RESERVE BANK
OF AUSTRALIA

Research Discussion Paper

Home Price Beliefs in Australia

Callan Windsor, Gianni La Cava and James Hansen

RDP 2014-04

The Discussion Paper series is intended to make the results of the current economic research within the Reserve Bank available to other economists. Its aim is to present preliminary results of research so as to encourage discussion and comment. Views expressed in this paper are those of the authors and not necessarily those of the Reserve Bank. Use of any results from this paper should clearly attribute the work to the authors and not to the Reserve Bank of Australia.

ISSN 1320-7729 (Print)

ISSN 1448-5109 (Online)

Home Price Beliefs in Australia

Callan Windsor, Gianni La Cava and James Hansen

Research Discussion Paper
2014-04

May 2014

Economic Research Department
Reserve Bank of Australia

We would like to thank Katherine Leong for her valuable contribution to earlier versions of this work. We are also grateful to Alexandra Heath, Christopher Kent, Tony Richards and Peter Tulip, as well as Roger Wilkins and other seminar participants at the biennial HILDA Survey Research Conference 2013. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Reserve Bank of Australia. The authors are solely responsible for any errors.

Authors: windsorc, lacavag and hansenj at domain rba.gov.au

Media Office: rbainfo@rba.gov.au

Abstract

We document some new stylised facts about how Australian homeowners value their homes using household panel data and unit-record data on home sale prices. We find that homeowners' price beliefs are unbiased at the postcode level, on average, although there is considerable dispersion in the difference between beliefs and prices across postcodes. Household characteristics, such as age and tenure, and the regional unemployment rate are correlated with differences between beliefs and prices. We also find evidence that the difference between beliefs and prices has explanatory power for average household consumption, leverage and portfolio decisions after controlling for the market-inferred value of the home. These facts provide empirical evidence to support recent literature on the importance of belief formation for household decision-making.

JEL Classification Numbers: C33, D8, E21, G11, G12

Keywords: housing, consumer behaviour, beliefs

Table of Contents

1.	Introduction		1
2.	Background and Existing Literature		4
3.	Data		7
	3.1	Home Sale Prices	7
	3.2	Self-assessed Home Values	7
4.	Hedonic Methodology		8
5.	Measuring Home Valuation Differences		11
	5.1	Distribution of Home Valuation Differences	11
	5.2	Significance of Home Valuation Differences	13
6.	Determinants of Home Valuation Differences		15
	6.1	Panel Regressions	15
	6.2	Quantile Panel Regressions	18
7.	Financial Decisions and Home Valuation Differences		20
	7.1	Homeowners	20
	7.2	Renters	23
8.	Robustness Tests		25
9.	Conclusion		27
Appendix A: Hedonic Model Output			29
Appendix B: Estimating Homeowner Recollection Bias			30
Appendix C: Quantile Regression Output			32
Appendix D: Economic Decisions Output			33

Appendix E: Additional Robustness 34

References 38

Copyright and Disclaimer Notices 41

Home Price Beliefs in Australia

Callan Windsor, Gianni La Cava and James Hansen

As the real estate agent said, 'Location, location, location', and we're right next door to the airport. It will be very convenient if we ever have to fly one day. Dale Kerrigan, *The Castle* (1997)

1. Introduction

Housing is the largest component of household wealth in Australia. Variation in housing prices has been shown to be important for household leverage, portfolio allocation decisions and consumption (Ellis, Lawson and Roberts-Thomson 2003; Kohler and Smith 2005; Berger-Thomson, Chung and McKibbin 2009; Windsor, Jääskelä and Finlay 2013). However, timely data on the prices of individual homes are not readily available. For this reason, households are typically required to infer or form a belief about the value of their home when making these economic decisions. As the quote above from the Australian film *The Castle* suggests, these beliefs can be very subjective.

This paper explores homeowners' beliefs about housing prices in Australia. Our goal is to understand differences between homeowners' beliefs and market-inferred home sale prices and whether these differences are important for economic decision-making. To this end, we estimate the average homeowner belief at the postcode level, and compare the distribution of these beliefs with the distribution of average prices inferred from transactions data. We refer to these differences between homeowner beliefs and market-inferred home sale prices as 'home valuation differences'. We also explore whether these valuation differences are correlated with certain demographic and economic variables and whether valuation differences are significant in explaining household consumption, leverage and portfolio allocation decisions across postcodes.

Our paper makes three contributions:

1. We estimate the difference between beliefs and prices (hereafter 'home valuation differences') in a way that is free of recollection bias. That is, unlike the previous literature, we construct our comparisons of beliefs to prices using

a method that does not rely on the ability of surveyed households to recall the purchase price of their homes. To do this, we use separate hedonic regressions on household survey data and on unit-record sale price data.

2. We regress our measure of home valuation differences on various household characteristics (e.g. age, income and education), the local area unemployment rate and a proxy for housing market information (the tenure of ownership).

3. We investigate whether the size of home valuation differences across postcodes is correlated with household spending, leverage and the share of risky assets held in households' financial portfolios.

Our approach allows us to focus not just on average beliefs about home prices, but also higher moments of the distribution of beliefs, and to relate any differences between beliefs and market values to households' economic decisions.

Understanding how well Australian homeowners assess the value of their homes is important for a number of reasons. First, self-assessed home values sourced from household surveys are the main source of data used to measure the distribution of household wealth (and related financial indicators, such as leverage) in Australia. If homeowners do not accurately value their homes, then survey measures of household wealth may be biased. For example, if home valuation differences vary systematically with age then the estimated age profile of household wealth using self-assessed home values will be biased, giving a misleading picture of the actual distribution of wealth by age.

Second, by focusing on the distribution of average differences in beliefs and prices across postcodes, our approach provides insight into alternative theories of homeowner belief formation. In particular, we consider whether beliefs are unbiased on average (rational) or whether there is skewness in beliefs that could reflect optimism or pessimism. Some models of decision-making under uncertainty that focus on factors such as robust control (Hansen and Sargent 2008; Bidder and Smith 2012) and ambiguity (Epstein and Schneider 2008) predict that some households may hold pessimistic beliefs and therefore undervalue their homes.

In contrast, Genesove and Mayer (2001) show that loss aversion may cause some homeowners to hold optimistic beliefs relative to market-inferred prices when prices are declining. Likewise, the recent literature on optimism and other rational biases (see, for example, Van den Steen (2004) and Brunnermeier and Parker (2005)) predicts that some households may hold optimistic beliefs and hence overvalue their homes. In particular, households may trade off the utility gains from optimism with any costs from making distorted decisions because of overvaluation. An appealing feature of our paper is that we can provide empirical evidence on these alternative theories of belief formation.

To begin with, we document some new stylised facts about Australian homeowners' ability to value their own homes. Home valuation differences, measured at the postcode level, are defined as the difference between the average of homeowners' beliefs about the value of their homes and the average price in the same postcode based on transactions data. Both measures are constructed by controlling for the differing characteristics of properties over which beliefs are formed, or that are sold, in any given period. This is done in two steps. First, we estimate average homeowner housing price beliefs across postcodes using hedonic regressions on household survey data. Second, we estimate average market values across postcodes using separate hedonic regressions on home sales data.

We find that homeowners' home price beliefs are unbiased on average across postcodes. In terms of the absolute differences, we find that half of the average home valuations fall within 11 per cent of the average market value across postcodes. Although beliefs are unbiased on average, we do find statistically significant differences between average beliefs and average sale prices for many postcodes. In particular, a relatively large share of postcodes are undervalued (have a significant negative valuation difference) and a relatively large share of postcodes are overvalued (have a significant positive valuation difference).

Certain average household characteristics are correlated with valuation differences. In particular, postcodes with older homeowners are more prone to overvalue their homes, on average. In contrast, postcodes in which homeowners have lived in their homes for a relatively long time or in regions with relatively high unemployment are more likely to undervalue their homes, on average.

We also explore how home valuation differences are associated with households' consumption and financial decisions. We find evidence that valuation differences are positively associated with spending, leverage and the allocation of wealth to 'risky' assets, such as equities, after controlling for a number of other factors, including average income and the average sale price of homes in the postcode.

Importantly, we show that our results are unlikely to be due to omitted characteristics in the hedonic regression. Our key findings also hold under an alternative approach to estimating home valuation differences using repeat sales.

In Section 2 we discuss some of the existing literature and the motivation for our research. In Section 3 we discuss the data and in Section 4 we outline the hedonic regression modelling. In Section 5 we document the key stylised facts about the distribution of home valuation differences before turning to the determinants of these differences in Section 6. We explore the correlation between valuation differences and household decision-making in Section 7 before we consider the robustness of our findings in Section 8. We then draw conclusions in Section 9.

2. Background and Existing Literature

The idea that households form subjective beliefs about the value of their own home is intuitive. Unlike financial assets such as equities, housing is an asset that is relatively hard and costly to value.[1] The main reasons for this are:

1. Housing is a heterogenous asset that is not sold on a large centralised market. Instead, there are non-trivial search costs for buyers and sellers to successfully match either through bilateral negotiation or an auction.
2. Selling or valuation costs can be non-trivial.
3. Housing is traded infrequently. On average over the past decade, around 6 per cent of the total number of homes in Australia are sold each year;[2] suggesting an average holding period of about 16 years.
4. Housing price data are generally only available with a lag.

---

1 This is gradually changing over time with the advent of internet-based home valuation tools.

2 This estimate does not include all newly built homes and transfers within families that are not sales. This measure also masks significant differences in turnover across regions.

5. Obtaining detailed price information about values in one's locality can be expensive, and especially so if one wants to obtain updates of this information regularly.

The fact that housing is an asset that is comparatively more difficult to price than financial assets, and is a large component of households' wealth, implies that households infer or form beliefs about the value of this asset when making economic decisions. In particular, standard theory would suggest that when making either a portfolio allocation decision or choosing a consumption path, beliefs about the value of the home will influence these decisions.

This raises the main questions of interest in this paper. First, how well do homeowners' beliefs match objective measures of home prices? This is addressed in Section 5. Second, are certain household characteristics correlated with beliefs? More specifically, what characteristics explain whether a household overvalues, undervalues or correctly values their home? This is addressed in Section 6. Third, does the self-reported belief or the market-inferred value of the home matter for household economic decisions? That is, do households that undervalue their homes choose different consumption, leverage and portfolio allocations to households that overvalue their homes? This is addressed in Section 7.

A novel feature of our paper is the data we use. We gauge the accuracy of self-assessed home values using a near-census of housing sale prices as the benchmark measure of valuation. The early research instead compared estimates of housing prices by homeowners and professional appraisals (see Kish and Lansing (1954) and Kain and Quigley (1972)). However, Robins and West (1977) showed that homeowners and professional appraisers assess the value of homes with the same degree of inaccuracy.

The literature has also compared self-assessed home values to recalled sale prices. In these studies, homeowners that have recently moved are surveyed and asked to make an assessment of the current value of their homes, as well as recall the original sale price of their homes (e.g. Goodman and Ittner 1992). Local housing price indices are typically used to control for the passage of time between the current estimate and the initial purchase price. For instance, Goodman and Ittner (1992) find an implied home valuation difference of around 6 per cent (for similar approaches see Ihlanfeldt and Martinez-Vazquez (1986) and Kiel and Zabel (1999)).

In a similar vein, Melser (2013) assesses home valuation differences for Australia using household panel data. He compares self-assessed current home prices to the initial purchase prices recalled by surveyed homeowners. He finds that Australian homeowners have a positive bias of around 4 per cent in estimating the value of their homes.

The limitation of these approaches are the small samples of sale prices (generally less than 1 000 observations); their inability to distinguish between valuation bias and recollection bias;[3] and their reliance on external indices to update self-assessed home values.[4]

Relative to previous studies, one advantage of our approach is that we can focus on valuation bias, abstracting from either recollection bias or the use of external indices to infer a belief. Another advantage is that we have a near-census of sale prices as the benchmark measure of valuation. Finally, the timing of these sale prices are matched to the timing of self-assessed home values, thereby avoiding the use of external benchmarks to 'update' self-assessed home values.

---

3 In Appendix B we directly estimate the degree of recollection bias and find that surveyed homeowners understate the purchase price of their homes by about 3.4 per cent, on average.

4 A recent paper by Henriques (2013) is an exception. For a panel of non-moving US homeowners, Henriques compares the growth in self-assessed home values over the period 2007 to 2009 for each homeowner to the growth in regional house price indices. She finds that the median home valuation difference on the rate of change in housing prices is around 2.5 per cent. Agarwal (2007) is another exception. The benchmark housing price data used in this US study comes from homeowners' financial institutions, with the financial institution's estimate of the market value coming from the Case-Shiller repeat sales index. However, despite both of these studies having access to market values, they only have access to self-assessed home values for particular periods. Agarwal uses observations from households who engaged with a particular financial institution in 2002, while Henriques uses observations from homeowners who were surveyed in 2007 and 2009, in the midst of the US housing downturn.

3. Data

Our benchmark measure of the market value of the home is derived from unit-record data provided by Australian Property Monitors (APM).[5] Our measure of self-reported home valuations is obtained from the Household Income and Labour Dynamics in Australia (HILDA) Survey. Details on the exact location of properties are available in the APM sales dataset, but are not available in the HILDA Survey. Accordingly, we compare self-assessed home values to market values at the postcode level rather than at the level of the individual home.

To account for the fact that homes that are sold each period are potentially different to the homes that are self-assessed, we use hedonic regressions as discussed in Section 4 to control for the different characteristics of properties that are sold or valued.

3.1 Home Sale Prices

Data from APM – which constitutes a near-census of all housing sales – are used to estimate the value of households' homes. These data provide the benchmark for inferring the accuracy of self-assessed home values. The dataset contains nominal sale price information, including the exact location and home characteristics, for homes sold in Sydney, Melbourne and Brisbane. Table E1 details the construction of our samples, and Appendix E shows that our results are robust to the imposition of exclusion restrictions that reduce the effects of either small samples or outliers.

The sale price data cover the period 1992 to 2012 and comprise around 3.9 million observations. In comparing these prices to self-assessed home values we match the data by the postcode and the quarter in which the sale and valuation occurred.

3.2 Self-assessed Home Values

To measure households' valuations over their own home, data from the HILDA Survey over the period 2002 to 2011 are used. The survey is a nationally representative annual household panel that began in 2001 with around 7 700

5 For more information about these data see the Copyright and Disclaimer Notices at the end of this document.

responding households.[6] It contains questions on household finances, employment and wellbeing.

Every year surveyed homeowners are asked to provide a self-assessed home value by responding to the question:

Do you know what the approximate value of your home is? I mean, how much would it bring if you sold it today? Include land, home improvements, and fixtures (such as curtains and light fittings) usually sold with a home. Exclude home contents.

In total, over the period 2002 to 2011 there are 18 964 self-assessed home valuations.

4. Hedonic Methodology

A simple metric for measuring home valuation differences would be to compare, at a given point in time, the average self-assessed home value of each surveyed homeowner within a postcode (from the HILDA Survey) to the average price of all homes sold in that postcode (from the APM dataset). For example, using the available data for Sydney and quarterly time periods over the period 2002 to 2011, we would obtain 5 840 (= 40 quarters $\times$ 146 postcodes) price comparisons; the mean of which would measure the overall home valuation difference. However, at any point in time there could be systematic differences in the composition of properties self-assessed and sold within postcodes, thereby complicating this simple comparison.

To account for this, we use hedonic regression methods, which are common in the home-price measurement literature. The hedonic adjustment method determines, for each postcode and time period, the mean price of homes sold (or self-assessed) conditional on the characteristics of homes sold (or self-assessed). These 'conditional mean' estimates are designed to capture the level of home prices, holding constant any compositional changes in the homes sold or self-assessed each period. Therefore, the estimates measure the change in home values solely due to the passage of time.

6 We deliberately choose a longer sample period for the APM data than for the HILDA data in order to increase the precision of our hedonic estimates and to allow us to better construct repeat-sales estimates of home prices, which we use later in the paper.

We use postcode-time dummy hedonic models. The hedonic models for the sale price of homes and self-assessed home values are given by Equation (1) and Equation (2) respectively.[7]

$$ln(S_{ipt}) = \sum_{c=1}^{C} \beta_c^S y_{cit} + \sum_{t=1}^{T} \sum_{p=1}^{P} \lambda_{pt}^S D_{ipt} + \varepsilon_{ipt}^S \quad (1)$$

$$ln(V_{jpt}) = \sum_{c=1}^{C} \beta_c^V y_{cjt} + \sum_{t=1}^{T} \sum_{p=1}^{P} \lambda_{pt}^V D_{jpt} + \varepsilon_{jpt}^V \quad (2)$$

In Equation (1), S_{ipt} denotes the sale price of each home i in postcode p in quarter t. In Equation (2), V_{jpt} denotes the value of a self-assessed home j in postcode p in quarter t.

A home's characteristic is denoted by y_c for $c = 1, 2, ...C$. These characteristics are common across the sale and valuation equations and include dummy variables for the number of bedrooms and housing type (i.e. house or unit). The coefficient β_c captures the value of a given characteristic.[8]

The dummy variable, D_{ipt}, takes the value of 1 if the home sold or self-assessed is from postcode-time pt, and 0 otherwise. Our interest is in the coefficients on the postcode-time fixed effects for the log of sale prices (λ_{pt}^S) and the log of self-assessed home values (λ_{pt}^V). In a given time period, these measure the average value of homes sold or self-assessed within a postcode, after controlling for the characteristics of homes sold or self-assessed.

The effect of hedonic (or composition) adjustments on sale prices and self-assessed home values is shown in Figure 1 (regression output is shown in Table A1). The compositionally adjusted scatter plot (left panel) shows the

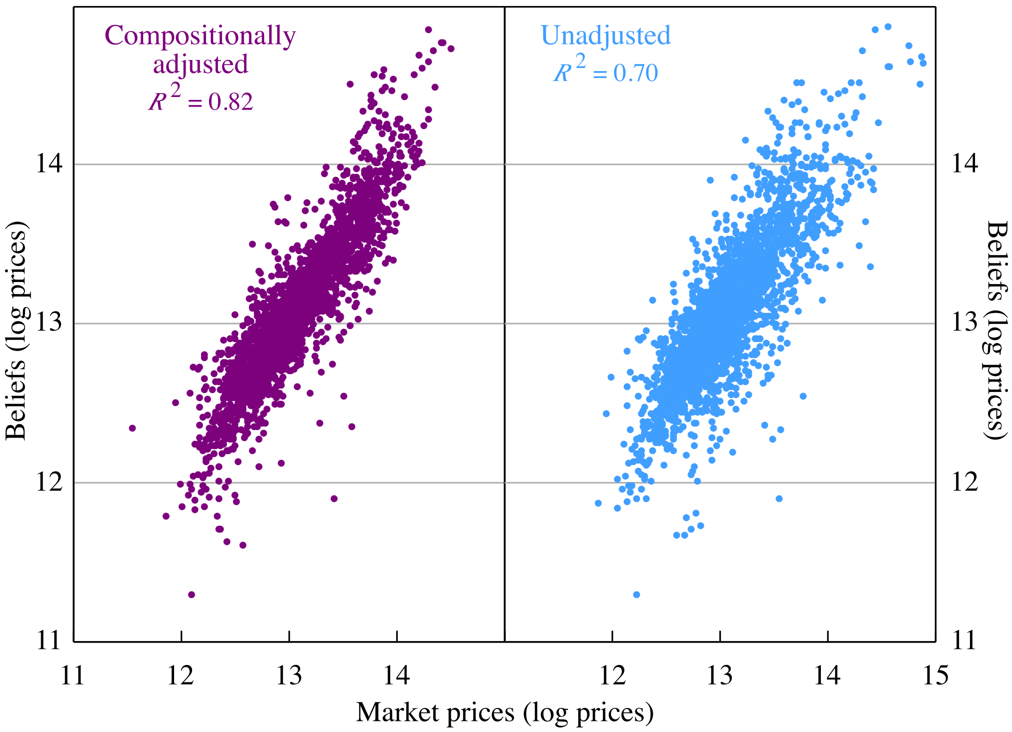
7 Following Hansen (2009), each capital city is treated as a separate market with the hedonic models estimated separately for Sydney, Melbourne and Brisbane.

8 In order to match the hedonic model that could be estimated for self-assessed home values using the HILDA dataset, a number of relevant housing characteristics – available in the APM dataset – have been excluded from Equation (1). These include the number of bathrooms, parking spaces and sale mechanism (i.e. auction or private treaty). Therefore, estimates of the value of a bedroom and the value of a postcode may be biased if these variables are correlated with any omitted characteristics in Equations (1) and (2). The robustness of our results to this potential source of bias is examined in detail in Section 8.

association between average sale prices (λ_{pt}^S) from Equation (1) and average self-assessed home values (λ_{pt}^V) from Equation (2) across postcodes. An unadjusted comparison is presented in the right panel, and shows the association between unconditional mean sale prices, $ln(\bar{S}_{pt}) = ln(\frac{1}{I_{pt}}\sum_{i=1}^{I_{pt}} S_{ipt})$, and self-assessed home values, $ln(\bar{V}_{pt}) = ln(\frac{1}{J_{pt}}\sum_{j=1}^{J_{pt}} V_{jpt})$, across postcodes.

Figure 1: Home Value Beliefs and Market Prices

By postcodes and time periods, 2002 to 2011



Note: Compositionally adjusted values refers to the mean of a three-bedroom house within postcode-time periods

Sources: APM; HILDA Release 11.0; authors' calculations

The correlation between adjusted prices is stronger than the correlation between unadjusted prices (0.91 versus 0.84). This suggests compositionally adjusting the data is important when estimating the difference between average home valuations and sales prices across postcodes.

5. Measuring Home Valuation Differences

5.1 Distribution of Home Valuation Differences

Using home valuation differences to make inferences about household belief formation is the focus of this paper. As discussed in the previous section, we could construct these valuation differences by comparing the (unconditional) average sale price to the average self-assessed home value within each postcode and time period. This comparison is given by Equation (3):

$$U_{pt} = ln(\bar{V}_{pt}) - ln(\bar{S}_{pt}). \tag{3}$$

However, this measure confounds a comparison of 'true' prices with changes in the composition of homes sold or self-assessed each period. For example, if the homeowners that drop out of the HILDA Survey over time are more likely to own smaller homes, then the size of homes self-assessed each year in the HILDA Survey will be increasing (attrition due to refusal, death or inability to locate the respondent is an issue in all panel surveys). Accordingly, for each postcode the change in average self-assessed home values each year will reflect 'true' price changes as well as the change in the composition of homes. This would, in turn, lead to valuation differences that are artificially inflated, with the bias increasing over time.

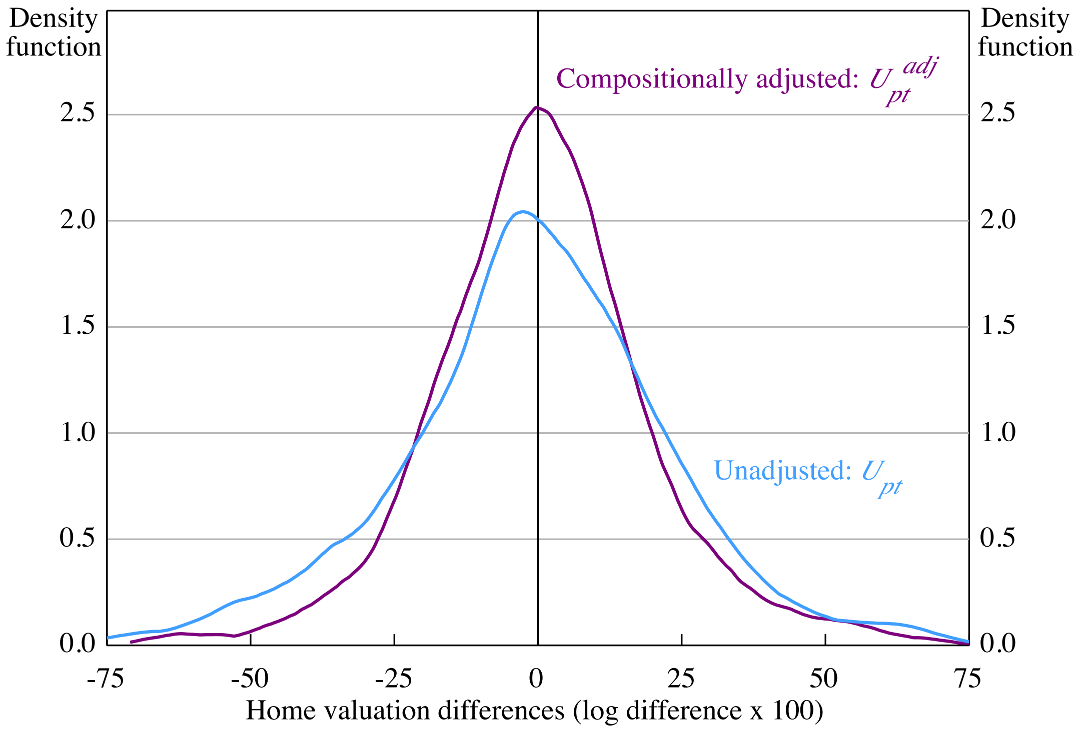
Compositionally adjusted estimates can be constructed by comparing the postcode-time dummy coefficients from Equation (1) and Equation (2):

$$U_{pt}^{adj} = \widehat{\lambda_{pt}^{V}} - \widehat{\lambda_{pt}^{S}}. \tag{4}$$

The density of U_{pt} versus U_{pt}^{adj} is plotted in Figure 2. In comparing these densities, two characteristics can be observed. First, both densities appear centered around zero. Second, the density of home valuation differences is less dispersed on a compositionally adjusted basis than on a raw basis with a standard deviation of 20 per cent versus 26 per cent.

Figure 2: Home Valuation Differences

Postcode level, 2002 to 2011



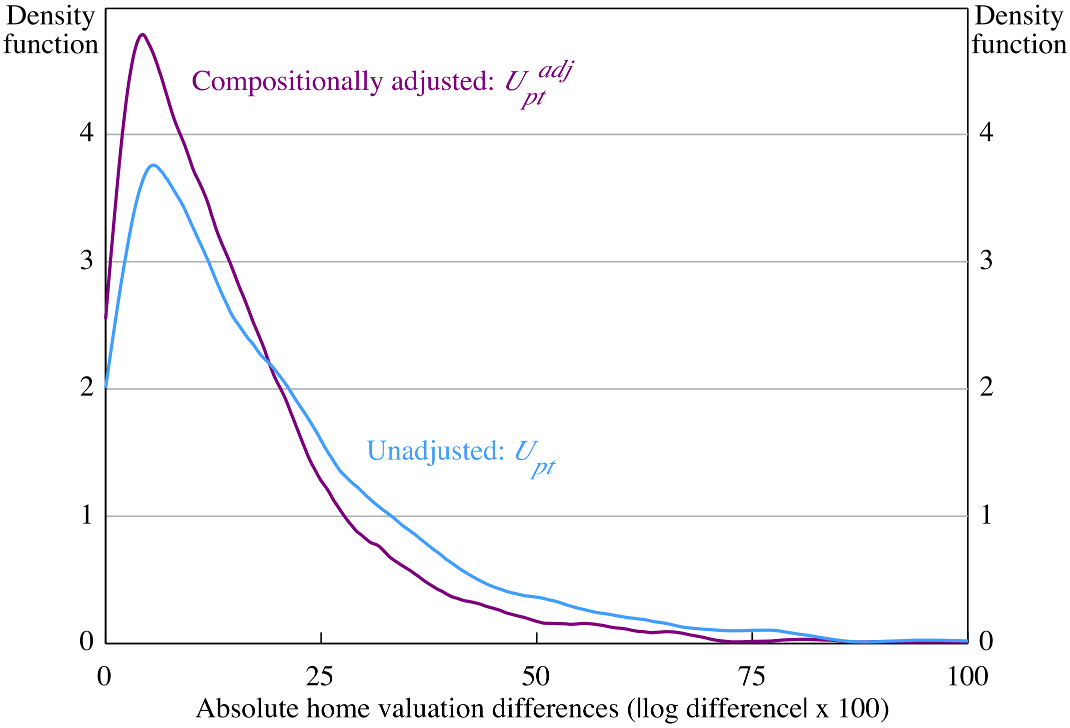
Note: Kernel density estimator

Sources: APM; HILDA Release 11.0; authors' calculations

The density of the absolute value of the home valuation differences is shown in Figure 3. On a compositionally adjusted basis, half of all postcodes provide an average valuation within 11 per cent of the average market value and three-quarters provide a valuation within 20 per cent.

The characteristics of home valuation differences (U_{pt}^{adj}) are summarised in Table 1. On average, across all cities, the differences have a mean that is positive at 0.5 per cent, but this is not statistically different to zero. This suggests that the average homeowner is unbiased in their subjective home valuation. However, this aggregate picture masks some small biases in different capital cities. In Sydney, for instance, home valuation differences are significantly positive on average, at around 1.8 per cent. Likewise in Brisbane, there is a significantly positive home valuation difference of 2.8 per cent. But, in Melbourne there is a negative home valuation difference of around 2.2 per cent, on average.

Figure 3: Absolute Home Valuation Differences
Postcode level, 2002 to 2011



Note: Kernel density estimator
Sources: APM; HILDA Release 11.0; authors' calculations

Table 1: Valuation Differences – Postcode Level

	All cities	Sydney	Melbourne	Brisbane
Moments				
Mean (%)	0.5	1.8***	–2.2***	2.8***
95% confidence interval	(–0.3 to 1.3)	(0.6 to 3.1)	(–3.3 to –1.1)	(0.7 to 4.9)
Standard deviation (%)	20.3***	19.7***	17.8***	24.9***
95% confidence interval	(19.4 to 21.3)	(18.6 to 20.8)	(16.5 to 19.0)	(22.1 to 27.7)
Skewness	–0.2	0.3**	–0.1	–0.8**
95% confidence interval	(–0.6 to 0.2)	(0 to 0.5)	(–0.7 to 0.5)	(–1.6 to –0.0)

Notes: Confidence intervals calculated via a bootstrap; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively
Sources: APM; HILDA Release 11.0; authors' calculations

5.2 Significance of Home Valuation Differences

We now examine whether the differences between self-assessed values and sale prices across different postcodes are statistically significant. To determine the statistical significance of the postcode differences it is helpful to nest Equations (1) and (2):

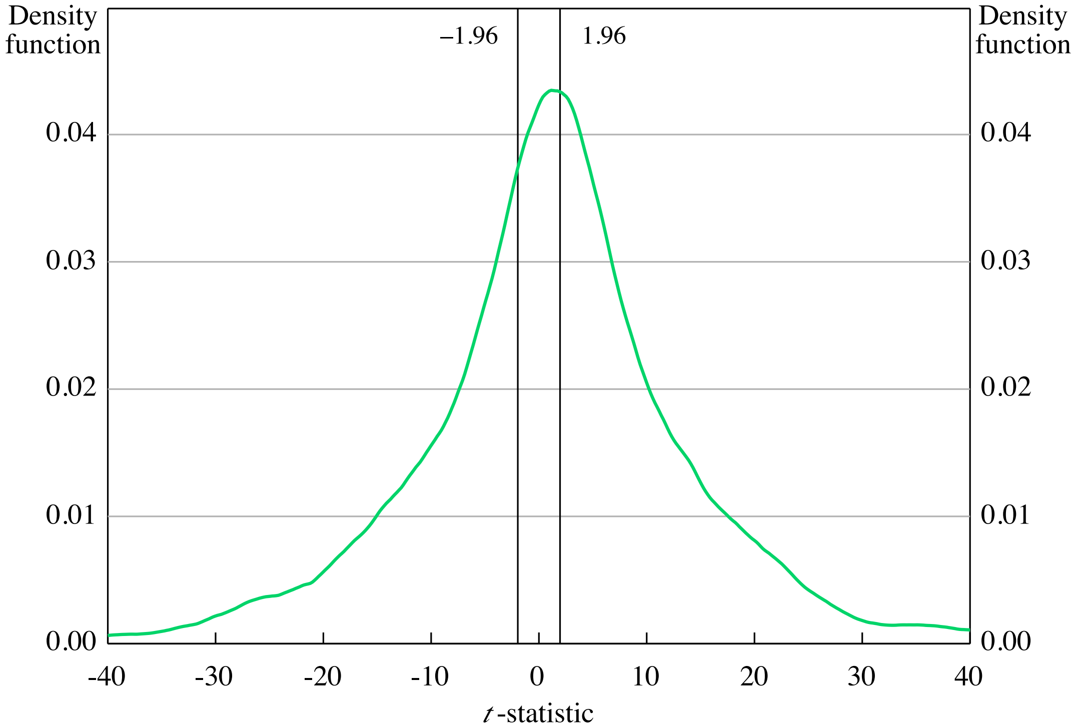
$$ln(P)_{dpt} = \sum_{c=1}^{C} \beta_c^S y_{cdt} + \sum_{c=1}^{C} \beta_c^V y_{cdt} S_{dpt} + \sum_{t=1}^{T} \sum_{p=1}^{P} \lambda_{pt} D_{dpt} + \sum_{t=1}^{T} \sum_{p=1}^{P} \kappa_{pt} D_{dpt} S_{dpt} + \varepsilon_{dpt}, \quad (5)$$

where P_{dpt} is the price of home d in postcode p in time period t (both self-assessed and sold) and where the dummy variable S_{dpt} takes the value of 1 if the observed price is self-assessed and a value of 0 if the observed price is a sale price. All other notation is as defined previously. Equation (5) is estimated separately for each city, with the κ_{pt} coefficients providing an estimate of the home valuation difference for each postcode and time period.

Our main interest is in the significance of the κ_{pt} coefficients, which allow us to make inferences about the mass of statistically significant valuation differences. To this end, in Figure 4 we plot the t-statistics for each home valuation difference κ_{pt} estimated in Equation (5) in each postcode and time period.

We find that about four-fifths of the home valuation differences are statistically significant at conventional levels. This suggests that our finding of no bias on average partly reflects offsetting, significant, differences. In other words, there is a distribution of valuation differences across postcodes; at any point in time, some postcodes have self-assessed valuations that appear high relative to market-inferred prices while some postcodes have valuations that appear low relative to market-inferred prices, on average.

Figure 4: The Significance of Home Valuation Differences



Notes: Kernel density estimator; distributions truncated at 40; t-statistics calculated with robust standard errors clustered at the postcode level

Sources: APM; HILDA Release 11.0; authors' calculations

6. Determinants of Home Valuation Differences

6.1 Panel Regressions

We now examine whether home valuation differences are systematically related to household and regional characteristics. To do so, we estimate several panel regressions of the following form:

$$U_{pt}^{adj} = \alpha + X'_{pt}\beta + \theta_p + \delta_t + \varepsilon_{pt}, \qquad (6)$$

where the dependent variable is the estimate of the average valuation difference (U_{pt}^{adj}) in each postcode and time period. A higher value for the dependent variable indicates greater overvaluation relative to the market sale price for a postcode. The specification includes a set of control variables (X_{pt}) capturing average household characteristics in each postcode and period.

The control variables can be separated into three groups that potentially determine household beliefs. First, there are demographic variables for each postcode, such as the average age of the household head (*Age*) and average household tenure (*Tenure*) (i.e. the number of years households have spent at their current address). Household tenure is a proxy for the level of market information – or 'housing experience' – of each postcode. Second, there are business cycle variables, such as the regional unemployment rate (*Unemployment*), which proxies for the probability of being unemployed. Third, the controls include proxies for the economic resources (or affluence) of households, such as (the log of) average household disposable income (*Log income*), the share of households with an education level above year 12 (*Education*) and the (log of the) compositionally adjusted home sale prices (λ_{pt}^{S}). With the exception of home sale prices, all the control variables are available from the HILDA Survey.

The regression shown by Equation (6) also includes postcode fixed effects (θ_p) that capture unobservable factors that vary across postcodes but not time. The specification also includes year dummies (δ_t) that control for factors that are common to all postcodes at a point in time (for example, the business cycle). The results are shown in Table 2.

Table 2: Explaining Home Valuation Differences across Postcodes

Age	0.009*	0.022***
Age^2	–0.0000	–0.0002***
Tenure	–0.007**	–0.006*
Tenure^2	0.0011*	0.0001
Log income	0.145***	0.032
Unemployment	–0.029***	–0.084***
Education	0.085***	–0.001
Sale price (λ_{pt}^{S})	–0.132***	–0.598***
Time fixed effects	No	Yes
Postcode fixed effects	No	Yes
R^2	0.139	0.721
Within R^2		0.271
Observations	2 551	2 551

Notes: Robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively

Sources: APM; HILDA Release 11.0; authors' calculations

Our preferred estimates are based on the postcode and time fixed effects regression results (column 2 of Table 2). The coefficient on the age and age squared variables suggests that home valuation differences are more positive for postcodes with older homeowners, but this effect becomes smaller as average age increases. At the mean, an increase by one year in the average age of homeowners within a postcode is associated with a higher average home valuation difference of about 0.25 percentage points.

In contrast to age, the coefficient on tenure suggests that postcodes with longer tenure on average are more likely to have negative valuation differences. This effect becomes smaller as average tenure rises. At the mean tenure of 12.5 years, an additional year of tenure, on average within a postcode, results in a predicted valuation difference that is 0.4 percentage points lower.

Higher regional unemployment is also associated with lower valuation differences. An increase in the average rate of unemployment in the region by 1 percentage point results in a decline in the predicted average home valuation difference of 8 percentage points.[9] We interpret this result as consistent with the idea that the degree of overvaluation (undervaluation) becomes smaller (larger) as the probability of becoming unemployed increases.

One reason that such an effect is plausible is that the expected costs of overvaluation – for example, spending too much, taking on more leverage or choosing a more risky financial portfolio (as discussed further below) – are likely to be higher the more likely a household is to experience unemployment. In particular, the most costly possible outcome is an extended period of unemployment that culminates in outright default. If households overvalue their home and have only small financial buffers – such as equity in the home or other assets – then the expected costs of optimism can be very high. The idea that households are less likely to be optimistic, or even are conservative (pessimistic), the more likely they are to experience an unemployment shock is consistent with recent literature on the optimal formation of household beliefs (see, for example, Brunnermeier and Parker (2005)).

9 The regional unemployment rate is measured at the city level. These results are robust to estimates of the unemployment rate at a more disaggregated level (the Australian Statistical Geography Standard, SA4).

Finally, it should be noted that the within-postcode explanatory power of the fixed effects regression is higher than the overall explanatory power of the least squares regression (27 per cent versus 14 per cent). This is due to the inclusion of time fixed effects, capturing the common trend in home valuation differences across all postcodes, due to, for example, the national housing cycle. Abstracting from these effects, the fixed effects model is able to account for around 10 per cent of the variation in home valuation differences within postcodes.

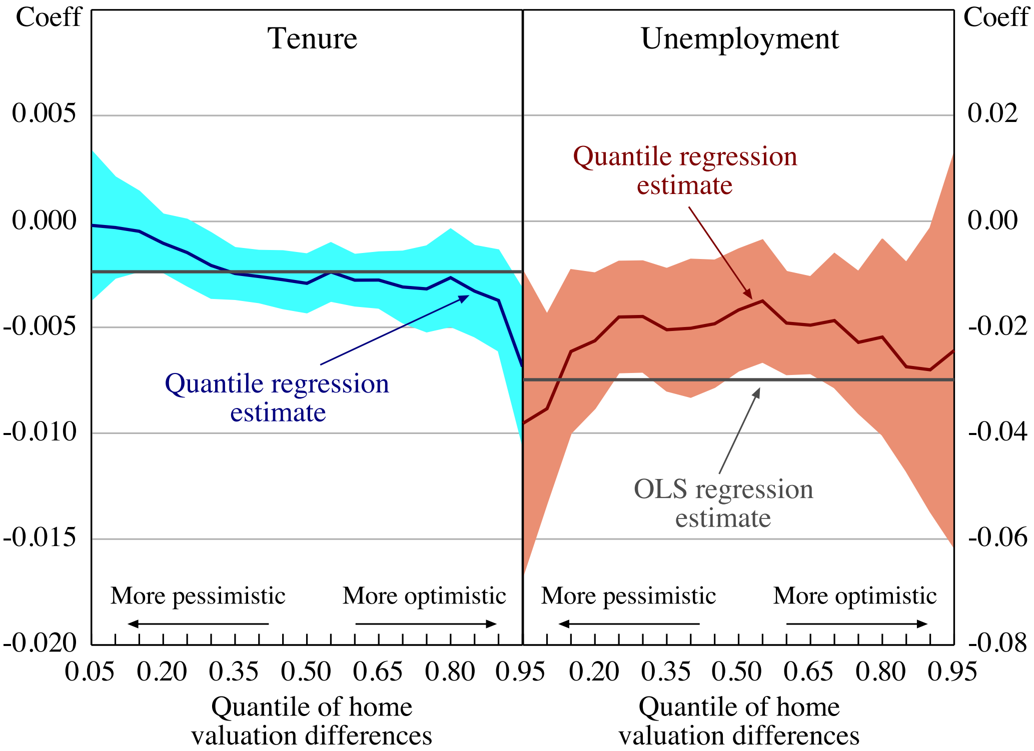
6.2 Quantile Panel Regressions

The previous results show the effect of household characteristics on average home valuation differences assuming these effects are constant across the distribution of beliefs. However, the effect of household characteristics on average home valuation differences may vary across postcodes depending on the extent to which they have a positive or negative valuation difference.

To account for this, a more complete analysis is undertaken by estimating a quantile function, which allows us to examine variation in the parameters over the full distribution of home valuation differences. This is useful because it allows us to see whether certain household characteristics affect either the bias or accuracy of average homeowners' beliefs. More specifically, if the household characteristic solely affects the level of bias, then the sign on the estimated coefficient should stay the same across the full distribution of beliefs. If the household characteristic affects the accuracy of beliefs, then the estimated coefficients should reverse sign between the top and bottom halves of the distribution.

To this end, for two selected characteristics – tenure and the local unemployment rate – Figure 5 plots the quantile regression estimates for U_{pt}^{adj} ranging from the 5th percentile of the home valuation difference distribution (the far left of each panel, where valuation differences are large and negative) to the 95th percentile (the far right of each panel, where valuation differences are large and positive).

Figure 5: The Determinants of Beliefs



Note: Bootstrapped robust standard errors clustered at the postcode level
Sources: APM; HILDA Release 11.0; authors' calculations

For example, focusing on the left panel, at the 5th percentile – a postcode with homeowners that appear very pessimistic on average – the estimated marginal effect of tenure is zero. In contrast, at the 95th percentile – a postcode with homeowners that appear very optimistic on average – the marginal effect of tenure is about -0.7 percentage points. This suggests that greater tenure attenuates optimism for postcodes with homeowners that appear very optimistic. The weak change in coefficient sign suggests that tenure may be capturing housing market information – or experience – within postcodes. In other words, this suggests that tenure affects the accuracy of home valuations rather than bias *per se*.

In contrast, the effect of the regional unemployment rate is estimated to be consistently negative across the distribution of beliefs, and so is about bias rather than accuracy. Interestingly, unemployment matters most (has the largest estimated marginal effect) at both tails. That is, in postcodes that are either very optimistic or very pessimistic, a higher local unemployment rate will reduce average valuation differences by more. For those postcodes who are unbiased or near the 50th percentile, the estimated effects of a higher local unemployment rate are statistically significant and negative, but smaller than the effects at the tails.

7. Financial Decisions and Home Valuation Differences

7.1 Homeowners

We now examine whether home valuation differences are associated with the economic decisions made by homeowners. In other words, in a regression framework, we now treat the valuation differences across postcodes as an independent variable. Specifically, we assess whether valuation differences are correlated with household spending, leverage and portfolio decisions. For instance, if optimistic homeowners typically overestimate the value of their homes we might expect that they spend more than pessimistic homeowners because they believe their lifetime wealth to be relatively high. Moreover, we might expect these optimistic homeowners to hold relatively more housing debt, on average, and to allocate a higher fraction of their financial portfolios to risky assets.[10]

To demonstrate the effect of valuation differences on household decisions, we first estimate postcode-level regressions of the following form:

$$Y_{pt} = \alpha + \widehat{\lambda}_{pt}^{V}\beta + X_{pt}^{\prime}\gamma + \delta_t + \varepsilon_{pt}, \tag{7}$$

where Y_{pt} is a measure of financial decisions for the average homeowner in postcode p in period t. The measures of household decisions include a spending measure – the (log) level of household consumption expenditure ($SPENDING_{pt}$) – measures of household leverage – the (log) level of total debt ($DEBT_{pt}$) and the (log) level of total housing debt ($HDEBT_{pt}$) – as well as portfolio allocation measures – the average share of wealth held in financial assets ($FINSHARE_{pt}$) and the average share of wealth held in equities ($EQSHARE_{pt}$). The key explanatory variable is the average self-assessed home value ($\widehat{\lambda}_{pt}^{V}$) in each postcode and year.

10 This is consistent with the findings of Brunnermeier, Gollier and Parker (2007).

We also estimate similar regressions in which we decompose the average self-assessed home value into the average home valuation difference (U_{pt}^{adj}) and the average home sale price ($\widehat{\lambda}_{pt}^{S}$):

$$Y_{pt} = \alpha + \underbrace{(\widehat{\lambda}_{pt}^{V} - \widehat{\lambda}_{pt}^{S})}_{U_{pt}^{adj}}\beta_1 + \widehat{\lambda}_{pt}^{S}\beta_2 + X_{pt}'\gamma + \delta_t + \varepsilon_{pt}. \tag{8}$$

Equation (8) is a more general specification than Equation (7) because we allow the estimated coefficients on the average valuation difference (β_1) and the average home sale price (β_2) to potentially differ from each other.

The key explanatory variable is the average home valuation difference made by homeowners in each postcode and year ($U_{pt}^{adj} = \widehat{\lambda}_{pt}^{V} - \widehat{\lambda}_{pt}^{S}$). The specification includes a set of control variables for the average household in each postcode and year (X_{pt}), which is similar to that used in Equation (6). The estimates of Equations (7) and (8) are shown in Table 3 for each dependent variable. For brevity, we do not report the coefficient estimates on the control variables. A more detailed analysis of the regression results, including a discussion of the estimated effect of the control variables, is available in Appendix D.

For postcode-level spending, the positive coefficient on average self-assessed home values ($\widehat{\lambda}_{pt}^{V}$) (column 1, block 1) indicates that household spending and self-assessed home values are positively correlated. A 1 per cent increase in the level of housing prices is associated with a 0.31 per cent increase in household spending. This result implies a marginal propensity to consume of around five cents for a one dollar increase in the self-assessed value (using an average spending-to-housing wealth ratio of 16 per cent), which is broadly in line with the magnitude of 'housing wealth effects' documented in previous studies for Australia (see, for example, Dvornak and Kohler (2007) and Windsor *et al* (2013)).

Decomposing the average self-assessed home value into the average home valuation difference (U_{pt}^{adj}) and the average home sale price ($\widehat{\lambda}_{pt}^{S}$) (column 2), we find that the positive wealth effect reflects two factors: first, that higher market values are associated with more spending and, second, that greater overvaluation is also associated with more spending. A 1 percentage point increase in the valuation difference is associated with a 0.34 per cent increase in household spending, on average. To gauge the economic significance of this result, the regression estimates

Table 3: Homeowner Decisions and Home Valuation Differences

$SPENDING_{pt}$		
Self-assessed home values	0.308***	
Sale prices		0.304***
Valuation difference		0.338***
$DEBT_{pt}$		
Self-assessed home values	0.383***	
Sale prices		0.298**
Valuation difference		0.817**
$HDEBT_{pt}$		
Self-assessed home values	0.455***	
Sale prices		0.466***
Valuation difference		0.384***
$FINSHARE_{pt}$		
Self-assessed home values	0.012	
Sale prices		–0.017
Valuation difference		0.159***
$EQSHARE_{pt}$		
Self-assessed home values	0.037***	
Sale prices		0.031***
Valuation difference		0.072***
Time fixed effects	Yes	Yes

Notes: Bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; the dependent variables $SPENDING_{pt}$, $DEBT_{pt}$ and $HDEBT_{pt}$ are in log levels; the dependent variables $FINSHARE_{pt}$ and $EQSHARE_{pt}$ are measured as ratios; the $SPENDING_{pt}$ regression is estimated over the period 2006 to 2011, for which there was comprehensive expenditure data; the $DEBT_{pt}$, $FINSHARE_{pt}$ and $EQSHARE_{pt}$ regressions are estimated on the wealth module years of 2002, 2006 and 2010; the $HDEBT_{pt}$ regression is estimated over the period 2002 to 2011

Sources: APM; HILDA Release 11.0; authors' calculations

imply that a 1 percentage point increase in the valuation difference would generate about the same effect on spending as a 2 per cent increase in household income, so the effect seems reasonably large.

We also find that postcodes in which the average homeowner appears to overestimate the value of their homes are those that typically have more debt (column 2, blocks 2 and 3). This is true for total debt, ($DEBT_{pt}$, block 2) and, more specifically, housing debt ($HDEBT_{pt}$, block 3).[11] For example, the estimates

11 The results also hold for other measures of leverage, including the ratio of total debt to total assets, the ratio of housing debt to housing assets and the ratio of housing debt to household income. These unreported results are available upon request.

indicate that a 1 percentage point increase in the valuation difference is associated with a 0.8 per cent rise in debt, on average.

We find that postcodes in which households generally overestimate the value of their homes are those in which households allocate a relatively high share of their wealth to financial assets, such as equities (column 2, blocks 4 and 5). The estimates indicate that a 1 percentage point increase in home valuation differences is associated with a higher average share of wealth held in financial assets of about 16 basis points. In terms of economic significance, this is roughly the same effect as 2 per cent more household income. We also find that higher average housing prices do not significantly affect the share of wealth held in financial assets (block 4), but do affect (increase) the share of wealth held in equities (block 5).

In sum, we find that, at the postcode level, valuation differences are positively correlated with the level of spending, debt and the share of wealth held in financial assets. In other words, postcodes that appear to overvalue their homes typically spend more, have higher leverage and choose riskier portfolios than postcodes that do not. Since this effect exists after controlling for the average level of prices, it suggests that these valuation differences may be capturing underlying optimism or pessimism.

7.2 Renters

To check whether home valuation differences are likely to be capturing optimism or pessimism among homeowners, we estimate the same postcode-level regressions on the decision variables of renting households. If valuation differences are capturing true differences between homeowner beliefs and sale prices then the valuation differences should only affect the decisions of homeowners and not renters. If, instead, we find that renters' decisions are also affected by our measure of valuation differences then this could be evidence that the measures are capturing unobserved heterogeneity in the characteristics of properties sold or valued in each postcode.

To do this, Equations (7) and (8) are re-estimated for each relevant decision variable, constructed using only data obtained from renting households. Specifically, we consider the average level of renters' spending, non-housing debt (credit cards and personal loans) and the average share of wealth held in equities (Table 4).

Table 4: Renter Decisions and Home Valuation Differences

$SPENDING_{pt}$		
Self-assessed home values	0.103***	
Sale prices		0.183***
Valuation difference		–0.115*
$DEBT_{pt}$		
Self-assessed home values	0.220	
Sale prices		0.202
Valuation difference		0.286
$EQSHARE_{pt}$		
Self-assessed home values	0.046***	
Sale prices		0.058***
Valuation difference		0.002
Time fixed effects	Yes	Yes

Notes: Bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; the dependent variables $SPENDING_{pt}$ and $DEBT_{pt}$ are in log levels; the dependent variable $EQSHARE_{pt}$ is measured as a ratio; the $SPENDING_{pt}$ regression is estimated over the period 2006 to 2011, for which there was comprehensive expenditure data; the $DEBT_{pt}$ and $EQSHARE_{pt}$ regressions are estimated on the wealth module years of 2002, 2006 and 2010

Sources: APM; HILDA Release 11.0; authors' calculations

At the postcode level, the spending of renters appears to be positively associated with market-inferred prices (column 2). This could reflect the fact that housing prices and spending are commonly associated with a third factor, such as income expectations (this is consistent with the findings in Attanasio *et al* (2009), for example). However, the postcode-level spending of renters is *negatively* correlated with homeowner valuation differences (column 2). Moreover, there is no significant relationship between renters' debt levels and either market-inferred housing prices or valuation differences. Finally, we find that the share of renters' total assets held in equities is responsive to market-inferred housing prices, which could be due to equity holdings and housing prices being commonly associated with a third factor. But, unlike homeowners' portfolio allocations, valuation differences do not explain the share of renters' wealth held in equities.

Taken together, these results are not consistent with the alternative explanation that our estimates of the homeowner valuations differences are contaminated by unobservable factors that affect the beliefs and decision-making of *all households*. Instead, our results support the notion that our estimated home valuation differences are capturing only the sentiments of homeowners, insofar as these valuation differences affect the decision-making of homeowners and not renters.

8. Robustness Tests

A condition for consistent estimation of the home valuation differences in our benchmark model is that any omitted housing characteristics in the hedonic models (for example, the number of bathrooms) should not have differential effects on market prices as compared with self-assessed values. To determine the extent to which our main results are robust to such omitted variables we construct the valuation differences using an alternative repeat-sales methodology. This is a useful alternative as time-invariant characteristics are differenced out of the regression model, reducing the potential for omitted characteristics to induce bias.

Specifically, we estimate the following model for sales prices:

$$ln(S)_{ipt} - ln(S)_{iph} = \sum_{\tau=1}^{T}\sum_{p=1}^{P}\tilde{\lambda}^{s}_{p\tau}G_{ip\tau} + e^{s}_{ipt}, \tag{9}$$

where $h < t$ and $G_{ip\tau}$ is a dummy equal to 1 for home i in postcode p in the period that resale occurs (i.e. if $\tau = t$); –1 for home i in postcode p in the period the previous sale occurred (i.e. if $\tau = h$); and 0 otherwise. Similarly, the model for self-assessed home values is:

$$ln(V)_{jpt} - ln(V)_{jpt-1} = \sum_{t=1}^{T}\sum_{p=1}^{P}\tilde{\lambda}^{v}_{pt}D_{jpt} + e^{v}_{jpt}, \tag{10}$$

where D_{jpt} is a dummy equal to 1 for self-assessed home j in postcode-time pt.

Comparing estimates of $\tilde{\lambda}^{s}_{pt}$ with $\tilde{\lambda}^{v}_{pt}$ provides home valuation differences on the *growth* in home values, which we refer to as Θ_{pt}.[12]

12 The drawback with using the repeat-sales methodology is that we cannot estimate the valuation differences in levels, but only in growth rates.

The determinants of the valuation differences can then be re-estimated accordingly,

$$\Theta_{pt} = \Delta X'_{pt}\beta + \Delta\delta_t + \Delta\varepsilon_{pt}, \tag{11}$$

with the results provided in Table 5 and compared with a first-differenced version of Equation (6).[13]

Table 5: Explaining Home Valuation Differences across Postcodes

	Least squares		**Feasible generalised least squares**	
	Hedonic	Repeat-sales	Hedonic	Repeat-sales
Δ Age	0.022***	0.007*	0.019***	0.007**
Δ Age^2	–0.0002***	–0.0001	–0.0002***	–0.0001**
Δ Tenure	–0.007*	–0.010***	–0.005*	–0.010***
Δ $Tenure^2$	0.0001	0.0001***	0.0001**	0.0001***
Δ Log income	0.012	–0.016	0.004	–0.005
Δ Unemployment	–0.021***	–0.003	–0.025***	–0.005
Δ Education	–0.002	0.006	0.016	0.004
R^2	0.030	0.016		
Observations	2 122	1 948	2 095	1 919

Notes: Where necessary, robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively

Sources: APM; HILDA Release 11.0; authors' calculations

Across both models, we find similar results to those presented in Section 6. Namely, the change in home valuation differences is positively associated with the change in age and negatively associated with the change in tenure and the regional unemployment rate, although the unemployment rate effect is now insignificant. We do not observe any statistically significant association between the home valuation differences and income or education.

Our results regarding financial decisions and home valuation differences can be checked in a similar way. Specifically, Equations (7) and (8) can be re-estimated in first differences. The results are provided in Table 6.[14]

13 There is serial correlation in the error term ($\Delta\varepsilon_{pt}$). Accordingly, in columns 1 and 2, robust standard errors are clustered at the postcode level, which accounts for this. We also report feasible generalised least squares that produce asymptotically more efficient (though possibly biased) estimates (Woolridge 2002) in columns 3 and 4.

14 We do not report results for $DEBT_{pt}$, $FINSHARE_{pt}$ and $EQSHARE_{pt}$ because these dependent variables are only available in HILDA wealth module years 2002, 2006 and 2010. As such, first differencing only yields two time periods to examine: 2006 and 2010.

Table 6: Homeowner Decisions and Home Valuation Differences

	Least squares		Feasible generalised least squares	
$\Delta SPENDING_{pt}$				
Δ Self-assessed home values	0.177		0.169***	
Δ Sale prices		0.304**		0.235*
Δ Valuation difference (ppt)		0.165		0.162***
$\Delta HDEBT_{pt}$				
Δ Self-assessed home values	0.341**		0.338**	
Δ Sale prices		0.645***		0.521**
Δ Valuation difference (ppt)		0.271*		0.283**
Δ Time fixed effects	Yes	Yes	Yes	Yes

Notes: Where necessary, bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; the dependent variables $SPENDING_{pt}$ and $HDEBT_{pt}$ are in log levels; the $SPENDING_{pt}$ regression is estimated over the period 2006 to 2011, for which there was comprehensive expenditure data; the $HDEBT_{pt}$ regression is estimated over the period 2002 to 2011

Sources: APM; HILDA Release 11.0; authors' calculations

For postcode-level spending, the positive coefficient on self-assessed home values (column 3, block 1) again indicates that household spending and self-assessed home values are positively correlated. And, again, we find that the positive wealth effect reflects two factors: higher market values are associated with more spending and greater overvaluation is associated with more spending. But, this effect is smaller than the comparable results presented in Section 7. Likewise, for the level of housing debt, we find similar results to those presented earlier; the change in home valuation differences is positively associated with the change in housing debt (column 4, block 2).

9. Conclusion

In contrast to the existing literature, this paper provides an approach to measuring home valuation differences that is free of recollection bias. We also study the determinants of valuation differences and whether valuation differences are important for economic decisions.

We find that homeowners' housing price beliefs – measured at the postcode level – are generally unbiased. However, there is significant variation around this unbiased mean: while around half of all postcodes provide an average valuation within 11 per cent of the average market value, about one-quarter of postcodes provide valuations that are more than 20 per cent away from the average market value.

Household characterisics including age and tenure, and local area characteristics, such as the regional unemployment rate, are able to explain some of the variation between homeowner beliefs and home prices. The results suggest that the unemployment rate affects the bias of beliefs, while tenure affects the accuracy of beliefs.

The extent of overvaluation is positively correlated with household spending and leverage. Homeowners that appear to overvalue their homes typically spend more and are more leveraged than owners who appear unbiased. In contrast, homeowners that appear to undervalue their homes spend less and are less leveraged relative to the same reference group. Our findings suggest that beliefs about home values affect household financial decisions.

Appendix A: Hedonic Model Output

Table A1: Hedonic Model Output

	Sydney	Melbourne	Brisbane
Sale prices: 1992–2012			
1 bed	–0.664***	–0.714***	–0.602***
2 beds	–0.270***	–0.265***	–0.232***
4 beds	0.242***	0.231***	0.249***
5 beds	0.427***	0.410***	0.452***
6 beds	0.516***	0.491***	0.556***
7 beds	0.576***	0.417***	0.646***
Unit	–0.311***	–0.199***	–0.217***
Constant: 3-bedroom house[(a)]	13.23***	12.86***	12.86***
R^2	0.795	0.802	0.666
Observations	609 499	683 323	222 067
Self-assessed home values: 2002–2011			
1 bed	–0.476***	–0.521***	0.119
2 beds	–0.234***	–0.154***	–0.210***
4 beds	0.181***	0.162***	0.195***
5 beds	0.291***	0.337***	0.258***
6 beds	0.490***	0.383***	0.323**
7 beds	0.511***	0.520*	0.376***
Unit	–0.342***	–0.257***	–0.223*
Constant: 3-bedroom house[(a)]	13.21***	12.94***	12.81***
R^2	0.767	0.789	0.745
Observations	7 171	7 971	3 821

Notes: Postcode-time dummies omitted; robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively
(a) For a particular postcode and time period

Sources: APM; HILDA Release 11.0; authors' calculations

Appendix B: Estimating Homeowner Recollection Bias

The advantage of our data is that we can estimate any valuation bias, abstracting from recollection bias. In this appendix, we estimate the degree of home price recollection bias. Home price recollection bias occurs when surveyed homeowners incorrectly recall the purchase price of their homes. This could happen if, for example, the owner simply has imperfect recall because they bought the home a long time ago. In the HILDA Survey, homeowners are asked to recall the purchase price of their home every four years (specifically in the 2002, 2006 and 2010 surveys). About 13 per cent of homeowners report a different purchase price from one interview to the next despite apparently not moving home between interviews. This suggests that recollection bias might be a problem.

A nice feature of our dataset is that we can directly estimate such a bias. We proxy recollection bias as the difference between what the homeowner recalls for their purchase price and the purchase price inferred from the APM dataset using a hedonic regression model. To the best of our knowledge, we are the first to directly estimate recollection bias in home prices.

We estimate the following hedonic regression model:

$$ln(S)_{i^*pt} = \alpha * PURCHASE_{i^*pt} + \sum_{c=1}^{C} \beta_c y_{ci^*p} + \sum_{t=1}^{T} \sum_{p=1}^{P} \lambda_{pt}^{S} D_{i^*pt} + \varepsilon_{i^*pt}, \quad \text{(B1)}$$

where S_{i^*pt} represents the sale price of home i^* in postcode p in year t (where i^* includes sales prices from the APM transactions dataset and recalled sale prices from the HILDA Survey).[15] The key explanatory variable is a dummy variable ($PURCHASE_{i^*pt}$) which is equal to 1 if the sale price is a reported purchase price from the HILDA Survey and is equal to 0 if the sale price is from the APM dataset.

The specification includes the same set of hedonic controls as before (for example, bedrooms and the type of housing) as well as postcode-time fixed effects. The specification effectively 'stacks' all the sales prices from APM with the reported purchase prices from the HILDA Survey. Our main interest is in the coefficient on the intercept term $PURCHASE_{i^*pt}$, which captures the average difference

15 This model specification uses annual data rather than quarterly data because the HILDA respondents only report the year of purchase and not the exact date.

between sale prices and reported purchase prices after controlling for the location and characteristics of the property at the time of purchase. If the coefficient on this term is positive it indicates that surveyed homeowners overstate the purchase price of their homes, on average. If the coefficient is negative, surveyed homeowners understate their home purchase prices, on average. The intercept therefore provides an estimate of the average recollection bias.

The results of estimating Equation (B1) suggest that homeowners understate the purchase price of their homes by 3.4 per cent on average (Table B1). This negative bias is statistically significant. Melser (2013) finds an average positive valuation difference of about 4.5 per cent by comparing current estimates of home prices to initial purchase prices in the HILDA Survey. But if homeowners undervalue the initial purchase price by 3.4 per cent, then this suggests that Mesler's estimated valuation difference would mainly reflect recollection bias. Taking the Melser estimates at face value, and adjusting for this recollection bias, we would find an average positive valuation difference of about 1 per cent. This is very similar to our own estimate of the average valuation difference across all postcodes.

Table B1: Homeowner Recollection Bias

Purchase	–0.034**
1 bed	–0.690***
2 beds	–0.268***
4 beds	0.239***
6 beds	0.520***
7 beds	0.549***
Unit	–0.249***
Constant: 3-bedroom house[(a)]	13.01***
R^2	0.793
Observations	1 519 719

Notes: Postcode-time dummies omitted; robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively
(a) For a particular postcode and time period

Sources: APM; HILDA Release 11.0; authors' calculations

Appendix C: Quantile Regression Output

Table C1: Explaining Home Valuation Differences across Postcodes

Appears pessimistic – 25th percentile	
Age	0.004
Age^2	–0.000
Tenure	–0.002**
Log income	0.183***
Unemployment	–0.018***
Education	0.060***
Sale price (λ_{pt}^{S})	–0.154***
Constant	–0.244
50th percentile	
Age	0.005*
Age^2	–0.000
Tenure	–0.003***
Log income	0.131***
Unemployment	–0.016***
Education	0.069***
Sale price (λ_{pt}^{S})	–0.117***
Constant	–0.080
Appears optimistic – 75th percentile	
Age	0.007*
Age^2	–0.000
Tenure	–0.003***
Log income	0.105***
Unemployment	–0.022***
Education	0.092***
Sale price (λ_{pt}^{S})	–0.119***
Constant	0.293
Time fixed effects	No
Postcode fixed effects	No
Observations	2 551

Notes: Bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively

Sources: APM; HILDA Release 11.0; authors' calculations

Appendix D: Economic Decisions Output

The regression output from Equation (8) is presented in Table D1.

Table D1: Homeowners' Decisions and Home Valuation Differences

	$SPENDING_{pt}$	$DEBT_{pt}$	$HDEBT_{pt}$	$FINSHARE_{pt}$	$EQSHARE_{pt}$
Sale price (λ_{pt}^{S})	0.304***	0.298**	0.466***	–0.017	0.031***
Valuation difference	0.338***	0.817**	0.384***	0.159***	0.072***
Adults	0.101***	–0.110	–0.071	–0.010	–0.007
Children	0.105***	0.320***	0.116***	–0.028**	–0.008
Age	0.034***	–0.026	–0.019	0.023***	0.005**
Age^2	–0.0004***	0.0001	0.0001	–0.0002***	–0.0000
Tenure	–0.000	–0.082***	–0.048***	0.006***	0.002
Tenure^2	–0.0000	0.0014***	0.0005***	–0.0001**	–0.0000
Log income	0.174***	0.167	0.077*	0.072***	0.024***
Unemployment	–0.017	–0.021	0.048	0.040***	0.023***
Education	0.073	0.101	–0.083	0.066***	0.015
Time fixed effects	Yes	Yes	Yes	Yes	Yes
R^2	0.499	0.243	0.293	0.320	0.209
Observations	1 708	731	2 423	747	747

Notes: Bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; the dependent variables $SPENDING_{pt}$, $DEBT_{pt}$ and $HDEBT_{pt}$ are in log levels; the dependent variables $FINSHARE_{pt}$ and $EQSHARE_{pt}$ are measured as ratios; the $SPENDING_{pt}$ regression is estimated over the period 2006 to 2011, for which there was comprehensive expenditure data; the $DEBT_{pt}$, $FINSHARE_{pt}$ and $EQSHARE_{pt}$ regressions are estimated on the wealth module years of 2002, 2006 and 2010; the $HDEBT_{pt}$ regression is estimated over the period 2002 to 2011

Sources: APM; HILDA Release 11.0; authors' calculations

The coefficients on the control variables for the spending regression are unsurprising. The sign and significance of the age variables – capturing the average age of homeowners within postcodes – as well as the variables for the average number of children and adults per family within postcodes, suggest that these variables capture the life cycle of spending. Furthermore, the coefficients on education and disposable income levels within postcodes show that postcode-level spending is associated with our proxies for lifetime income.

Looking at the estimated coefficients on the controls for the other dependent variables, we find that, on average, postcodes with older homeowners hold more wealth in financial assets; higher-income postcodes hold more housing debt and more wealth in financial assets while more educated postcodes also hold more wealth in financial assets.

Appendix E: Additional Robustness

E.1 Data Restrictions

The exclusion restrictions applied in this paper are detailed in Table E1. In this appendix, we examine the sensitivity of our main results to additional exclusion restrictions on sample sizes and outliers.

Table E1: Sample Selection

	Dropped	Remaining
Criteria for selection: sale prices – 1992–2012		
Recorded private final sale price		3 877 815
\$28 000 $\leq$ sale price $\leq$ \$50 000 000	6 812	3 871 003
Non-reported bedrooms in dataset	2 352 132	1 518 871
$\leq$ 7 bedrooms	1 117	1 517 754
$\geq$ 2 sales per postcode-quarter	2 865	1 514 889
Criteria for selection: self-assessed home values – 2002–2011		
Self-assessed home valuation		47 820
Respondent located in Sydney, Melbourne or Brisbane	27 679	20 141
\$28 000 $\leq$ home valuation $\leq$ \$50 000 000	3	20 138
Non-reported bedrooms in dataset	14	20 124
$\leq$ 7 bedrooms	18	20 106
$\geq$ 2 assessments per postcode-year	1 142	18 964

Sources: APM; HILDA Release 11.0; authors' calculations

In this robustness test, the number of sales per postcode per quarter is further restricted to at least 20 sales and the number of self-assessed home valuations per postcode per year is further restricted to be at least 6 valuations. We also trim the top and bottom 1 per cent of home valuation differences.

The original results from Table 2 are shown again in Table E2 based on the additional exclusion restrictions. The local unemployment rate result is robust to the additional restrictions. So too are the results regarding the effect of tenure on home valuation differences. While the sign on the age variable is consistent with the results presented in Table 2, the effect is now statistically insignificant.

Table E2: Explaining Home Valuation Differences across Postcodes

Age	0.007	0.013
Age^2	–0.0000	–0.0001
Tenure	–0.013**	–0.008*
Tenure^2	0.0002**	0.0002**
Log income	0.151***	0.023
Unemployment	–0.023***	–0.056***
Education	0.110***	0.070*
Sale price λ_{pt}^{S}	–0.190***	–0.502***
Time fixed effects	No	Yes
Postcode fixed effects	No	Yes
R^2	0.195	0.800
Observations	1 323	1 323

Notes: Robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively

Sources: APM; HILDA Release 11.0; authors' calculations

The original results from Table 3 are shown again in Table E3 based on the additional exclusion restrictions. We find that home valuation differences have very similar effects on all the dimensions of household decisions considered in the main text.

Table E3: Homeowner Decisions and Home Valuation Differences

$SPENDING_{pt}$		
Self-assessed home prices	0.221***	
Sale prices		0.225***
Valuation difference		0.174**
$DEBT_{pt}$		
Self-assessed home prices	0.231	
Sale prices		0.203
Valuation difference		0.486*
$HDEBT_{pt}$		
Self-assessed home prices	0.455***	
Sale prices		0.450***
Valuation difference		0.519***
$FINSHARE_{pt}$		
Self-assessed home prices	0.002	
Sale prices		–0.016
Valuation difference		0.164***
$EQSHARE_{pt}$		
Self-assessed home prices	0.040**	
Sale prices		0.035**
Valuation difference		0.092***
Time fixed effects	Yes	Yes

Notes: Bootstrapped robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; the dependent variables $SPENDING_{pt}$, $DEBT_{pt}$ and $HDEBT_{pt}$ are in log levels; the dependent variables $FINSHARE_{pt}$ and $EQSHARE_{pt}$ are measured as ratios; the $SPENDING_{pt}$ regression is estimated over the period 2006 to 2011, for which there was comprehensive expenditure data; the $DEBT_{pt}$, $FINSHARE_{pt}$ and $EQSHARE_{pt}$ regressions are estimated on the wealth module years of 2002, 2006 and 2010; the $HDEBT_{pt}$ regression is estimated over the period 2002 to 2011

Sources: APM; HILDA Release 11.0; authors' calculations

E.2 Weighted Least Squares

A further robustness check, which accounts for estimation uncertainty without restricting the sample, is to weight each estimated home valuation difference by the uncertainty around its estimate.

For example, to examine the robustness of our results regarding the determinants of home valuation differences, Equation (6) can be re-estimated after pre-multiplying both the left- and right-hand side variables by the inverse of the standard errors of the estimated home valuation differences (as estimated in Equation (5)).

The results from this weighted least squares (WLS) approach are shown in Table E4. The results are similar to those presented in Table 2. Reiterating, home valuation differences are positively associated with age; negatively associated with tenure (although this effect is now insignificant in the regression with fixed effects); and negatively associated with the regional unemployment rate.[16]

Table E4: Explaining Home Valuation Differences across Postcodes – Weighted Least Squares Estimates

Age	0.0136**	0.0232***
Age^2	–0.0001	–0.0002***
Tenure	–0.0091**	–0.0043
Tenure^2	0.0001	0.0001
Log income	0.0997***	0.0440*
Unemployment	–0.0245***	–0.0098*
Education	0.0634	–0.0129
Sale price (λ_{pt}^{S})	–0.107***	–0.0710***
Time fixed effects	No	Yes
Postcode fixed effects	No	Yes
R^2	0.104	0.736
Observations	2 376	2 376

Notes: Time and postcode dummies omitted from fixed effects column; robust standard errors clustered at the postcode level; ***, ** and * indicate significance at the 1, 5 and 10 per cent level, respectively; this table shows the results from re-estimating Equation (6) in the main text after pre-multiplying each variable by $\frac{1}{s.e.(\kappa_{pt})}$, where κ_{pt} denotes the estimated home valuation difference obtained from Equation (5) in the main text

Sources: APM; HILDA Release 11.0; authors' calculations

16 The results regarding home valuation differences and household decision-making were also robust to a WLS approach, and are available upon request.

References

Agarwal S (2007), 'The Impact of Homeowners' Housing Wealth Misestimation on Consumption and Saving Decisions', *Real Estate Economics*, 35(2), pp 135–154.

Attanasio OP, L Blow, R Hamilton and A Leicester (2009), 'Booms and Busts: Consumption, House Prices and Expectations', *Economica*, 76(301), pp 20–50.

Berger-Thomson L, E Chung and R McKibbin (2009), 'Estimating Marginal Propensities to Consume in Australia Using Micro Data', RBA Research Discussion Paper No 2009-07.

Bidder RM and ME Smith (2012), 'Robust Animal Spirits', *Journal of Monetary Economics*, 59(8), pp 738–750.

Brunnermeier MK, C Gollier and JA Parker (2007), 'Optimal Beliefs, Asset Prices, and the Preference for Skewed Returns', *The American Economic Review*, 97(2), pp 159–165.

Brunnermeier MK and JA Parker (2005), 'Optimal Expectations', *The American Economic Review*, 95(4), pp 1092–1118.

***The Castle* (1997),** Motion picture, Working Dog Productions, Melbourne. Distributed by Village Roadshow (Australia), and starring Michael Caton, Anne Tenney and Stephen Curry.

Dvornak N and M Kohler (2007), 'Housing Wealth, Stock Market Wealth and Consumption: A Panel Analysis for Australia', *The Economic Record*, 83(261), pp 117–130.

Ellis L, J Lawson and L Roberts-Thomson (2003), 'Housing Leverage in Australia', RBA Research Discussion Paper No 2003-09.

Epstein LG and M Schneider (2008), 'Ambiguity, Information Quality, and Asset Pricing', *The Journal of Finance*, 63(1), pp 197–228.

Genesove D and C Mayer (2001), 'Loss Aversion and Seller Behavior: Evidence from the Housing Market', *The Quarterly Journal of Economics*, 116(4), pp 1233–1260.

Goodman, Jr JL and JB Ittner (1992), 'The Accuracy of Home Owners' Estimates of House Value', *Journal of Housing Economics*, 2(4), pp 339–357.

Hansen J (2009), 'Australian House Prices: A Comparison of Hedonic and Repeat-Sales Measures', *Economic Record*, 85(269), pp 132–145.

Hansen LP and TJ Sargent (2008), *Robustness*, Princeton University Press, Princeton.

Henriques AM (2013), 'Are Homeowners in Denial about their House Values? Comparing Owner Perceptions with Transaction-Based Indexes', Board of Governors of the Federal Reserve System Finance and Economics Discussion Series No 2013-79.

Ihlanfeldt KR and J Martinez-Vazquez (1986), 'Alternative Value Estimates of Owner-Occupied Housing: Evidence on Sample Selection Bias and Systematic Errors', *Journal of Urban Economics*, 20(3), pp 356–369.

Kain JF and JM Quigley (1972), 'Note on Owner's Estimate of Housing Value', *Journal of the American Statistical Association*, 67(340), pp 803–806.

Kiel KA and JE Zabel (1999), 'The Accuracy of Owner-Provided House Values: The 1978–1991 American Housing Survey', *Real Estate Economics*, 27(2), pp 263–298.

Kish L and JB Lansing (1954), 'Response Errors in Estimating the Value of Homes', *Journal of the American Statistical Association*, 49(267), pp 520–538.

Kohler M and K Smith (2005), 'Housing and the Household Wealth Portfolio: The Role of Location', RBA Research Discussion Paper No 2005-10.

Melser D (2013), 'How Well Do Australian Home-Owners Know the Value of Their Home?', *The Australian Economic Review*, 46(1), pp 31–44.

Robins PK and RW West (1977), 'Measurement Errors in the Estimation of Home Value', *Journal of the American Statistical Association*, 72(358), pp 290–294.

Van den Steen E (2004), 'Rational Overoptimism (and Other Biases)', *The American Economic Review*, 94(4), pp 1141–1151.

Windsor C, J Jääskelä and R Finlay (2013), 'Home Prices and Household Spending', RBA Research Discussion Paper No 2013-04.

Wooldridge JM (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge.

Copyright and Disclaimer Notices

APM Disclaimer

The Australian property price data used in this publication are sourced from Australian Property Monitors Pty Limited ACN 061 438 006 of level 5, 1 Darling Island Road Pyrmont NSW 2009 (P: 1 800 817 616).

In providing these data, Australian Property Monitors relies upon information supplied by a number of external sources (including the governmental authorities referred to below). These data are supplied on the basis that while Australian Property Monitors believes all the information provided will be correct at the time of publication, it does not warrant its accuracy or completeness and to the full extent allowed by law excludes liability in contract, tort or otherwise, for any loss or damage sustained by you, or by any other person or body corporate arising from or in connection with the supply or use of the whole or any part of the information in this publication through any cause whatsoever and limits any liability it may have to the amount paid to the Publisher for the supply of such information.

New South Wales Land and Property Information

Contains property sales information provided under licence from the Department of Finance and Services, Land and Property Information.

State of Victoria

The State of Victoria owns the copyright in the Property Sales Data and reproduction of that data in any way without the consent of the State of Victoria will constitute a breach of the Copyright Act 1968 (Cth). The State of Victoria does not warrant the accuracy or completeness of the Property Sales Data and any person using or relying upon such information does so on the basis that the State of Victoria accepts no responsibility or liability whatsoever for any errors, faults, defects or omissions in the information supplied.

State of Queensland

© State of Queensland (Department of Natural Resources and Mines) 2012. In consideration of the State permitting use of this data you acknowledge and agree that the State gives no warranty in relation to the data (including accuracy, reliability, completeness, currency or suitability) and accepts no liability (including without limitation, liability in negligence) for any loss, damage or costs (including consequential damage) relating to any use of the data. Data must not be used for direct marketing or be used in breach of the privacy laws.

HILDA

The following Disclaimer applies to data obtained from the HILDA Survey and reported in this RDP.

Disclaimer

The Household, Income and Labour Dynamics in Australia (HILDA) Survey was initiated and is funded by the Australian Government Department of Social Services (DSS), and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute). The findings and views based on these data should not be attributed to either DSS or the Melbourne Institute.

RESEARCH DISCUSSION PAPERS

These papers can be downloaded from the Bank's website or a hard copy may be obtained by writing to:

Mail Room Supervisor
Information Department
Reserve Bank of Australia
GPO Box 3947
SYDNEY NSW 2001

Enquiries:

Phone: +61 2 9551 9830
Facsimile: +61 2 9551 8033
Email: rbainfo@rba.gov.au
Website: http://www.rba.gov.au

2013-08	International Business Cycles with Complete Markets	Alexandre Dmitriev Ivan Roberts
2013-09	Terms of Trade Shocks and Incomplete Information	Daniel Rees
2013-10	Stochastic Terms of Trade Volatility in Small Open Economies	Patricia Gómez-González Daniel Rees
2013-11	Issues in Estimating New-Keynesian Phillips Curves in the Presence of Unknown Structural Change	Mariano Kulish Adrian Pagan
2013-12	Central Counterparty Links and Clearing System Exposures	Nathanael Cox Nicholas Garvin Gerard Kelly
2013-13	Inventory Investment in Australia and the Global Financial Crisis	Gianni La Cava
2013-14	Reserves of Natural Resources in a Small Open Economy	Isaac Gross James Hansen
2013-15	Trends in Funding and Lending Behaviour of Australian Banks	Chris Stewart Benn Robertson Alexandra Heath
2014-01	Macroeconomic Consequences of Terms of Trade Episodes, Past and Present	Tim Atkin Mark Caputo Tim Robinson Hao Wang
2014-02	Fiscal Policy and the Inflation Target	Peter Tulip
2014-03	Household Saving in Australia	Richard Finlay Fiona Price

RESERVE BANK OF AUSTRALIA