HOUSING LEVERAGE IN AUSTRALIA

Luci Ellis, Jeremy Lawson and Laura Roberts-Thomson

Research Discussion Paper
2003-09

July 2003

Economic Group
Reserve Bank of Australia

The authors thank Gianni La Cava for providing invaluable assistance with the imputation of household income for some respondents in the Household, Income and Labour Dynamics in Australia (HILDA) Survey, and Steven Stillman of the New Zealand Department of Labour for providing Stata code implementing our main estimation technique. We have benefited from helpful conversations with Alex Heath and other colleagues at the Reserve Bank, and from participants at the 2003 HILDA Conference. This paper would not have been possible without the hard work of the developers and sponsors of the survey. The views expressed in this paper are those of the authors and should not be attributed to the Reserve Bank.

Abstract

A home is the single largest purchase that most households make, and it is one that usually requires some debt financing. Because housing debt is such a large component of households' balance sheets, it is important to understand the financing decision. In this paper, we use household level data from the HILDA survey to relate households' leverage to their observed characteristics using both graphical and econometric techniques. We also model the decisions to own a home and to have debt against it. We correct for any possible selection bias arising from these decisions before drawing conclusions about population behaviour.

Much of the variation in leverage is attributable to the passage of time, as borrowers pay down their loans on schedule and the value of their homes rise. On top of these largely exogenous effects, we find evidence that some households make conscious decisions that strongly affect leverage. For example, Australian homeowners generally plan to pay off their mortgage before its contracted end date, and many are therefore ahead of schedule in paying off their housing debt. On the other hand, a minority of households have higher leverage than similar households because they have engaged in leveraged investment in both owner-occupied and rental housing.

JEL Classification Numbers: D12, G21, R21
Keywords: household survey, housing debt, leverage

Table of Contents

HOUSING LEVERAGE IN AUSTRALIA

Luci Ellis, Jeremy Lawson and Laura Roberts-Thomson

1. Introduction

Australia's housing sector has long been characterised by relatively high homeownership rates and a predominance of variable-rate mortgages. Thus, it might be expected that fluctuations in housing prices would have a relatively strong effect on consumption, and that households would be quite sensitive to interest rates (McLennan, Muellbauer and Stephens 1999). When mortgage rates rise, it impinges on the cash flows of indebted households. This will tend to reduce their consumption, unless they are able to offset this cash-flow effect with further borrowing, or reductions in savings or excess repayments of principal.

In this paper, we focus on a particular dimension of households' balance sheets, the leverage on owner-occupied housing. Although households' debt-income ratios determine the relative effect of a given-sized change in interest rates on their cash flows, we focus on leverage – the debt-assets ratio – because this might help determine the level of debt that households are willing to bear, in addition to the burden of repayments.

Leverage might be expected to influence households' and intermediaries' behaviour through a number of channels. Data from the UK suggests that some households will try to reduce their leverage in response to a negative wealth shock such as a fall in housing prices (Smith, Sterne and Devereux 1994). Households' desired leverage might itself be endogenously affected by the business cycle or uncertainty if choices about balance sheets reflect precautionary savings motives (Carroll and Dunn 1997).

Leverage is also important because of its likely implications for credit supply. Increases in interest rates might not induce households to reduce consumption if they can borrow additional funds, but intermediaries' willingness to lend more to households might be reduced if leverage is particularly high. Households with higher leverage might therefore be less likely to offset fluctuations in their cash flow with further borrowing. Their ability to smooth their consumption during downturns might thus be constrained by asset-price developments associated with that downturn (Bernanke and Gertler 1995; Carroll and Dunn 1997). In addition, since real estate is widely used as collateral for loans, the level of leverage is a determinant of the balance-sheet risk of financial institutions (Kent and Lowe 1997; Schwartz 2002).

Finally, the interaction of leverage with movements in house prices will determine the prevalence of negative equity, which in turn has implications for labour market flexibility (Henley 1999) and the behaviour of the real estate market (Genesove and Mayer 1997). Although we do not cover these issues in this cross-sectional study, an understanding of housing leverage may help explain housing market features such as pricing inertia and the correlation between falling prices and a larger stock of unsold homes.

Despite the range of reasons why households' leverage might be important, most previous literature either focuses on the narrower decision on housing tenure (Bourassa 1994, 1995), or uses broader wealth data as yet unavailable for Australia to examine household portfolio behaviour more generally (e.g., the papers in Guiso, Haliassos and Jappelli (2002)). In this paper, we focus on cross-sectional, microeconomic aspects of households' housing leverage, with a view to understanding which households are most likely to be affected by changes in interest rates or falls in housing prices. In Section 2, we discuss the HILDA dataset and our approach to imputing missing income data; the econometric results underlying our imputation methods are presented in Appendix A. After undertaking some preliminary graphical analysis in Section 3, we set up our core econometric model in Section 4 and discuss its implications. A brief conclusion follows in Section 5.

2. The HILDA Dataset

In this study we use data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. HILDA is a household-based panel or longitudinal survey that aims to track all members of an initial sample of Australian households over time. The survey was commissioned by the Department of Family and

Community Services and is directed by the Melbourne Institute. This study uses data from the first wave of the survey, collated from interviews conducted with some 14 000 individuals living in almost 7 700 households over the second half of 2001. The survey contains a wealth of possible explanatory variables, covering four broad areas: economic wellbeing, labour market dynamics, family dynamics, and subjective wellbeing.

2.1 Constructing the Housing Leverage Variable

Housing leverage is typically expressed as a loan-to-valuation (LTV) ratio, in contrast to the debt-equity ratios commonly used in analysis of corporate finance. We can construct LTV ratios for homeowners' principal residence using information from the household questionnaire in Wave 1 of HILDA. Homeowners were asked to estimate the current value of their home, and to report the amount they currently owed on loans taken out against that home, including institutional mortgages, loans from family, friends and other members of the community, and home equity loans. To calculate the LTV ratio, we total the outstanding amounts of all borrowings against the principal residence and divide it by the estimate of the home's value. Households who rented, or who occupied their home rent-free but did not own it, were not asked these questions. Information on other properties such as investment properties or holiday homes was not included, so our study relates specifically to owner-occupiers' principal residences.

The use of subjective valuations raises the question of their accuracy. Goodman and Ittner (1993), using US survey data, find that there is a small positive bias of about 6 per cent in homeowners' estimates, but that the mean absolute error of estimates tends to be larger at about 15 per cent. This may be due to rounding errors: 42 per cent of households reporting estimated home values in HILDA reported a figure that was a multiple of $50 000. We believe our analysis is unlikely to be significantly biased by homeowners' subjectivity. If the bias in their estimates is small and, as Goodman and Ittner find, unrelated to owners' characteristics, then our point estimates will not be significantly biased even if these rounding errors make them less precise. In any case, households' behaviour presumably depends on their perceptions of their leverage rather than realised leverage, especially if they are not intending to sell their homes in the near future.

2.2 Missing Data and Income Imputation[1]

Compared with similar international household surveys, HILDA does not suffer greatly from problems of missing data (Watson and Wooden 2003). For example, only 4 per cent of households with mortgages fail to report the value of their loans, while the value of the principal residence is missing for only 6 per cent of owning households; these households are excluded from the results below. However, there is a relatively high incidence of missing data for income-related questions. We can separate the most common reasons for non-response into 'item non-response' and 'incomplete households'. *Item non-response* occurs when a member of a selected household agrees to be interviewed, but then either refuses, or is unable, to answer some of the questions asked. This is the main source of missing data, accounting for 64 per cent of the missing household income information. Most of the missing income data is due to item non-response for income sourced from business (missing 23.5 per cent of people with business income) and investments (missing 8.1 per cent of recipients of this kind of income). Wages and salaries (missing 7.2 per cent of wage earners) and government benefits and pensions (missing 1.4 per cent of benefit recipients) have lower incidences of missing data.

The other major source of missing data is the 810 *incomplete households*, accounting for 10.5 per cent of the household sample and 36 per cent of the missing household income information; these are households in which not all eligible adult members agreed, or were able, to be interviewed. The HILDA dataset as distributed does not include an entry for household income if any of its eligible members were not interviewed, or did not report complete income information; in all, 29 per cent of households have a missing value for household income, which is clearly an unacceptable data loss.

In such circumstances we have two choices. We can drop the 29 per cent of households for which income data is missing from the sample, or impute the income of the individuals with missing data. Our choice to impute income for missing individuals is shaped by two factors. First, because income non-response is not random or uncorrelated with the variable(s) of interest, the missing cases cannot be safely dropped from the sample (Watson and Wooden 2003). For example, men, individuals outside the labour force, and people with large amounts

[1] This section and Appendix A report work by Gianni La Cava and Jeremy Lawson.

of leisure time (and generally have low incomes) were more likely to offer complete income information than other individuals. Second, we have a large cross-section of information from the HILDA Survey that presumably permits us to do a reasonable job of imputing income where the information is missing.

2.2.1 Imputation methodology

Following the recommendations of the HILDA Survey team and methods adopted in the British Household Panel Survey (BHPS), we impute income using the predictive mean matching method (Little 1988; ISER 2002; Watson and Wooden 2003). This is a stochastic imputation technique that has the advantage of maintaining the underlying distribution of the data by allowing the imputation of error around the mean. Appendix A outlines the method in detail and shows the regression results for the three models.

The nature of the missing data leaves us with the need to impute income for three separate types of missing cases:

1. Individuals that did not complete a person questionnaire and therefore did not report any income information (Type I) (n = 1 158).

2. Individuals that completed a person questionnaire but did not provide information on wage income (Type II) (n = 673).

3. Individuals that completed a person questionnaire but did not provide information on non-wage income (Type III) (n = 1 621).

Three separate models are estimated to impute income for each type of missing case. For Type I respondents we have information on the characteristics of their household (e.g., value of the dwelling, geographic location, the number of bedrooms) and a limited range of personal information from the household questionnaire. We also have personal information collected about other respondents in the household. These 'family variables' include the income, labour force status and occupation of other household members. Both the household and family variables are likely to be correlated with both personal and household income and hence act as useful explanatory variables in the model. We impute total gross financial year income for these individuals. For Types II and III respondents we also have additional personal information obtained from items that they did complete – labour force status, age, gender, English-speaking background – including information about the sources of their income. This allows us to predict wage and non-wage income, and add to it the income that individuals report from other sources. For example, for Type III individuals we add their imputed non-wage income to any actual reported wage and salary income.

2.2.2 Imputation results

In the regression model for Type I households our model explains nearly 32 per cent of the variation in total gross household income. The root mean square error (RMSE) is about $26 000. In the regression model for Type II households our model explains about 46 per cent of the variation in individuals' wage and salary income and the RMSE is nearly $19 000. In the regression model for Type III households our model explains nearly 21 per cent of the variation in individuals' non-wage income and the RMSE is about $20 500. Although these errors are quite large, we regard the imputation as being relatively successful, not least because it allows us to use reported income from other income sources and household members that would otherwise be lost.

Our income imputation strategy allowed us to recover household income estimates for all but 201 households (about 3 per cent of the sample), ensuring that any bias introduced by dropping missing observations from the sample is minimised. However, because our imputed household income estimates are likely to diverge from the true income that households did not report, we also construct a dummy variable for those households with imputed household income. This dummy was not significant in any of our three equations, implying that inclusion of households with imputed income did not significantly distort our results.

2.3 User Cost of Housing

The model outlined in Section 4 includes a model of households' tenure decisions. These depend on the utility they gain from owning rather than renting, and the relative costs of each tenure type. The relative cost of owning compared to renting is calculated by multiplying housing *i*'s user cost of housing, u_{im}, by the price-rent ratio in the relevant geographical area m. User cost captures the net per-unit cost of owning for owner-occupiers. We calculated it similarly to Bourassa (1995) as Equation (1), so as to take account of the details of Australia's tax system – specifically, that mortgage interest payments are not deductible.

$$u_{im} = (1 - t_{im})(1 - v_i)r + v_i r - \pi_m + \delta \tag{1}$$

As is standard in the literature, per-unit user costs include the rate of depreciation δ, and the interest repayments on any mortgage, which in turn depend on interest rates r. Because interest payments are not deductible, this cost also depends on the mean expected leverage for household i's age group, v_i. As in Bourassa (1995), this depends on the household's age and its permanent income, as estimated using its observable characteristics. The use of a group mean, not actual leverage, minimises any endogeneity concerns. There are also tax benefits to owner-occupation because the flow of housing services (imputed rent) is not taxed but actual rent is paid from post-tax income. Therefore there are tax and income costs to moving from owning to renting, as shown in the first term in Equation (1).[2] Owner-occupiers also accrue the benefit of expected capital gains on their home, π_m, which are calculated here using past housing price inflation in market m. Bourassa (1994, 1995) contains more detail on the calculation of the components.

3. Preliminary Analysis

Before we outline an estimated model of household housing leverage in Section 4, we first present some stylised facts about housing leverage in Australia, identify some of the household characteristics likely to influence households' housing leverage, and undertake a brief graphical analysis of how three key variables – age of the household head, household income, and household housing wealth – are related to leverage.

2 The income cost per unit of housing from moving from owning to renting is $(1\text{-}v_i)r$. t_{im} is the ratio of the change in tax payable to the change in income in moving from owning to renting, where the difference between the payable taxes in the tenure states are calculated using the household's permanent income. Thus, $(1\text{-}t_{im})(1\text{-}v_i)r$ is the per unit income and tax cost of moving from owning to renting.

According to the HILDA Survey the homeownership rate in 2001 was 68 per cent, broadly consistent with both the 2001 Census and the 1998 Household Expenditure Survey (HES). Of these owner-occupying households, 42 per cent were still paying off their home and a further 9 per cent had other kinds of debt secured against it, such as a second mortgage or home equity loan. Households with some leverage therefore represented around one-third of the households surveyed. The average level of household leverage implied by the HILDA data is a little below the level implied by the aggregate credit data, at around 15 per cent for all owner-occupying households. This is despite the inclusion of loans from friends and family that are not in the aggregate data; these account for just under half of a percentage point of the average. The leverage of only those households with outstanding debt on their owner-occupied home averages around 48 per cent, which given the definitional differences in the debt measure, is roughly consistent with aggregate figures.

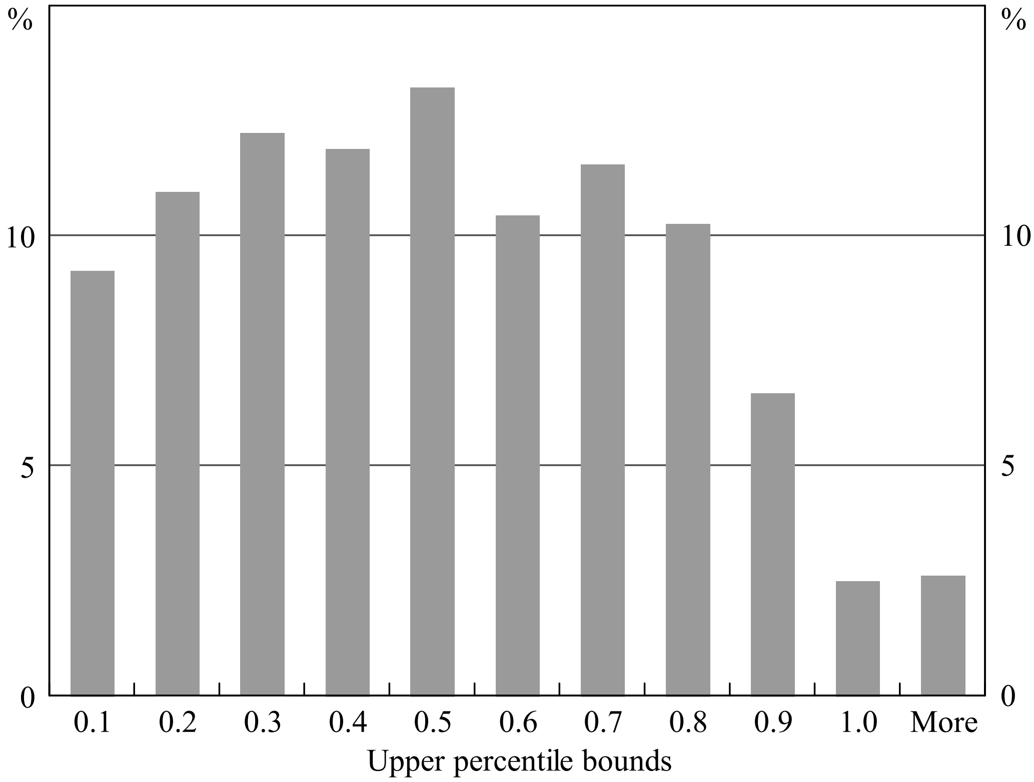
Although these figures confirm that, on average, housing debt is much lower than the value of housing assets, we are also interested in how this leverage is distributed across Australian households. Figure 1 shows that the distribution of leverage is fairly evenly spread, but few Australian households have high leverage, that is, an LTV ratio greater than 80 per cent.[3] Only 11.4 per cent of households with housing loans (less than 4 per cent of all households) had an LTV ratio that exceeded 80 per cent, and less than 3 per cent had negative equity in their homes.[4]

3 Our use of an 80 per cent threshold is consistent with evidence in Genesove and Mayer (1997) that household behaviour alters when leverage exceeds this level, and it allows us to consider how the incidence of negative equity might change if house prices were to fall by a significant amount. It is also the threshold above which borrowers require mortgage insurance.

4 There may be other households with high leverage against other properties not captured here, such as an investment property. Unfortunately this possibility cannot be explored until Wave 2 HILDA data, including information on wealth, are released in late 2003.

Figure 1: Distribution of Loan-to-valuation Ratios

Percentage of households with housing loans



3.1 Leverage and Household Characteristics

We can use the wealth of cross-sectional information about Australian households contained in the HILDA dataset to investigate whether the diversity of household leverage is related to observable characteristics.

There are a number of household characteristics that can be expected to influence housing leverage. First, we would expect leverage to be affected by a household's stage in the life cycle. Leverage should be low for young households because their homeownership rates are low; few younger households would have saved enough to purchase a home while they are still completing their education and establishing their careers. Indeed, younger people may not have formed households at all for these reasons. Leverage should then rise as households move into their peak family formation years. This usually occurs early in their working lives when they have accumulated little wealth, so they must borrow to purchase large assets such as the family home. This debt is then steadily reduced over their working lives so that they are relatively debt-free when their incomes drop sharply upon retirement.

Variables that capture a household's position in the life cycle, such as the household head's *age*, *labour force status*, and *marital status* should therefore be important in explaining housing leverage. We also expect that the time since the mortgage was taken out, which we proxy using *time the household has lived in the home*, will be an important determinant of household leverage.

A household's means, and in particular its *income*, should also affect its leverage by influencing its willingness, need and ability to take on housing debt. We expect leverage to increase with income for three main reasons. First, higher-income households are more likely to be homeowners in the first place, because they are more likely to have been able to save the necessary downpayment. Second, we expect households that have paid off their mortgages to have lower incomes than households with existing mortgages, because retired lower-income households prefer to have paid off their debt. Finally, housing leverage may be expected to increase with income if financial institutions apply easier lending criteria to high-income households. The expected importance of household means in predicting leverage also leads us to anticipate that alternative indicators of household means, such as self-assessed *income adequacy* and the past occurrence of *family breakdown* should also influence leverage. Offsetting this, high-income households have more scope to pay off their loan quickly and thus own their home outright.

Finally, we would expect households' attitude to homeownership and debt to influence their housing leverage. For example, households that are *uncomfortable about taking on debt* may accumulate more savings before purchasing a home than households that feel more comfortable taking on debt. They may also choose to pay their loans off faster. Other variables that may pick up different attitudes to homeownership or debt across households include households' *ethnic background*, *aversion to risk*, and *credit card usage*.

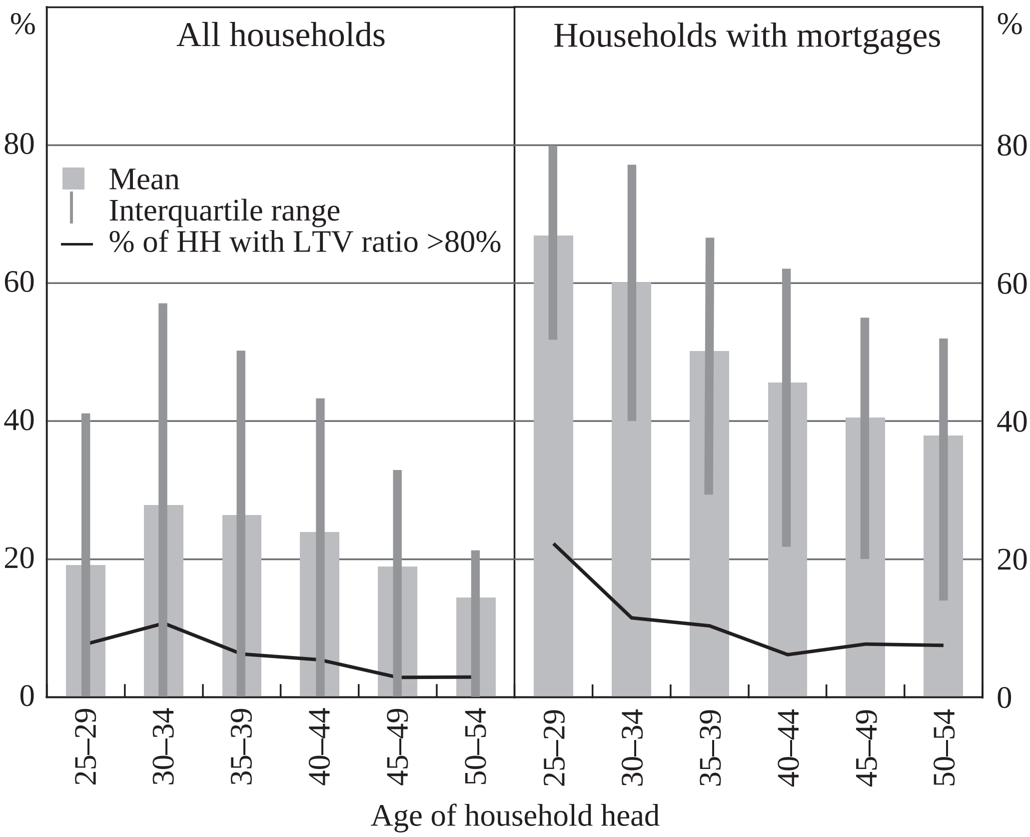
3.2 Graphical Analysis

In Section 3.1 we identified a number of household characteristics that can be expected to influence a household's housing leverage. Before using this information in an econometric model, it is worth examining graphically how three key variables – age of the household head, household income, and household housing wealth – are related to housing leverage, treating households without mortgages as having zero leverage.

3.2.1 Age of the household head

Figure 2 shows how housing leverage varies across age cohorts. In addition to mean leverage for each cohort we also include information about the distribution of leverage – the interquartile range, and the percentage of households with an LTV ratio greater than 80 per cent. The right-hand panel contains data only for households with mortgages, while the left-hand panel includes all households.

Figure 2: Housing Leverage by Age Group



Looking first to the right panel, younger households with mortgages appear to have higher leverage than their older counterparts. Average leverage was almost 70 per cent for households where the household head was aged between 25 and 29, compared with 40 per cent for households with a head aged between 45 and 49. Similarly, younger households are more likely to have high leverage. Just under one-quarter of leveraged owner-occupying households with heads aged between

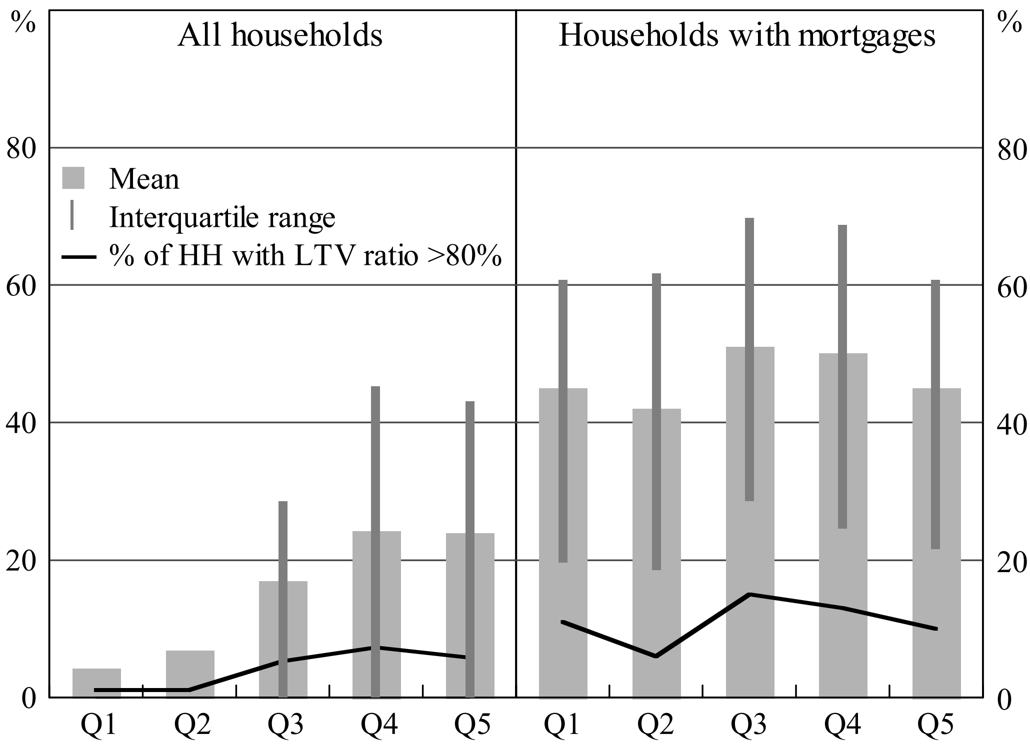
25 and 29 had leverage exceeding 80 per cent, a number that falls to just under 8 per cent for households with heads aged between 45 and 49. Because there is some association between age of the household head and time spent in homeownership, it is difficult to distinguish whether this is a pure age effect, or a reflection of the time the household has had to pay off their initial loan. This issue will be investigated in our econometric analysis in the following sections.

The left panel of Figure 2 shows that when younger households' lower homeownership rate is taken into account, their apparent higher leverage is reduced, and the nature of the relationship between age and leverage changes. Average leverage now increases to a peak when the household head is aged 30 to 34 before falling away as households pay off their housing debt. This observation is important because it shows that if we use a sample that only includes households with mortgages, we may exaggerate the incidence of leverage for categories of households with low homeownership rates, or higher rates of outright ownership where the mortgage has already been paid off. In our graphical analysis we make use of both samples so that we can observe pockets of high leverage in categories that otherwise have low debt. Similarly, in our empirical analysis, we use a technique that enables us to capture both the marginal effect of variables on leverage for those who have debt and the effect of these variables on the probability of having debt, giving us the total effect of variables on leverage.

3.2.2 Household income

Figure 3 suggests that there may be a hump-shaped relationship between household income and leverage. Average leverage is greatest among upper-middle income households, and falls away somewhat for both low-income and high-income households. When the lower homeownership rate and share of households with existing loans in the bottom two income quintiles is taken into account (left panel), their average leverage looks particularly low. This may be partly an age effect; the average age of household heads in quintiles 1 and 2 is higher (at 55) than in the three higher income quintiles (at about 42).

Figure 3: Housing Leverage by Income Quintile

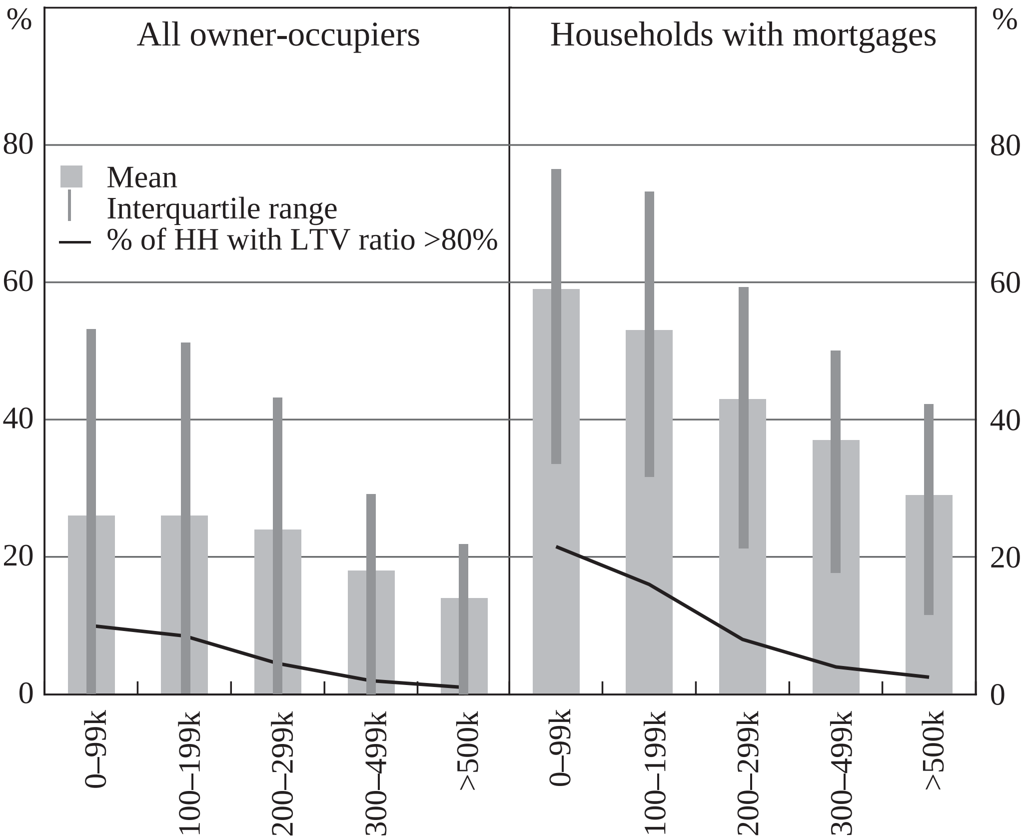


3.2.3 Housing wealth

Although the Wave 1 HILDA Survey did not ask households about their non-housing wealth, their estimates of the current market value of their homes should capture a large proportion of their assets. According to Figure 4 there is a strong negative relationship between households' housing leverage and the value of their properties. Households with properties worth less than \$100 000 had an average leverage (at almost 60 per cent) roughly double that of households with properties worth in excess of \$500 000. The low average incidence of leverage among households with expensive properties is reflected in the much lower likelihood of their being leveraged above 80 per cent, at 3 per cent compared to over 20 per cent for households with homes worth less then \$100 000. However, because owner-occupying households that have paid off their mortgages were more likely to have homes worth less than \$100 000, their average leverage falls to about the same amount as households with homes worth between \$100 000 and \$199 000 when this is taken into account (left panel).

There are a couple of possible explanations for the negative relationship between housing wealth and housing leverage. First, if a household's housing wealth is positively correlated with its age and tenure length, then the observed relationship may only suggest that younger households, which are more likely to be highly leveraged, are also more likely to own cheaper homes. Second, if more expensive homes have appreciated more rapidly than less expensive homes, households that own more expensive homes will have experienced a larger passive reduction in their leverage, for a given length of tenure. Shocks to the price of a particular home have the opposite effect on the leverage ratio. If households update their assessments of their homes' current value for past price changes without adjusting their housing debt, this could generate a negative relationship between current estimated housing wealth and leverage.

Figure 4: Housing Leverage by Value of Principal Home



One possible implication of the negative relationship between leverage and housing wealth is that leverage might be lower in locations with the highest housing prices, holding everything else constant. It is these areas – New South Wales and Victoria, and the inner suburbs of state capitals – that have experienced the strongest rises in housing prices recently, and thus where passive reductions in leverage are likely to have been greatest. In general, the data support this argument. Average leverage is relatively low in New South Wales and Victoria, and particularly in their inner-city suburbs, as is the incidence of high leverage. For example, in the late 2001 fieldwork period of this survey, average leverage for households with leverage in inner Melbourne was around 35 per cent (38 per cent in inner Sydney), compared to over 50 per cent in inner Brisbane and Adelaide. Indeed, only 1 per cent of households living in inner Melbourne in late 2001 had an LTV ratio above 80 per cent.[5] Provided the average length of tenure (i.e., the frequency of households' decisions to trade up) is not greatly affected by the rate of house price inflation, this provides some comfort that those areas where concerns about future reversals of past price increases might be greatest are those where the negative equity consequences are least. This strengthens our general conclusion that the moderate level of aggregate housing leverage in Australia does not seem to be masking important pockets of potential vulnerability.

4. Econometric Model and Results

4.1 Model and Methodology

The only households with positive leverage are those that currently have debt against their home. By definition, leverage cannot be negative, so its distribution is censored at zero. The full sample of households also contains a large fraction of renters and outright homeowners who have no owner-occupied leverage. Although tobit techniques can be used to estimate some censored samples, they are inappropriate in the current context. Tobit models require that the explanators of whether the household has any leverage at all are the same as the determinants of its level. However, characteristics of the *loan* can be (and here turn out to be) important determinants of the level of leverage, but they are undefined for households without housing debt. Therefore they cannot also be used to explain whether or not a household has any leverage.

On the other hand, estimating a model of leverage only over the sub-sample with debt does not allow assessment of the whole population's behaviour.

5 Classifying postcodes into inner and outer city groups requires some subjective judgment. We defined suburbs as inner city if they were fairly close to the CBD and considered prestige or fashionable places to live; the specific allocation of postcodes is available from the authors.

Gronau (1974) and Lewis (1974) showed that a censored sample may have unrepresentative characteristics relative to the whole population. Since this can bias OLS parameter estimates (Heckman 1976, 1979), we therefore model leverage using an appropriately specified selection model, which corrects for this bias.

Table 1 illustrates that the leveraged sub-sample of the dataset used here could be unrepresentative. Households with housing debt have quite different age and income characteristics than the whole sample. The probability of having a mortgage is greatest for those aged 35–39, irrespective of household income. As income increases, so does the probability of having a mortgage. These two features imply that the sub-sample of households with a mortgage is likely to have a higher median income than the general population, and an over-representation of middle-aged household heads.

Table 1: Effect of Age and Income on Having a Mortgage

Per cent of people in each category with a mortgage

Age group	**Income quintile**				
	Q1	Q2	Q3	Q4	Q5
20–24	9.2	9.9	17.6	29.8	39.3
25–29	12.3	14.2	35.0	44.5	43.9
30–34	21.6	32.5	45.0	57.1	65.5
35–39	28.4	36.7	49.8	66.7	**68.4**
40–44	20.3	36.5	46.0	63.0	63.6
45–49	12.1	27.5	39.6	50.0	58.1
50–54	23.1	27.8	35.2	41.1	46.0
55+	**3.7**	4.5	12.7	21.6	23.2

Heckman (1976) proposed a two-step procedure to correct for selection bias, involving modelling of the rules that determine inclusion into the sample. The results from these selection models are then used to adjust the estimates for the equation of interest so that they capture population responses rather than those of the sub-sample over which they are estimated.[6] Households with owner-occupied

6 In an uncorrected model, the marginal effects only apply to those that actually have mortgages, while the bias-corrected (or population) marginal effects apply to all households if they were to have mortgages. The derived 'population' marginal effects are thus still conditional on the household having a mortgage.

leverage select themselves into this group by both owning their home and having some debt secured against it. Thus our model involves two selection rules: the *tenure decision* on whether or not the household owns rather than rents their home; and the *mortgage decision* on whether or not the household has housing debt.

The error structure of the model depends crucially on whether the two decisions are made sequentially or jointly. The distinction between sequential and joint decisions relates to the interdependency of the two decisions, rather than their timing, as illustrated in Figure 5. If the two decisions are defined over the entire set of observations, so that all four cells are logically possible, then they are made jointly, regardless of whether all the outcomes represented by the cells are actually observed (Lee and Maddala 1985; Tunali 1986). The standard case of choices about education and labour supply is an example of a joint decision process (Fishe, Trost and Lurie 1981); an individual could choose to enter the labour force or not regardless of whether they had previously completed a particular level of education. Sequential decisions, on the other hand, are characterised by one decision only being defined given a particular *outcome* of the other decision (Maddala 1986). This is the case with our sample selection rules; a household cannot 'decide' to have a mortgage on an owner-occupied property if they are not owner-occupiers; cell 3 is infeasible.

Figure 5: Selection Rule Matrix

		Selection equation 2	
		Yes	No
Selection equation 1	Yes	1	2
	No	3	4

Given this structure for the selection rules, we have a system in which the tenure selection equation is used to adjust the mortgage selection equation, which is only defined if the household is a homeowner ($T_i = 1$). These selection equations are estimated using maximum-likelihood probit techniques. As in all probit models, the selection equations assume that the observed decisions T_i and M_i reflect the values of unobserved latent variables T_i^* and M_i^*. If $T_i^* > 0$, $T_i = 1$ (the household is an owner-occupier) while if $T_i^* \le 0$, $T_i = 0$ and the household rents. Similarly if a home-owning household *i* has housing debt, then $M_i = 1$, implying $M_i^* > 0$.

The results from the selection equations can then be used to adjust the equation explaining leverage (lev_i), as shown in Equation (2), so that OLS estimates of β are consistent, although they are inefficient compared to maximum likelihood techniques. The second-stage OLS estimates of the error variance and covariance matrices are biased, but we make an approximate correction for this using the delta method (Heckman 1979; Tunali 1986). The parameter estimates are also quite sensitive to violations of the underlying normality assumptions; our preferred specifications preserve normality, but possibly at the cost of some efficiency.

$$\begin{aligned} T_i^* &= Z_{1i}\gamma_1 + e_{1i} \\ M_i^* &= Z_{2i}\gamma_2 + \lambda_T\gamma_3 + e_{2i} \\ lev_i &= X_i\beta + \lambda_1\alpha_1 + \lambda_2\alpha_2 + v_i \end{aligned} \tag{2}$$

The additional variables λ_1 and λ_2 in the leverage equation capture the effect of selection on the error term: $\alpha_1\lambda_1 + \alpha_2\lambda_2 = E\left(u_i \middle| T_i^* > 0, M_i^* > 0\right)$. Because this is a double-selection model, they are not identical to the inverse Mills ratio in a standard Heckman-Lee single-selection model.

Precise estimation of parameters depends on the three equations being sufficiently identified, which requires that the sets of explanators Z_1, Z_2 and X have at least some non-overlapping variables, so that the regressors for one equation are not simply capturing the same information as the selection terms (λ_i). Our preferred specification involves several variables, such as user cost of housing, that are specific to one of the selection equations. It is nonetheless difficult to be certain that robust identification has been achieved; the significance of the selection terms in the leverage equation was sensitive to the inclusion in the selection equations of some statistically insignificant variables. The need to ensure identification and avoid multicollinearity result in a tendency to favour parsimonious specifications for our model, relative to the richness of the HILDA dataset.

The residual in the tenure equation, e_{1i}, is assumed to be distributed as a standard normal; e_{2i} is also standard normal, but defined only on the subpopulation $T_i^* > 0$;

and v_i is normally distributed on the subpopulation $T_i^* > 0$, $M_i^* > 0$ with variance σ_v^2 (Lee and Maddala 1985). The three errors may be correlated; the technique used here allows for this. Many of the unobserved factors explaining past mortgage repayments, and thus whether a mortgage has now been paid off, will be unrelated to the past tenure decision. There are, however, good reasons to suspect that e_{1i} and e_{2i} could be correlated. At least some households' decisions to rent are driven by their inability to obtain housing finance. Financial institutions' willingness to lend might depend on variables that are not observed in the HILDA dataset such as savings history, past loan defaults and so on. Their influence would therefore end up in the residuals of both selection equations.

4.2 Estimation Results

Table 2 shows the parameter estimates and Table 3 the prediction outcomes for the two selection equations. The fit and prediction rates are reasonable, and within only a few percentage points of results in previous studies of tenure choice in Australia (Bourassa 1994, 1995), despite the absence of information about important components of households' wealth portfolios in HILDA. As expected, age, income and other indicators of life-cycle such as marital status are important determinants of homeownership. Age, income and other financial indicators were also strong determinants of whether home-owning households had housing debt. The indicator variables of whether anyone in the household made a profit or loss in the form of rental income are highly significant in the mortgage equation. Households that made a profit on their rental properties were significantly less likely to have a mortgage on their owner-occupied property than households without such income. In contrast, households that made a rental loss – that is, their expenses, including interest, on any mortgages against the investment property exceeded rental income – were significantly more likely to also have a mortgage on their home.

Table 2: Household Tenure and Mortgage Decision Regression Results

	Tenure	Mortgage		Tenure	Mortgage
Personal characteristics of the household head and demographic variables					
Age	1.4***	–2.6***	Married	30.4***	
Retired		–41.2***	Born in UK	–10.5	
Born in Europe[(a)]	–22.3		Father born in UK	–2.3	10.2
Father born in Europe[(a)]	27.2**		Student	–8.5	
Group household	–71.1***		Children at home		7.5***
Only speaks English at home		25.5***	Receives youth allowance		–118.0***
Household means variables					
Income ($'000)	1.6***	0.5***	Income squared ($'000)	–0.8***	–1.0***
Income cubed	1.5**		Interest income		–7.5***
Casual job		–18.9**	Income adequacy		–10.8***
Made profit on rental property		–22.1**	Made loss on rental property		30.5***
Financial and housing variables					
Time at address	6.5***	–2.7***	Short saving time horizon		12.1*
Condition of home[(b)]	–26.0***		Importance of home[(b)]	7.5***	
Moves in last 10 years		11.0***	Not moved recently		2.2
Housing adequacy[(b)]	13.0***		Value of home		0.02*
Small apartment block	–58.3***		Tall apartment block	–91.0***	
Semi-detached dwelling	–59.1***		Pays off credit card on time[(b)]		–5.6**
Has credit card		–35.6***	No of bedrooms	21.3***	
Took out institutional loan		167.0***	Relative cost of owning	–10.3***	
Location variables					
Inner Sydney		24.5*	Non-metro NSW		22.7**
Outer Melbourne		29.8**	Non-metro VIC		38.5***
Outer Brisbane		45.9***	Non-metro QLD		46.1***
Outer Adelaide		35.0***	Inner Perth		53.9*
Outer Perth		46.7***	Non-metro WA		58.9***
Log-likelihood	–4 064.955		Correct prediction rate	85.6%	85.8%

Notes: ***, **, and * represent significance at 1, 5 and 10 per cent levels.
(a) Europe refers to Italy, Greece, Netherlands or Germany.
(b) Denotes ordered categorical variables.

Table 3: Prediction Rates for Selection Equations

Tenure equation

		Actual		
		Own	Rent	Total
Predicted	Own	4 662	641	5 303
	Rent	404	1 543	1 947
	Total	5 066	2 184	7 250

Mortgage equation

		Actual		
		Mortgage	No mortgage	Total
Predicted	Mortgage	2 197	463	2 660
	No mortgage	258	2 148	2 406
	Total	2 455	2 611	5 066

Prediction rate for tenure equation = 85.6%
Prediction rate for mortgage equation = 85.8%

The central results from the estimation of the equation for household housing leverage are contained in Table 4. Overall the results are reasonable, with our model explaining about 30 per cent of the variation in households' housing leverage and an *F*-test of the null hypothesis that all the coefficients are equal to zero rejected at all conventional levels of significance. The estimated parameters from the model correspond to the total effect of a one-unit change in that variable on leverage, expressed in percentage points of leverage. For variables that are also in the selection equations, this total effect includes the variable's effect on leverage through its effect on the probability that the household has leverage at all. Table 4 therefore also shows the *marginal effects* of these variables on leverage, conditional on the household having leverage, which can be calculated from the estimated coefficients on the selection terms using standard methods. For ease of interpretation, we split explanators in the equations into three categories of household characteristics – the stage of the household in the life cycle, household means, and characteristics relating to finance and housing – and a group of location dummies to control for region-specific effects and the effect of unexpected changes in housing prices on leverage. Other demographic variables that might be expected to be associated with leverage, such as the education level, occupation or gender of the household head, were insignificant and thus excluded from the final specification.

Table 4: Household Housing Leverage Regression Results

	Total	Marginal		Total	Marginal
Personal characteristics of household head and demographic variables					
Age	–0.6***	–0.5	Retired	–11.8**	–10.5*
Married	3.3	3.2	Divorced	–1.5	–1.5
De facto	3.8	3.9	Separated	11.0***	11.0***
Born in UK	–5.0*	–5.0*	Born in Europe(a)	1.9	1.9
Father born in UK	4.9**	4.6*	Father born in Europe(a)	–6.8***	–6.9**
Student	7.2	7.3			
Household means variables					
Income ($'000)	0.07(c)	0.04	Income squared ($'000)	–0.2	–0.1
Income adequacy(b)	–1.1	–0.7	Satisfaction with pay(b)	0.8***	0.8***
Casual job	–7.1***	–6.5***	Fixed-term contract	–2.8	–2.8
Likelihood of losing job in next 12 months(b)	–0.03	–0.03	Had difficulty paying mortgage on time	4.7**	4.7**
Has rental income	14.2***	14.2***	Has imputed income	1.0	1.3
Financial and housing variables					
Time lived at address	–0.8***	–0.7**	Planned pay off date	1.0***	1.0***
Ahead of mortgage repayment schedule	–5.4***	–5.4***	Satisfaction with neighbourhood(b)	–1.0***	–1.0***
Long saving horizon	4.0***	4.0***	Took out institutional loan to buy house	18.2*	12.6
Attitude to borrowing	1.1	1.1	Attitude to risk(b)	–0.8	–0.8
Moves in last 10 years	1.2	0.8	No recent moves	–5.3***	–5.3***
Condition of home(b)	3.0***	3.1***	Not a first-home buyer	6.2***	6.2***
Semi-detached dwelling	5.1**	5.3**	Pays credit card on time(b)	–1.2***	–1.2**
Location variables					
Inner Sydney	5.4	4.5	Outer Sydney	13.9***	13.4***
Non-metro NSW	15.0***	14.2***	Outer Melbourne	14.7***	13.7***
Non-metro VIC	18.1***	16.8***	Inner Brisbane	22.1***	22.1***
Outer Brisbane	23.9***	22.4***	Non-metro QLD	23.2***	21.6***
Inner Adelaide	18.2***	18.2***	Outer Adelaide	21.0***	19.8***
Non-metro SA	28.5***	27.6***	Inner Perth	9.5	7.7
Outer Perth	21.0***	19.4***	Non-metro WA	19.2***	17.2***
Inner Hobart	2.8	2.8	Outer-city TAS	22.5***	22.5***
Non-metro TAS	42.2***	41.0***	ACT	24.0***	23.3***
NT	31.0***	31.0***	House price growth	–0.1	–0.1
Tenure selection term	2.4	2.4	Mortgage selection term	17.4(c)	17.4
Adjusted R^2	**0.278**		**No of observations**	**2 327**	

Notes: ***, **, and * represent significance at 1, 5 and 10 per cent levels.

(a) Europe refers to Italy, Greece, Netherlands or Germany.

(b) Represents ordered categorical variables.

(c) Represents significance at 11 per cent level.

4.2.1 Stage of the household in the life cycle

A household's stage in the life cycle should be a key explanator of a household's housing leverage. Taking the results in Table 1 at face value, one might expect housing leverage to rise as households purchase homes and acquire debt during their family formation years, and then fall as they repay this debt. Our results are, however, mainly suggestive of leverage falling monotonically as households move through the life cycle. A household head that is one year older is associated with up to 0.6 percentage points less leverage, while an age-squared term was negative but insignificant if included. This contrasts with the positive coefficient on age and negative coefficient on the square of age that would be required to obtain a hump-shaped profile.

A number of other variables are also suggestive of leverage falling as households move through the life cycle. For a given term for the mortgage, a move-in date that is one year earlier than an otherwise identical household implies that the household's housing leverage will be about 1.8 percentage points lower, since both the move-in date and the expected payoff date are then one year earlier. This suggests an important role for passive paydown of mortgages on schedule; households that are older and have lived in their homes for longer have lower leverage in part simply because they have had longer to pay their mortgage off. In addition, rising housing prices imply that mortgages taken out earlier were likely to be smaller, and thus a smaller proportion of the home's current estimated price. On the other hand, some households must be making explicit decisions about the end date of their mortgage; otherwise, move-in date and payoff date would not both be significant because they would then be more closely correlated. Indeed, the expected loan terms implied by the move-in and payoff dates are overwhelmingly shorter than the 20 to 25-year terms generally specified in loan contracts. These financial decisions are likely to be dependent on households' assessments of their ability to achieve the targeted payoff date, and thus on their incomes and means more generally. Nonetheless, the significance of the coefficient on expected payoff date in a multivariate setting indicates that these expectations are also in large part independent of income and other factors, instead capturing household preferences about debt duration and portfolios that are not explained by income and demographic variables alone.

The combined size and precision of the estimated coefficients on these two variables suggest that this passive paydown effect is more important than any pure age effect. On the other hand, the significance of the coefficient on age after controlling for these factors indicates that age on its own also plays a role in determining leverage. This might be because older households have had longer to accumulate wealth (something that we cannot measure using the first wave of HILDA). Moreover, other significant explanators of leverage also suggest a more explicit life-cycle interpretation. Married household heads and those in de facto relationships tend to have slightly higher leverage than homeowners that have never married, although the difference is not significant at conventional levels. Households with retired household heads were both significantly less likely to have a mortgage, and to have lower leverage than other households when they did.

4.2.2 The means of the household

Our expectation that leverage should rise with household means is generally supported by the data. For example, our main indicator of household means, household income, is positively related to leverage, although it is only statistically significant at the 11 per cent level. The magnitude of the effect is small, with each extra $10 000 of income associated with, at most, only 0.7 percentage points higher leverage. Increased income appears to be more closely associated with higher values of both debt and housing assets than with the ratio of those two variables. We also find tentative evidence that the effect of household income on leverage is non-linear, with the negative sign on the income-squared variable implying that leverage is increasing in income but at a decreasing rate. The coefficient on this term is so small that leverage does not begin falling until income reaches about $500 000.

Two measures of households' subjective views about their incomes – the ease with which they are making ends meet (income adequacy), and their satisfaction with their pay – were included in the model to determine whether such measures of income relative to perceived requirements add information above that of measured income. Neither variable is closely correlated with reported actual income. Reported satisfaction with pay is statistically significant with a small positive coefficient, perhaps indicating that respondents that are highly satisfied with their pay are less concerned about their vulnerability to future income shocks, and are thus more willing to take on higher gearing.[7]

Past loan repayments in excess of the contracted minimum will naturally result in lower leverage at a given point in time, all else equal. Although the HILDA Survey does not explicitly ask about the extent of these past overpayments, respondents were asked if the loan repayments were (currently) ahead of schedule, behind schedule, or about on schedule. Households that reported that they were ahead of schedule in the repayments on the main loan against their home had leverage 5.4 percentage points lower than that of otherwise similar households.

The means of the household captures more than just its income. Factors such as family breakdown, permanency of employment, negative income shocks and familial support could also influence leverage by affecting households' ability to take on and pay off debt. Our results suggest that such factors do add some information in explaining household housing leverage. For example, households with separated household heads have considerably higher leverage (11 percentage points) than other households. However, the fact that divorced household heads do not have higher leverage than other households suggests the impact of family breakdown on leverage may be only temporary; this possibility could be confirmed using the longitudinal aspects of the HILDA Survey to track households through the process of breakdown.

One indicator of household means, although not a causal factor, is that the results show that households whose homes are in poor condition have higher leverage than other households. This may be an example of reverse causation; households with high leverage may be too financially stretched to pay for renovations, and might be unable to borrow more to do so.

4.2.3 Household characteristics and attitudes

Besides the structural characteristics of households such as age and income, we also expect household attitudes to housing and debt to influence their housing leverage. For example, we may have expected that households that are more

---

7 The point estimates of other parameters are not sensitive to the inclusion of these subjective means variables, indicating that their possible endogeneity is not a major concern.

comfortable about taking on debt would have higher leverage than other households. However, a variable that summarised households' attitude to borrowing for items such as holidays, cars, and clothes, did not enter significantly into the model, although the point estimate did have the expected sign. Similarly, risk-averse households might be expected to prefer to pay housing debt off more rapidly and thus have lower leverage. However, our results indicated that households indicating a high aversion to risk in their saving decisions did not have significantly lower leverage on their homes, although the coefficient on this variable is again of the expected sign. This may suggest that these households' aversion to the riskiness of debt might be offset by a preference for housing assets over other assets they perceive as riskier; confirming this suspicion will not be possible until Wave 2 data on other kinds of assets and debt are available.

Other attitudinal variables do, however, enter the regression significantly. For example, households that pay their credit card off on time each month have lower leverage than households that do not. This could reflect either such households' preference for paying as little interest as possible on their debt, or perhaps greater financial sophistication and means. Households with long savings horizons have slightly higher leverage than households with shorter savings horizons.

The results also show that several other indicators of households' financial situation have significant associations with leverage outcomes. Although some of these indicators are symptomatic of the same causes as are influencing leverage, rather than themselves being causal factors, they provide some descriptive value in determining what kinds of households are most leveraged. For example, households who have owned more than one home – as opposed to first-home buyers – have slightly higher than average leverage, which suggests that households take on more debt when trading up, relative to the asset's value. Households that report some rental income, and therefore must own investment property, also have substantially higher leverage (14 percentage points) than other households. This result probably identifies a sub-group of the population that has actively engaged in leveraged asset accumulation, and is therefore willingly taking on the increased financial risks that this entails.

We also observe that households that took out loans from financial institutions at the time of purchase have higher leverage than the small proportion that did not.[8] Households that did not require a loan from a bank or other financial intermediary presumably acquired the property through inheritance, or used their own resources and bequests, gifts or loans from friends and family to fund their purchase. Loans from friends and family are included in the measure of leverage used here, but the other means of funding the purchase are clearly substitutes for debt that would reduce initial leverage at the time of purchase. Although some of these households might subsequently take out a loan secured against their home, in general they are likely to have little debt against their home.

People's cultural background may affect their leverage by influencing attitudes to debt, homeownership and intergenerational transfers; variations in homeownership rates amongst households of different ethnic origins have previously been observed in Australian data (Bourassa 1994, 1995). To test whether these differences reflect the migration experience or transmitted cultural values, we included variables representing the country of birth of both the household heads and their parents. Migrants will report both their own and their parents' birthplaces as being outside Australia, while second-generation Australians will report parental birthplaces outside Australia and their own birthplace as Australia.

We found that only the parental background variables were significant. Households with heads whose fathers were born in continental Europe had leverage 7 percentage points lower than households where the head's father was born in Australia. In contrast, households with heads whose fathers were born in the UK had higher leverage than other households.[9] Taken at face value, these results could be interpreted as indicating that a combination of cultural values and intergenerational transfers explains the pattern of lower than average leverage for the children of European migrants, rather than being a product of the migration experience.

8 Only 4 per cent of households who report having a mortgage did not take out an institutional loan at the time of purchase. However, over all owners, this increases to 28 per cent.

9 Using birthplace of mother or of either parent gave virtually identical results. The number of households where the heads or their parents were born in other regions was too small to produce statistically significant estimates. Variables representing birth in other regions were therefore excluded from our preferred specification.

4.2.4 Geographic variables

We also include a series of regional dummies to potentially capture two effects. First, they may capture different preferences for leverage across different regions, or differences in lending policies of banks. Second, they may proxy for different rates of housing price growth across regions, which can be expected to have exogenously influenced households' housing leverage. Although disaggregated data on growth in median house prices over the past two years resulted in the expected significant negative coefficient on its own, this was dominated by the inclusion of a suite of location dummies distinguishing the inner suburbs of the capital city, the outer suburbs of the capital city, and the non-metropolitan regions of each state.[10] Households move at different times and have thus experienced different degrees of inflation of the value of their home since they purchased it. Available data on Australian housing prices do not permit construction of regional level data on housing price growth over each individual household's holding period for their home, so housing price growth had to be calculated over a fixed window. The dummies may therefore be capturing variations in averages for both price growth and holding period, as well as other regional influences on leverage.

Using inner Melbourne as our base category because it is the region that has experienced the most rapid price growth in recent years, we can see that all other regions have higher leverage than inner Melbourne, and that the differences are broadly consistent with the recent pattern of relative growth rates for housing prices. In general, leverage is higher in non-metropolitan regions than metropolitan regions, and leverage is highest in Tasmania and Queensland. Brisbane's high average leverage is a puzzle, given that its price growth has been rapid in recent years, but this may be partly a base effect.

4.2.5 Effect of selection bias

The model with selection effects shown in Table 4 provides only a marginal improvement in fit over a model estimated over only those households that have leverage (adjusted R^2 of 0.278 versus 0.274). The coefficient on the tenure

---

[10] The housing price data used were a combination of regional-level median housing prices from Residex for the eastern seaboard states, and metropolitan and non-metropolitan dwelling prices compiled by the Commonwealth Bank for the HIA *Housing Report* for the other states.

selection term is not significant, while the coefficient on the mortgage selection term is significant only at the 11 per cent level and not at more conventional thresholds. Given the non-overlapping sets of explanators in the selection equations, it seems unlikely that identification problems drive this result, although multicollinearity could be an issue instead. It is more likely that the opposite effects that key variables, such as age and time the household has lived at the address, have on mortgage and tenure selection partially net out in the leverage equation. Since the selection terms combine information from both selection equations, these important variables may generate little net selection bias. Together with the bias arising from other variables in the model, this seems to result in only a minor degree of selection bias even though this is clearly a model with self-selection. The difference between the bias-corrected marginal effects and total effects are consistent with the mortgage equation dominating the tenure equation when variables counteract each other.

Despite the insignificance of the selection terms, the effects of the adjustment on coefficients for several key behavioural variables relating to life cycle and tenure characteristics are nonetheless large enough to conclude that this exercise enhances our understanding of population behaviour compared with the results implied by the non-bias adjusted model. Table 5 presents selected coefficients from the unadjusted model alongside the comparable marginal effects derived from the model results shown in Table 4. This enables a comparison with the model with selection effects and thus demonstrates where there are important differences.

The unadjusted model suggests a much smaller effect of age and time at address in diminishing the remaining loan size, and thus leverage, than the marginal effect implied by the model adjusted for selection bias. This unadjusted model also substantially understates the base effect of whether the household originally took out an institutional loan to help fund the purchase. Although neither specification generates an income effect that is significant at conventional levels, the point estimates are sufficiently different that economic interpretation might be affected.

In addition, our results suggest that this double-selection model is superior to a standard single-selection model with a Heckman correction. An alternative specification with only one selection equation, distinguishing mortgage-holders from other households and treating renters and homeowners without mortgages the same, produces different results that appear to fit the data less well than our preferred specification (these results are available from the authors). Although the measure of selection bias is not quite statistically significant, the results presented in this paper clearly show that such adjustments are useful when examining leverage across the whole household sector.

Table 5: Marginal Effects With and Without Selection Bias Correction

	With correction	Without correction	% difference
Household income	0.04	0.03	35.3
Age	–0.48	–0.36***	25.2
Has rental income	14.20***	14.10***	0.5
Satisfaction with pay	0.76***	0.75***	1.5
Took out institutional loan to buy home	12.60	–1.30***	108.1
Retired	–10.50	–4.50	57.0
Income adequacy	–0.70	–0.10	85.3
Married	3.16	2.30	27.1
Separated	11.00***	10.20***	7.1
Time lived at address	–0.75**	–0.44***	41.8
Father born in Europe	–6.88***	–6.90**	–0.34
Moves in last 10 years	0.84	0.99**	–19.13

Note: ***, **, and * represent significance at 1, 5 and 10 per cent levels.

5. Conclusions

The results presented here are descriptive, rather than being a fully causal model of household behaviour. Households' leverage at any point in time reflects the cumulation of many past decisions including the decision to purchase, to pay off the original mortgage or to make overpayments when it is feasible to do so. Cross-sectional data can be used to relate current leverage to observable characteristics, but it is not feasible to reconstruct all these past decisions.

The results suggest several descriptive conclusions about the pattern of housing leverage in Australia. The graphical results indicate that the households that are most highly leveraged are those most able to bear the debt – mid-life households with high income. Leverage is also higher for households living in areas least vulnerable to reversals in housing prices, the outer suburbs and non-metropolitan regions that have experienced relatively smaller price gains in recent years. Young homeowners are likely to have particularly high leverage, but young households in general are less likely to be homeowners. On the other hand, households that are negatively geared on investment property, and thus declaring a loss on their rental income, are much more likely to have a mortgage, and to have higher leverage when they do have a mortgage. This finding is indicative of a sub-group of the population that is willingly engaging in leveraged asset accumulation, and taking the associated financial risks. The general picture, however, accords with aggregate data in suggesting that leverage on the housing stock remains fairly moderate.

The econometric modelling also shows significant roles for age, life-cycle stage and time at address variables. This points to the importance of the passive paydown of debt as scheduled in the household's mortgage contract, in determining current leverage. Similarly, the pattern of coefficients on the locational variables suggests that increases in housing prices reduce leverage, and are not offset by households increasing their debt in response to the increase in wealth. These determinants of leverage can all be characterised as largely being beyond the control of individual households, suggesting that at least in the short run, households do not necessarily adjust their balance sheet to maintain a desired leverage ratio as predicted by some theoretical models. Against this, however, we find evidence of at least some households making explicit decisions about the end date of their mortgage, which might not be the date specified in the loan contract. As noted earlier, if some households did not make conscious decisions about the desired date on which their loan will be fully paid off, move-in date and payoff date would not both be significant in our estimated leverage equation.

Amongst the households making these explicit decisions about their payoff date, at least, we might expect deliberate portfolio reactions to developments in interest rates and housing prices, rather than simply adhering to a predetermined path for their remaining outstanding debt. Observing such reactions would, however, require tracking households through time. The longitudinal nature of the HILDA Survey will make it uniquely suited, amongst all datasets for Australia, to examination of these household responses. Future waves of the HILDA dataset will therefore be essential for further work on understanding households' decisions about their balance sheets, leverage and debt, and in particular their responses to interest rate changes and housing price movements through these channels.

Appendix A: Income Imputation and Results

As discussed in Section 2.2, the nature of the missing data leaves us with the need to impute income for three separate types of missing cases. For Type I individuals we impute total gross financial year income. For Type II individuals we impute gross financial year wage and salary income and add this imputed income to their reported gross financial year non-wage and salary income. For Type III individuals we impute gross financial year non-wage and salary income and add this imputed income to their reported gross financial year wage and salary income. Table A1 contains all the relevant results.

In all cases, missing values are imputed using the predictive mean matching (PMM) method outlined in Little (1988). In the first stage this involves estimating a regression on the variable to be imputed for individuals without missing values – in our case income. Next the model with the highest R^2 is used to predict the income of individuals with missing values. For every missing value we find the record with the nearest predicted value. The actual value of this 'donor' is then imputed for the missing value. The advantages of using the PMM method over other single imputation methods, such as simply imputing the conditional mean obtained from a regression, are that it ensures that only feasible values of the variable are imputed, and that a random error component is introduced so that imputed values have a similar variance to the reported values (ISER 2002).

Table A1: Income Imputation Results

I Total income ('000)		II Wage income ('000)		III Non-wage income ('000)	
Age	1.6***	Age	0.8***	Age	0.08***
Age squared	–0.15***	Age squared	–0.01***		
VIC	–0.3	VIC	–1.2**	VIC	0.6
QLD	–1.4*	QLD	0.0	QLD	0.0
SA	–2.4**	SA	–1.2*	SA	–0.8
WA	–0.9	WA	–2.4***	WA	1.3**
ACT	6.3***	ACT	4.8***	ACT	–1.5
Make ends meet	3.9***	Make ends meet	2.1***	Make ends meet	0.9***
Socio-economic	1.0***	Socio-economic	0.1***	Socio-economic	
Has disability	2.3***	Business income	–16.0***	Business income	20.9***
Lone person	7.9***	Govt benefit	–5.1***	Govt benefit	2.6***
Group household	6.7***	Receives interest	1.2***	Receives interest	4.1***
Sole parent, dependant children	6.8***	Receives rent	4.4***	Receives rent	4.5***
		Receives dividends	2.5***	Receives dividends	1.0**
Sole parent, no dependant children	4.4**	Non-metropolitan	–1.9***	Age pensioner	–2.3***
		Inner-city	1.5***	Receives royalties	3.9
Persons in h'hold	–1.4***	Union member	5.9***	Union member	–1.9***
Employed	13.0***	Employed	11.6***	Employed	–4.2***
Retired	–7.6***	Retired	–10.6***	Retired	4.3***
Home duties	–6.0***	Spouse's income	0.0***	Spouse's income	0.1***
Multifamily home	–3.7*	Multifamily home	–3.6**	Student	–1.7
		Household head	9.9***	Household head	6.2***
No of bedrooms	0.8***	Has disability	1.5***	Health	0.2
Home's condition	–1.1***	Home's condition	–0.6***	Home's condition	–0.3
Home value	0.03***	Home value	0.004***	Home value	0.004***
No of children	1.4***			No of children	0.3*
Married	8.1***	Education level 2	–10.9***	Never married	3.9***
Separated	4.7***	Education level 3	–11.1***	Separated	3.9***
De facto	10.0***	Education level 4	–11.6***	De facto	2.3***
Divorced	4.9***	Education level 5	–10.7***	Divorced	4.8***
Widowed	8.5***	Widowed	4.4*	Widowed	1.8*
Adjusted R^2	0.32	Adjusted R^2	0.46	Adjusted R^2	0.204
RMSE	\$26 000	RMSE	\$19 000	RMSE	\$20 500

Note: ***, ** and * represent significants at 1, 5 and 10 per cent levels.

References

Bernanke B and M Gertler (1995), 'Inside the black box: the credit channel of monetary policy transmission', NBER Working Paper No 5146.

Bourassa SC (1994), 'Immigration and housing tenure choice in Australia', *Journal of Housing Research*, 5(1), pp 117–137.

Bourassa SC (1995), 'A model of housing tenure choice in Australia', *Journal of Urban Economics*, 37(2), pp 161–175.

Carroll CD and WE Dunn (1997), 'Unemployment expectations, jumping (S,s) triggers, and household balance sheets', NBER Working Paper No 6081.

Fishe RPH, RP Trost and P Lurie (1981), 'Labor force earnings and college choice of young women: an examination of selectivity bias and comparative advantage', *Economics of Education Review*, 1(2), pp 169–191.

Genesove D and CJ Mayer (1997), 'Equity and time to sale in the real estate market', *American Economic Review*, 87(3), pp 255–269.

Goodman J and J Ittner (1993), 'The accuracy of home owners' estimates of house value', Board of Governors of the Federal Reserve System Working Paper No 131.

Gronau R (1974), 'Wage comparisons – a selectivity bias', *Journal of Political Economy*, 82(6), pp 1119–1143.

Guiso L, M Haliassos and T Jappelli (eds) (2002), *Household Portfolios*, MIT Press, Cambridge.

Heckman JJ (1976), 'The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models', *Annals of Economic and Social Measurement*, 5(4), pp 475–492.

Heckman JJ (1979), 'Sample selection as a specification error', *Econometrica*, 47(1), pp 153–161.

Henley A (1999), 'The economics of the crazy British housing market', Aberystwyth University Economic Research Paper No 99/8.

Institute for Social and Economic Research (ISER) (2002), 'Weighting, imputation and sampling errors', Chapter 5 in British Household Panel Survey (BHPS) User Documentation.

Kent C and P Lowe (1997), 'Asset-price bubbles and monetary policy', Reserve Bank of Australia Research Discussion Paper 9709.

Lee LF and GS Maddala (1985), 'Sequential selection rules and selectivity in discrete choice econometric models', in GS Maddala (ed), *Econometric Methods and Applications Volume II*, Edward Elgar Publishing Limited, Aldershot, pp 311–329.

Lewis HG (1974), 'Comments on selectivity biases in wage comparisons', *Journal of Political Economy*, 82(6), pp 1145–1155.

Little R (1988), 'Missing-data adjustments in large surveys', *Journal of Business and Economic Statistics*, 6(3), pp 287–301.

Maddala GS (1986), *Limited-dependent and Qualitative Variables in Econometrics*, Econometric Society Monographs No 3, Cambridge University Press, Sydney.

McLennan D, J Muellbauer and M Stephens (1999), 'Asymmetries in housing and financial market institutions and EMU', Centre for Economic Policy Research Discussion Paper No 2062.

Schwartz AJ (2002), 'Asset price inflation and monetary policy', NBER Working Paper No 9321.

Smith J, G Sterne and M Devereux (1994), 'Personal and corporate sector debt', *Bank of England Quarterly Bulletin*, 34(2), pp 144–155.

Tunali I (1986), ‘A general structure for models of double-selection and an application to a joint migration/earnings process with remigration’, in R Ehrenberg (ed), *Research in Labor Economics*, 8(B), JAI Press, Greenwich, pp 235–282.

Watson N and M Wooden (2003), ‘Toward an imputation strategy for Wave 1 of the HILDA Survey’, HILDA Project Discussion Paper Series No 1/03.