition prices from 1960 onward. We disregard observations with an acquisition price of less than EUR 1,000, where the acquisition price is missing, or where the ratio between OEV and acquisition price is less than 0.5.

We use 23,898 price pairs to estimate the ROA index.

B.3 Federal Reserve Bank of New York's Survey of Consumer Expectations

B.3.1 Description of the data

The Federal Reserve Bank of New York's Survey of Consumer Expectations (SCE) is an internet-based survey of a rotating panel of approximately 1,200 household heads from across the U.S., with the goal of eliciting expectations about a variety of economic variables, such as inflation and labour market conditions. Respondents participate in the panel for up to twelve months, with a roughly equal number rotating in and out of the panel each month. Respondents are invited to participate in at least one survey each month.

B.3.2 Elicitation of the OEV

Owners in the housing module are asked to report the current value of their home via the following question: *What do you think your home would sell for today?*

B.3.3 Measure of Expectations

Owners in the housing module of SCE were asked to report the current and the expected values of their home via the following questions: *What do you think your home would sell for today? And in one year from today?.* We use the response to the second question to measure individual expectations. We use a similar questions to elicit expectations regarding the future value of the house in five years when available.

[13] `https://www.bancaditalia.it/pubblicazioni/indagine-famiglie/index.html`, last accessed on May 29, 2019.

Table B.5: Descriptive Statistics - SCE national sample

	Mean	S.D.	Min	Max
Dependent Variable:				
OEV (log)	12.18	0.91	3.53	17.91
House Characteristics:				
Location of the Dwelling:				
City Center	0.24	0.37	0	1
Suburban (less than 20km from City Center)	0.36		0	1
Suburban (more than 20km from City Center)	0.10		0	1
Small Town	0.18		0	1
Rural	0.11		0	1
Other	0.01		0	1
Type of Residence:				
House	0.69		0	1
Apartment	0.19		0	1
Townhouse	0.06		0	1
Mobil Home	0.04		0	1
Other	0.02		0	1
Age of the Dwelling:				
0-5 years ago	0.05		0	1
5-20 years ago	0.28		0	1
20-40 years ago	0.30		0	1
40+ years ago	0.37		0	1

Notes: These numbers refer to the national sample of the Federal Reserve Bank of New York's Survey of Consumer Expectations.

B.4 Eurosystem Household Finance and Consumption Survey

B.4.1 Description of the data

The HFCS is an ex-ante harmonized survey on households' finances and consumption carried out by the national central bank or national statistical office of all participating countries. The survey is coordinated by the European Central Bank, which also publishes the harmonized survey data we use here.

Currently, there are two waves available, which were released respectively in April 2013 and December 2016. Since the fieldwork periods differ from country to country, the waves refer to different years (see Table B.6). Due to the heterogeneity of the data, we refrain from reporting classical summary statistics but refer the interested reader to the online documentation.[14]

In most countries, where a household wealth survey already existed (like Italy), the national surveys were usually amended and extended to fit into the HFCS. Usually, national surveys are more comprehensive than the harmonized HFCS version.

B.4.2 Elicitation of the OEV

Owner-occupier are asked to estimate the current value of their home via the following question: *What is the value of this property, i.e. if you could sell it now how much do you think would be the price of it?*

B.4.3 Methodological Notes on the S-RPPI

Country-specific ROA indices. The harmonized HFCS data lacks sufficient hedonic characteristics. In particular, there is no information on the location of the dwelling. Also, there is no panel component. Hence we only estimate an adjusted ROA index.

Table B.6 reports the number of waves that are available for each country as well as the respective fieldwork period. Additionally, the number of observations used for constructing the index are reported. These numbers differ strongly across countries, which is due to three reasons: there is substantial variation in the size of European countries, the share of owner-occupiers as compared to renter in the population varies a lot, and also the relative size of the survey sample differs across countries. Hence, the number of observations is related to the number of households as well as the number of owner-occupiers. From these shares, one expects less reliable indices for Germany and Poland. This is indeed what we find (see Figure A.1).

We construct adjusted ROA indices. Hence, index numbers referring to survey years are left out and substituted by interpolated values. Since some of the survey years coincides with the peak of a housing boom (e.g., in Spain or Greece), the indices are *per construction* unable to properly detect them. The more survey waves become available, the more acquisition prices will be available for neighbouring year, which will ultimately increase precision.

The euro area ROA index. The European Central Bank also compiles an aggregate RPPI for the euro area (EA). We use the index that reports changes in house prices for EA-17, i.e., the euro area as composed in 2011 including 17 countries. The EA-RPPI is a weighted average of individual, country-specific RPPIs. Weights are determined by the contribution of a country's GDP to the overall EA-17 GDP.

To create a subjective counterpart, we calculate a weighted average over all ROA indices be-

[14]`https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html`, last accessed on May 29, 2019.

Table B.6: Summary of HFCS Data Used to Construct a ROA

Country	Country Code	Obs. for Index	Share of households	Share of OOH	ROA Start	Fieldwork periods Wave 1	Wave 2	EA-17
Austria	AT	826	0.214	0.449	1970	2011	2014	✓
Belgium	BE	2035	0.424	0.603	1960	2010	2014	✓
Cyprus	CY	501	1.652	2.249	1995	2010	2014	✓
Germany	DE	2166	0.055	0.123	1968	2011	2014	✓
Estonia	EE	806	1.409	1.843	1993	–	2013	✓
Spain	ES	8939	0.513	0.617	1950	2009	2012	✓
France	FR	4462	0.154	0.262	1970	–	2014	✓
Greece	GR	1586	0.372	0.516	1960	2010	2014	✓
Ireland	IE	2256	1.335	1.894	1965	–	2013	✓
Italy	IT	6194	0.251	0.368	1950	2010	2015	✓
Luxembourg	LU	1047	4.963	7.341	1974	2010	2014	✓
Latvia	LV	458	0.553	0.727	1992	–	2014	.
Malta	MT	843	5.288	6.593	1970	2010	2014	✓
Netherlands	NL	1652	0.218	0.379	1971	2010	2014	✓
Poland	PL	884	0.066	0.085	1996	–	2014	.
Portugal	PT	4200	1.045	1.399	1960	2010	2013	✓
Slovenia	SI	471	0.574	0.778	1990	2010	2014	✓
Slovakia	SK	1928	1.039	1.217	1970	2010	2014	✓

Notes: shares are in per mille and according to the HFCS 2nd wave. France participated in the first waves, but did not collect acquisition prices; Finland (part of EA-17) does not collect acquisition prices. OOH: owner-occupied households, fieldwork period: sometimes overlaps two years - every wave is allocated to the year where the major part of the fieldwork was carried out

longing to the EA-17. The weights are held constant over the period and refer to the 2013 GDP at market prices. As the HFCS lacks data to compile a ROA for Finland (no original acquisition price is available), this country is excluded. Finland's share of euro area GDP amounts to roughly 2% only, which is why the impact of this exclusion is expected to be minor.

C Objective RPPIs

The objective RPPIs are predominantly taken from the BIS data warehouse, that includes RPPIs for a long list of countries. We give preference to indices representing the entire country and including all types of dwellings (apartments and houses). For some countries more than one index fulfills these criteria whereas for other countries no index exists that has such a broad coverage. Table C.1 summarizes details about all objective RPPIs used in this article.

Table C.1: Objective residential property price indices.

ID	Country	Coverage	Frequency	Time frame	Quality Adjustment	Compiler	Source (if source different to compiler)	Source ID
AT 1	Austria	All dwellings; Whole country	Quarterly	2000 – 2017	Hedonic	NCB	BIS	Q:AT:0:1:0:0:6:0
BE 1	Belgium	All dwellings; Whole country	Quarterly	2005 – 2017	Hedonic	NSI	BIS	Q:BE:0:1:0:1:6:0
BE 2	Belgium	Existing dwellings; Whole country	Quarterly	1973 – 2017	Geographical stratification	NCB	BIS	Q:BE:0:1:1:0:0:0
BE 3	Belgium	Diverse	Annual	1878 – 2012	Diverse	Knoll et al. (2017)		
CY 1	Cyprus	All dwellings; Whole country	Quarterly	2002 – 2017	Hedonic	NCB	BIS	Q:CY:0:1:0:0:6:0
DE 1	Germany	All owner-occupied dwellings; Whole country	Annual	1975 – 2015	Diverse	NCB	BIS	A:DE:0:1:0:0:6:0
DE 2	Germany	All dwellings; Whole country	Quarterly	2000 – 2017	Hedonic	NSI	BIS	Q:DE:0:1:0:1:6:0
DE 3	Germany	Diverse	Annual	1870–1938, 1962 – 2012	Diverse	Knoll et al. (2017)		
EE 1	Estonia	All Dwellings; Whole country	Quarterly	2005 – 2017	Stratification	NSI	BIS	Q:EE:0:1:0:1:1:0
ES 1	Spain	All dwellings; Whole country	Quarterly	1995 – 2018	Stratification	General Government	BIS	Q:ES:0:1:0:3:1:0
ES 2	Spain		Annual	1900 – 2016		Francisco Amaral	Macrohistory Database (Jordà et al., 2017)	
FR 1	France	All dwellings; Whole country	Quarterly	2006 – 2017	Hedonic	NSI	BIS	Q:FR:0:1:0:1:6:0
FR 2	France	Diverse	Annual	1870 – 2012	Diverse; Mainly Repeat-sales	Knoll et al. (2017)		
GR 1	Greece	All dwellings; Whole country excl. capital	Quarterly	1994 – 2017	No	NCB	BIS	Q:GR:1:1:0:0:1:0
GR 2	Greece	Flats; Capital city and suburbs	Quarterly	1997 – 2017	No	NCB	BIS	Q:GR:3:8:0:0:1:0
HU 1	Hungary	All dwellings; Whole country	Quarterly	2007 – 2017	Hedonic	NSI	BIS	Q:HU:0:1:0:1:1:0
IE 1	Ireland	All dwellings Whole country	Monthly	2005 – Feb 2018	Hedonic	NSI	BIS	M:IE:0:1:0:1:0:0
IE 2	Ireland	Dublin only;	Annual	1950 – 2014	Hedonic	Keely and Lyons (2019)		
IT 1	Italy	All dwellings; Whole country	Quarterly	1990 – 2017		NCB	BIS	Q:IT:0:1:0:0:6:0

IT 2	Italy	Whole country	Annual	1927 – 2012	Diverse	Cannari et al. (2016)		
LU 1	Luxembourg	All dwellings; Whole country	Annual	1974 – 2015		NSI	BIS	A:LU:0:1:0:1:0:0
LV 1	Latvia	All dwellings; Whole country	Quarterly	2006 – 2017	Hedonic	NSI	BIS	Q:LV:0:1:0:1:6:0
MT 1	Malta	All dwellings; Whole country	Quarterly	2000 – 2017	No	NCB	BIS	Q:MT:0:1:0:0:0:0
NL 1	Netherlands	All dwellings; Whole country	Quarterly	2005 – 2017		NSI	BIS	Q:NL:0:1:0:1:6:0
NL 2	Netherlands	Existing single-family; houses; Whole country	Quarterly	1995 – 2017		NSI	BIS	Q:NL:0:2:1:1:0:0
PL 1	Poland	Flats; Whole country	Quarterly	2010 – 2017	Stratification	NSI	BIS	Q:PL:0:8:0:1:6:0
PL 2	Poland	Single-family houses Whole country	Annual	2003 – 2016	Average price	NSI	BIS	A:PL:0:2:0:1:1:0
PT 1	Portugal	All dwellings; Whole country	Quarterly	2008 – 2017	Hedonic	NSI	BIS	Q:PT:0:1:0:1:6:0
PT 2	Portugal		Annual	1971 – 2016		F. Amaral	Macrohistory Database (Jordà et al., 2017)	
SI 1	Slovenia	All dwellings; Whole country	Quarterly	2007 – 2017	Hedonic	NSI	BIS	Q:SI:0:1:0:1:6:0
SK 1	Slovakia	All dwellings; Whole country	Quarterly	2005 – 2017	Per Square meter prices	NCB	BIS	Q:SK:0:1:0:2:1:0
US 1	US	Single-family homes Whole country	Monthly	1987 – 2018	Repeat-sales	Core Logic	Fed St. Louis	CSUSHPINSA
US 2	US	Single-family house Whole country	Quarterly	1975 – 2018	Repeat-sales	FHFE	Fed St. Louis	USSTHPI
US 3	US	All dwellings Whole country	Monthly	1996 – 2019	Hedonic Imputation	Zillow	Zillow Home Value Index	
US 4	US	Diverse	Annual	1890 – 2012	Diverse	Knoll et al. (2017)		
EA 1	Euro Area	EA-17; Average over national indices	Quarterly	1980 – 2018	Diverse	ECB		RPP.Q.I6.N.TD.00.3.00

Notes: The table provides overview information about the objective indices used in this article. Long series usually include structural breaks due to changes in methodologies and/or data sources. In-depth methodological descriptions are provided by the respective index compilers. NSI refers to *National Statistical Office* and NCB to *National Central Bank*. In the case of several indices, preference has been given to long series with the broadest coverage in terms of types of dwellings and geographical coverage. The end date refers to the last year of the index used in this article.

D What disturbs ROA indices? A Simulation

The ROA approach combines objective price information – the acquisition price – with subjective ones – the OEV. In contrast to the other S-RPPIs, the ROA index does not match O-RPPIs well as it shows unrealistic spikes in the years where the survey was carried out.

We simulate the scenarios, where either OEVs are shifted upwards due to a one-time premium (a "level effect," which would be consistent with an endowment effect), or OEVs are shifted upwards due to an inflated belief of actual appreciation rates.

We perform a simulation, which is calibrated to the US housing market. We simluate log-normally distributed house prices with parameters μ_p and σ_p,

$$P_0 = \exp\left(\mu_p + \sigma_p \cdot Z\right),$$

where $Z \sim \mathcal{N}(0, 1^2)$. The parameters μ_p and σ_p are chosen to match the observed mean and variance of reported acquisition prices for 2000 in the AHS survey, i.e.,

$$\mathbb{E}[P_0] = \exp\left(\mu_p + \frac{\sigma_p^2}{2}\right) = 175,601 \text{ USD} \quad \text{and}$$
$$\sqrt{Var[P_0]} = \sqrt{\Big(\exp(\sigma_p^2) - 1\Big) \cdot \exp(2\mu_p + \sigma_p^2)} = 155,636 \text{ USD}.$$

We simulate $n = 100,000$ prices from P_0 denoted by $P_0(i)$, for $i \in \{1, \ldots, n\}$. We randomly assign each home a year of acquisition $t_a(i)$ between 1975 and 2012, and a year where the owner is surveyed $t_s(i)$. We assume four survey waves taking place in 2006, 2008, 2010 and 2012. We drop all observations where $t_a > t_s$.

We impute the purchase price by projecting backwards $P_0(i)$ using the annualized FHFE index. We normalize the FHFE index to the year 2000 and denote the index numbers by $I(t)$, i.e., $I(2000) = 1$. Thus,

$$P_a(i) = P_0(i) \cdot I(t_a(i)).$$

Not that we here implicitly assume that the acquisition price is perfectly objective.

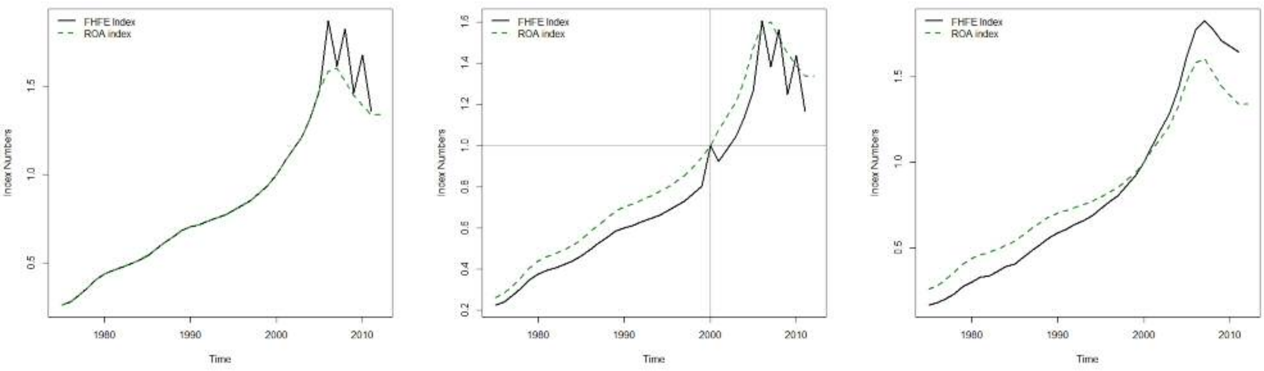
In contrast, for survey years we mimic an OEV, which is subjective and assumed to be affected by a systematic upward bias. We simulate this by adding a one-time premium that is specific for every person ("subjective") but on average positive ("upward bias"). Agarwal (2007) reports substantial average absolute mis-estimation of owners ranging between 14% and 25%. We hence choose $\rho_i \sim \mathcal{N}(0.2, 0.05^2)$ and compute the OEV via

$$P_s(i) = P_0(i) \cdot I(t_s(i)) \cdot (1 + \rho_i). \tag{11}$$

We use the resulting price pairs to estimate a ROA index. The result is shown in the left panel of Figure D.1. The pattern matches exactly the observed ones in the ROA indices for the US and Italy: in survey years the index jumps up due to the inflated subjective estimates. In non-survey years, the index is determined by the acquisition prices not confounded by the premium added to OEVs.

This has two implications: first, ROA index numbers for years where a survey took place are not reliable. In general, a ROA is particularly useful for the period preceding the first survey year. Second, a survey year may not be chosen as a base period used to normalize the index, as the entire index series will be shifted (see middle panel of Figure D.1). Preferably, a year preceding the first survey year should be used for normalization.

Figure D.1: Simulation: the Effect of a One-Time Premium.



Notes: The figures show an ROA index estimated from simulated data calibrated to the US housing market. We assume four survey years (2006, 2008, 2010 and 2012).
In the left and middle panel, the simulation includes a one-time premium for OEVs. The right panel includes a premium on annual appreciation rates. The middle panel includes an additional survey year 2000, which also acts as base period for the index.
All figures also show the U.S. Federal Housing Finance Agency All-Transactions House Price Index (FHFE) house price index.

To rule out an alternative explanation, namely that owners tend to think that their own asset, i.e., their home, outperformed the market, we also simulate a second type of systematic bias.

We therefore repeat the same exercise but change the way the OEV is constructed. Instead of a one-time premium added to the "objective market value," the premium depends on the ownership duration. Hence we simulate the setting that an owner beliefs that the annual appreciation rate is inflated.

Acknowledging that owners may have diverse opinions on appreciation rates, we assume that the annual actual appreciation rate is on average slightly inflated (here by 2%) but the variance of this type of premium is large and people may even have a negative premium. We implement this by setting $\rho'_i \sim \mathcal{N}(0.02, 0.05^2)$ and changing equation to (11)

$$P_s(i) = P_0(i) \cdot I(t_s(i)) \cdot (1 + \rho'_i)^{t_s - t_a}.$$

As a result, the ROA index is shifted, but the observed pattern including spikes is missing (see right panel of Figure D.1). We hence rule out an overly optimistic belief in appreciation rates as the prime explanation for the observed pattern in the ROA index.

E Tracking Sentiments: The Role of Expectations – Additional Tables

Table E.1: Descriptive Statistics - US Sample

	Mean	S.D.	Min	Max
Dependent Variable:				
OEV (log)	12.24	0.74	9.10	14.91
Expectations:				
Optimism	0.51		0	1
Pessimism	0.04		0	1
Neutral	0.46		0	1
Individual Characteristics:				
Female	0.56		0	1
Age	52.83	14.18	22	90
Married	0.74		0	1
Employed	0.54		0	1
Education:				
High School	0.10		0	1
Some College	0.32		0	1
College	0.58		0	1
Annual Household Income:				
Under 50k	0.28		0	1
50k to 100k	0.38		0	1
Over 100k	0.34		0	1
House Characteristics:				
Location of the Dwelling:				
City Center	0.17	0.37	0	1
Suburban (less than 20km from City Center)	0.39		0	1
Suburban (more than 20km from City Center)	0.12		0	1
Small Town	0.19		0	1
Rural	0.13		0	1
Other	0.01		0	1
Type of Residence:				
House	0.88		0	1
Apartment	0.06		0	1
Townhouse	0.06		0	1
Age of the Dwelling:				
0-5 years ago	0.05		0	1
5-20 years ago	0.29		0	1
20-40 years ago	0.30		0	1
40+ years ago	0.36		0	1

Notes: These numbers refer to the estimation sample of 2,661 observations coming from the Housing Module of the Federal Reserve Bank of New York's Survey of Consumer Expectations used in the empirical analysis in Table 3.

Table E.2: Descriptive Statistics - Italian Sample

	Mean	S.D.	Min	Max
Dependent Variable:				
OEV (log)	12.05	0.62	9.21	15.07
Expectations:				
Optimism	0.08		0	1
Pessimism	0.27		0	1
Neutral	0.65		0	1
Individual Characteristics:				
Female	0.47		0	1
Age	62.65	14.16	21	90
Employed	0.38		0	1
Marital Status:				
Married	0.59		0	1
Single	0.13		0	1
Divorced	0.07		0	1
Widow	0.21		0	1
Education:				
None	0.03		0	1
Primary School	0.21		0	1
Lower Secondary School	0.34		0	1
Vocational Diploma	0.28		0	1
Upper Secondary School	0.13		0	1
University Degree or Higher Education	0.01		0	1
Annual Household Income (log)	10.30	0.63	0	12.99
House Characteristics:				
Surface (square meter, log)	4.63	0.38	3	6.91
Age of the Dwelling (years)	53.13	38.12	0	326
More than 2 bathrooms	0.53		0	1
Rating of the Property:				
Luxury	0.01		0	1
Highly Desirable	0.16		0	1
Mid Range	0.68		0	1
Modest	0.11		0	1
Low-Income	0.04		0	1
Very Low-Income	0.01		0	1
Size of the Municipality				
Up to 20,000 inhabitants	0.27		0	1
From 20,000 to 40,000	0.19		0	1
From 40,000 to 500,000	0.46		0	1
More than 500,000	0.08		0	1
Location of the Dwelling:				
Isolated Area	0.04		0	1
Town Outskirts	0.33		0	1
Between Outskirts and Town Center	0.29		0	1
Town Center	0.20		0	1
Hamlet	0.13		0	1

Notes: These numbers refer to the estimation sample of 10,646 observations coming from the Survey on Household Income and Wealth used in the empirical analysis in Table 3

Table E.3: Owner Estimated Value (OEV) and Price Expectations: OLS results - Original Elicitation of Expectations

	Self-reported Housing Price (log)		
	Italy		U.S.
	2014	2016	
	(1)	(2)	(3)
Optimism	0.081***		
	(0.026)		
Pessimism	-0.023+		
	(0.013)		
More than 5% (percentage point)		0.001+	
		(0.000)	
Between 2% and 5% (percentage point)		0.001**	
		(0.000)	
Between -2% and -5% (percentage point)		-0.000	
		(0.000)	
More than -5% (percentage point)		-0.000	
		(0.000)	
Value of the house in the future$^+$ (log)			0.008***
			(0.002)
Value of the house in the future$^-$ (log)			-0.004
			(0.005)
Observations	5201	5445	2661
Adjusted R^2	0.577	0.653	0.495
House Characteristics	✓	✓	✓
Individual Controls	✓	✓	✓

Notes: Standard errors are clustered at the individual level. All regressions include year dummies. In the U.S. panel, the house characteristics includes dummies for the size of the the municipality, dummies for the type of residence, dummies for the age of the residence and state*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies, an employment dummy and household income categories. In the Italian panel, the house characteristics includes the surface of the dwelling (in log), the number of bathrooms, dummies for the rating of the property, dummies for the size of the municipality, the age and the age squared of the residence and large city*region*year fixed effects and the individual controls include the age, the age squared, the gender, the marital status, education dummies and the household income (in log). Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^+$ if the p-value is lower than 0.1.

Table E.4: Owner Estimated Value (OEV) and Price Expectations: OLS results with full controls - US

	OEV (log)			
	(1)	(2)	(3)	(4)
Optimism	0.114***	0.107***	0.061*	0.061**
	(0.027)	(0.023)	(0.024)	(0.021)
Pessimism	-0.204**	-0.094	-0.169**	-0.088
	(0.074)	(0.063)	(0.064)	(0.057)
City Center (ref. category)		.		.
		(.)		(.)
Suburban 20-		0.146***		0.090**
		(0.034)		(0.030)
Suburban 20+		0.228***		0.138***
		(0.045)		(0.040)
Small Town		-0.138***		-0.118***
		(0.040)		(0.035)
Rural		-0.006		0.007
		(0.046)		(0.041)
Other		-0.062		-0.128
		(0.168)		(0.148)
House (ref. category)		.		.
		(.)		(.)
Apartment		-0.230***		-0.062
		(0.053)		(0.048)
Townhouse		-0.159**		-0.081^{+}
		(0.052)		(0.046)
0-5 years ago (ref. category)		.		.
		(.)		(.)
5-20 years ago		-0.167**		-0.189***
		(0.056)		(0.050)
20-40 years ago		-0.379***		-0.347***
		(0.056)		(0.050)
40+ years ago		-0.543***		-0.431***
		(0.057)		(0.051)
Married			0.100***	0.101***
			(0.029)	(0.026)
Age of the respondent			0.007	0.008^{+}
			(0.006)	(0.005)
Age of the respondent Squared			0.000	-0.000
			(0.000)	(0.000)
Female			0.065**	0.059**
			(0.024)	(0.021)
Employed			-0.123***	-0.064*
			(0.028)	(0.025)
Under 50k (ref. category)			.	.
			(.)	(.)
50k to 100k			0.292***	0.228***
			(0.031)	(0.027)
Over 100k			0.734***	0.535***
			(0.035)	(0.032)
High School (ref. category)			.	.
			(.)	(.)
Some College			0.071^{+}	0.057
			(0.043)	(0.038)
College			0.303***	0.255***
			(0.042)	(0.037)
Constant	12.229***	12.578***	11.183***	11.593***
	(0.020)	(0.062)	(0.150)	(0.139)
Observations	2661	2661	2661	2661
Adjusted R^2	0.011	0.297	0.340	0.492

Notes: Standard errors are in parentheses. All regressions include year fixed effects. Columns (2) and (4) include region*year fixed effects. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^{+}$ if the p-value is lower than 0.1.

Table E.5: Owner Estimated Value (OEV) and Price Expectations: OLS results with full controls - Italy

	OEV (log)			
	(1)	(2)	(3)	(4)
Optimism	0.113***	0.068***	0.091***	0.066***
	(0.023)	(0.015)	(0.020)	(0.015)
Pessimism	-0.087***	-0.025**	-0.030*	-0.013
	(0.014)	(0.009)	(0.012)	(0.009)
Surface (in log)		0.715***		0.652***
		(0.013)		(0.013)
More than 2 bathrooms		0.183***		0.150***
		(0.010)		(0.010)
Rating of the House:				
Very Low-Income (ref. category)		0.000		0.000
		(.)		(.)
Low-Income		0.113*		0.091*
		(0.047)		(0.046)
Modest		0.245***		0.219***
		(0.044)		(0.043)
Mid Range		0.414***		0.359***
		(0.043)		(0.042)
Highly Desirable		0.642***		0.560***
		(0.044)		(0.044)
Luxury		0.813***		0.705***
		(0.057)		(0.056)
Up to 20,000 inhabitants (ref. category)		0.000		0.000
		(.)		(.)
From 20,000 to 40,000		-0.277***		-0.231***
		(0.018)		(0.018)
From 40,000 to 500,000		-0.250***		-0.218***
		(0.016)		(0.016)
Hamlet (ref. category)		0.000		0.000
		(.)		(.)
Isolated Area		-0.007		0.006
		(0.023)		(0.022)
Town Outskirts		0.016		0.010
		(0.013)		(0.013)
Between Outskirts and Town Center		0.033*		0.017
		(0.013)		(0.013)
Town Center		0.081***		0.056***
		(0.014)		(0.014)
Age of the dwelling		-0.002***		-0.001***
		(0.000)		(0.000)
Age of the dwelling Squared		0.001***		0.000***
		(0.000)		(0.000)
Female			0.000	0.003
			(0.011)	(0.008)
Age of the respondent			0.003	0.003
			(0.003)	(0.002)
Age of the respondent Squared			0.000	-0.000
			(0.000)	(0.000)
Employed			-0.085***	-0.038***
			(0.015)	(0.011)
No Diploma (ref. category)			.	.
			(.)	(.)
Primary School			0.067	0.049
			(0.045)	(0.032)
Lower Secondary School			0.202***	0.103**
			(0.043)	(0.032)
Vocational Diploma			0.289***	0.139***
			(0.044)	(0.032)
Upper Secondary School			0.425***	0.172***
			(0.046)	(0.033)
University Degree/Higher Education			0.869***	0.322***
			(0.056)	(0.041)
Widow (ref. category)			.	.
			(.)	(.)
Married			0.021	0.001
			(0.016)	(0.012)
Single			-0.054**	-0.013
			(0.021)	(0.015)
Divorced			0.020	-0.009
			(0.024)	(0.018)
Household Income (in log)			0.365***	0.106***
			(0.010)	(0.008)
Constant	12.067***	8.424***	8.344***	7.688***
	(0.007)	(0.072)	(0.136)	(0.118)
Observations	10650	10650	10650	10650
Adjusted R^2	0.007	0.599	0.263	0.615

Notes: Standard errors in parentheses are clustered at the individual level. All regressions include year fixed effects. Columns (2) and (4) include region*year fixed effects. Statistical significance is coded following the standard notation: *** if the p-value is lower than 0.001, ** if the p-value is lower than 0.01, * if the p-value is lower than 0.05, $^+$ if the p-value is lower than 0.1.