

Overview and early analysis of the new Wealth and Housing Asset Module (WHAM)

Natasha Bradshaw*
Jonathan Hambur**
Martin McCarthy**
Simon Nash*

ABS/RBA Conference

June 2025

*Macroeconomic Analysis and Policy Division, Australian Treasury

**Economic Research Department, Reserve Bank of Australia

The authors thank members of the Technical Working Group (including Eric Morris, Justin Holland, Oscar Lane, Tristram Sainsbury, Will Mackey, Joseph Manalo and Rachel Price) and the Project Board from Treasury, DISR, DSS, the RBA, the ATO and the ABS, participants at the 2025 ABS-RBA joint conference on housing, as well as others from the ABS for their input and feedback to the paper. The views expressed are those of the authors, and not the RBA, Treasury or the Australian Government.

Abstract

This paper introduces the Wealth and Housing Assets Module (WHAM) pilot, which will integrate data on housing, superannuation and firm ownership into the ABS's secure microdata ecosystem. These data represent a significant uplift in Australia's data landscape and will enable key insights into housing supply and affordability, inequality and entrenched disadvantage, entrepreneurship and competition, and the targeting of Australia's income support and aged-care systems. We present two case early case studies to highlight the potential value of the data, decomposing recent changes in household size by housing tenure, and measuring income dynamics of recent homebuyers. In the longer term, the vision is to expand the module over several phases to improve its coverage and scope.

1. Introduction

Some of the most significant policy challenges facing Australia revolve around housing, wealth, and business ownership, and their interactions with the tax and transfer system. This includes housing supply and affordability, inequality and entrenched disadvantage, entrepreneurship and competition, and the targeting of Australia's income support and aged-care systems. Addressing these issues has the potential to significantly improve macroeconomic outcomes, and Australians' welfare.

These policy challenges are complex. As such, having a good evidence base is valuable in trying to develop and assess policy responses. However, there is a lack of detailed and integrated data on housing and wealth available in Australia that can be used to build this evidence base. While some surveys exist, limitations in coverage, frequency and scope constrain their ability to answer many key policy questions, leaving a large and crucial gap in the data landscape that limits our ability to inform responses to these key policy challenges.

The Wealth and Housing Asset Module (WHAM) looks to fill this gap by integrating information on housing and locations, wealth, and business ownership into the Australian Bureau of Statistics' (ABS) secure anonymised microdata ecosystem. This represents a significant uplift in the data landscape in Australia, and will enable a range of new granular insights into high-value policy issues.

As a first step, the Treasury, Department of Social Services (DSS), Reserve Bank of Australia (RBA), Department of Industry, Science and Resources (DISR), Australian Tax Office (ATO), and the ABS are collaborating on a WHAM pilot.¹ The pilot will bring in additional information on the asset-holdings of Australians, particularly housing, superannuation and business ownership. This will facilitate new insights into wealth, housing markets, and business dynamism.

The pilot is also building the foundation for further capabilities, particularly by accelerating work in the ABS to improve location data. This will increase the value of already integrated housing, individual and business data. And it also has the potential to enable other types of geospatial data to be integrated, such as state and local government data on zoning and lots, or environmental data, opening up the potential for further policy-relevant research and insights.

The pilot represents a significant first step towards the production of a fuller WHAM, and in and of itself represents a significant uplift to Australian administrative microdata, bringing it close to the global frontier. As such, in addition to facilitating a range of new insights into key policy challenges, it has the potential to attract additional research on the Australian economy evaluation of Australian public policy.

In the longer term, the vision is to expand the pilot over several phases to improve its coverage and scope. This could include: data from state and local authorities that would facilitate new and powerful insights into housing policies, affordability and housing market dynamics; information from not-for-profit social housing and homelessness services that would allow for additional analysis of homelessness and entrenched disadvantage; and information on debt and insurance from various sources that could help to understand household vulnerability.

¹ The pilot is jointly funded by Treasury, DSS, RBA and DISR who are also providing advice and 'in-kind' support. The ABS is sourcing and linking the data. The ATO are assisting with data provision and providing technical expertise.

This paper provides an overview of WHAM and its pilot. It begins by discussing the current data landscape and gaps. It then gives information on the WHAM framework and pilot datasets, and how they can be used to answer key policy questions, before discussing two initial case studies that demonstrate this value. The paper concludes by discussing the way forward for the pilot as well as the broader WHAM project, highlighting high-value directions for future investment.

2. There are significant gaps in existing wealth, ownership and housing data in Australia

Current wealth, housing and business ownership data available in Australia provide significant value but have limitations that prevent researchers from answering key policy questions. The WHAM pilot will significantly build on this capability and bring Australian wealth microdata close to the global frontier.

2.1 Improvements to the current wealth data landscape

There are three main sources of wealth data in Australia: The ABS's Survey of Income and Housing (SIH), the Household, Income and Labour Dynamics in Australia (HILDA) survey, and the Australian System of National Accounts (ASNA).

The ASNA data provide quarterly macroeconomic aggregates of household assets and liabilities, providing a comprehensive view of overall wealth and housing trends at the national level. However, it lacks the demographic detail required for distributional analysis, and is an aggregated dataset that cannot be used to understand individual or household dynamics.

HILDA and SIH provide individual-level data but with some important limitations. First, the ability to investigate wealth and housing dynamics is limited by the cross-sectional nature of SIH and the infrequent collection of wealth data in HILDA. The SIH provides biennial cross-sectional data, enabling detailed snapshots of wealth across the distribution and for different demographic groups. However, the cross-sectional nature limits the ability to use the data for some purposes. HILDA is an annual longitudinal survey, but only collects information on wealth every four years. This enables the tracking of households over time to understand trends in wealth accumulation and mobility over the life course, but on a relatively infrequent basis. A lot can change for individuals and households over four years, which makes it difficult for researchers to draw conclusions on responses to cyclical factors or sudden shocks.

Second, the relatively small sample sizes in HILDA and SIH limit the ability of researchers to conduct some types of detailed analysis. For example, it can be difficult to look at outcomes for small population groups, such as First Nations Australians or first-home buyers. It is also difficult to analyse events that occur relatively infrequently across the population, such as property transactions, starting a business, retiring, or moving house. While infrequent, these significant life events can have large macroeconomic implications, and matter for individuals' financial outcomes and wellbeing. Similarly, policies, such as those relating to superannuation or housing, may have large implications for individuals and the economy but only impact a small number of Australians in any given year.

Third, HILDA and SIH are subject to survey bias and attrition, and may suffer from misreporting. Both HILDA and SIH exclude some cohorts of Australians, in particular rough sleepers and those

living in sparsely populated areas.² Household surveys also encounter difficulty surveying very wealthy individuals, and under-survey the most disadvantaged, with implications for our understanding of wealth inequality. Both HILDA and SIH also rely on self-reported information, which can be subject to recall bias or mis-valuation, particularly for complex or infrequently valued assets such as property or shares. In HILDA, non-response rates for wealth data are also higher than for most other questions (Melbourne Institute, 2023).

WHAM can help address these issues. It will provide annual longitudinal information on wealth, should have near universal coverage of Australians, and, as administrative data, should be less prone to misreporting.

However, it is important to note that administrative data, while extremely useful, are not a panacea. Administrative data are typically collected for program administration rather than research purposes, which has implications for data quality and coverage, and can introduce its own set of biases. For example, data from income tax returns are commonly used to measure individuals' income, but many Australians are not required to lodge a tax return because they earn below the tax-free threshold, and some people may not report their income accurately. Similarly, coverage could be poorer for some geographic areas, like regional or remote communities, or certain groups who are less likely to have interacted with government services (Productivity Commission, 2024).

Some of these issues can be resolved by researchers linking multiple data sources, including payment summary and single touch payroll data that employers report to the ATO, and reported income and payments from the Department of Social Services' DOMINO dataset. They can also potentially be addressed by further investments in WHAM.

2.2 Improvements to the housing market data landscape

There are currently various datasets providing specific information on housing market transactions and dynamics. These include datasets of housing transactions, rental agreements, and derived measures of prices and rents, often provided by commercial institutions.

While these sources are highly useful, they have limitations which prevent researchers from answering several key policy questions. A particular shortcoming is that they tend not to be integrated with other relevant information that is needed to understand and assess housing policies. For example, data on housing transactions is generally not integrated into information on who was transacting. This limits the ability to use such data to understand the effect of policies such as first-homebuyer grants, or understand the dynamics of how housing filters between different types of households, which has important implications for the targeting of housing supply policies.³ Similarly, information on rents is not integrated with income data, limiting the ability to assess rental stress over time.

² In HILDA, attrition rates are also higher for some cohorts, including those born in a non-English speaking country, of Aboriginal and Torres Strait Island descent, single, unemployed, or working in low skilled occupations. While weighting the sample does go some way in controlling for these issues, any adjustments are limited by the ability to measure differential attrition: Melbourne Institute, 2023.

³ An exception is [Hansen and Rambaldi \(2022\)](#).

2.3 Improvements in our understanding of business dynamism and ownership

Data on business ownership and entrepreneurship is currently very limited. While HILDA has some information on which individuals run businesses, this is subject to many of the limitations discussed above. Other data linking individuals' circumstances to business decisions are very limited.

This is a very large gap in the data landscape. It limits the ability to understand changes in the nature and drivers of entrepreneurship and business dynamism, including the interaction between housing and housing prices and these decisions. These dynamics have very important implications for productivity. It also limits our ability to understand certain competition dynamics, such as common ownership, which have been shown to be important overseas (see, for example, Azar 2018). The WHAM pilot can begin to fill this gap.

3. The WHAM pilot will fill data gaps, facilitating new insights into high-value policy questions

The WHAM pilot will integrate new administrative data on housing, superannuation, firm ownership (including dividend, trust and partnership payments) into the ABS's existing Person-Level Integrated Data Asset (PLIDA) and Business Longitudinal Analysis Data Environment (BLADE). The former covers anonymised information on health, education, government payments, income and taxation, employment, migration, and demographics. The latter contains anonymised business taxation, trade and other information.

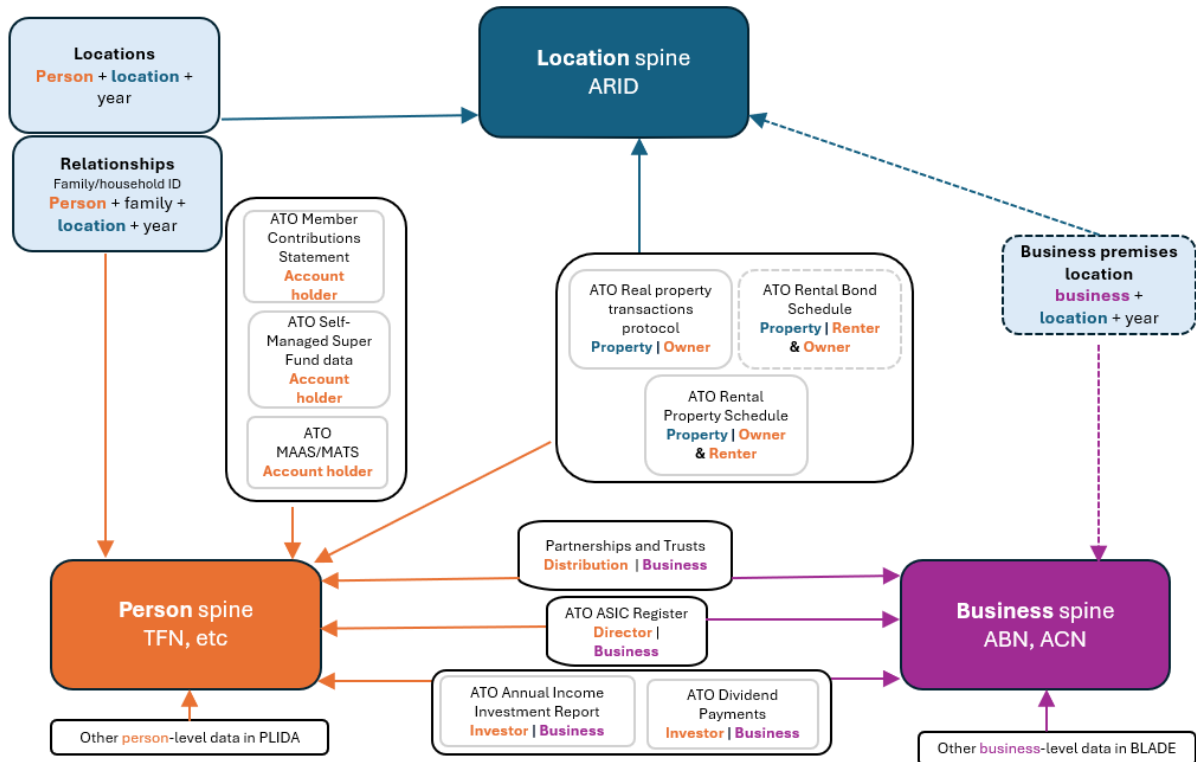
By combining these datasets, we will begin to fill some of the existing gaps in the Australia data landscape, facilitating important policy-relevant research. Moreover, these additional datasets will enrich and improve existing data. For example, data on housing transactions can be used to improve locational information, and family linkages.

As the new data are being brought into the existing environments, they will have the same strong privacy measures in place as for PLIDA and BLADE. These include safeguards around use being restricted to approved and experienced researchers, for approved purposes, in a secure environment with no identifiable information being able to be extracted.⁴

Figure 1 outlines the new data assets the make up the WHAM pilot and their linkages. Broadly these datasets can be thought of as three new modules: housing; superannuation; and firm ownership and other assets. These modules are linked through the person, business and location spines. These modules cover the majority of Australians' asset holdings, which are dominated by land and dwellings (56 per cent) and superannuation (21 per cent).⁵ Shares and other equity make up 8 per cent, and currency and deposits 9 per cent. Remaining financial and non-financial assets make up the remaining 7 per cent.

The following sections outline the new data in the WHAM pilot, and highlight key questions that they can be used to address. It also discusses investments being made in location data which will increase the value of these new datasets, as well as opening new, high-value future directions.

⁴ For more information see the [ABS website](#).

Figure 1

3.1 Housing

Improving housing affordability is a major policy challenge for Australian governments, but researchers' ability to generate insights on the housing market and understand the implications of policy have been limited by data availability. As shown in Figure 1, the housing component of the WHAM pilot will bring two key data sources into the existing environment:

- **Rental property schedules:** Includes information on income, expenses and location for individual rental properties reported on income tax returns. These data can be linked to both the owner of the property through their person-level identifier, and to the renting household, through the location identifier. This covers the 2015-2023 financial years.⁶
- **Property transaction data:** Provides information on who is transacting a property, and for how much. These data will be linked to both the transactors of the property and the address. This will provide information on the ownership of properties, as well as their value. When combined with other location information, it also enables a measure of whether the house is owner-occupied. The pilot will integrate data from 2017 to 2024. Future work could be done to work with states and territories to integrate more historical data (see final section).

The combination of these data will allow for researchers to understand which properties are rented and which are owned, as well as housing dynamics around transactions and rents. This will allow examination of a number of important policy areas and challenges.

⁶ Rental bond schedule data are also being considered, which would provide information directly linking landlords and rents, and cover details such as weekly rent and some information on the dwelling.

First, ownership data, when combined with information on housing prices, will be a key component of measuring household wealth. This in turn will facilitate research into wealth inequality, static and over time, and possibly even over generations. It will also allow for analysis to be undertaken exploring interaction between housing wealth and the tax and transfer system.

The ability to understand changes in ownership will also allow for detailed examination of housing market dynamics, and how they are influenced by policies. For example, by analysing outcomes for first home buyers, versus other individuals, researchers could explore the effects of first homebuyer grants and similar policies. Similarly, changes in land tax and stamp duty that affected some groups, like investors, differently could be assessed, as could macroprudential policies such as investor lending caps introduced by APRA during the 2010s. Relatedly, the data could be used to assess whether houses tend to filter down from higher income to lower income households, or not. This has implications for whether housing supply policies need to target certain housing types.

Information on renters and their rents, when combined with data on their income from PLIDA, will provide scope to better understand rental stress. Moreover, when combined with data on Commonwealth Rent Assistance from PLIDA it can be used to consider how the benefits of these policies accrue to both the renter and the landlord, which has important implications for the design of such policies.

These data also have the potential to answer important questions regarding household and financial stability. For example, they would allow for additional insights into the extent to which investors can amplify housing cycles, which is often put forward as a reason for the use of macroprudential policies targeting investors. Moreover, they could help researchers and policymakers understand what happens to investors' ability to repay debts when they lose a renter.

More generally, these data will also enable an enhanced understanding of the interaction between housing and households' major life decisions and responses to shocks. This includes the drivers of changes to household size and housing demand, whether housing wealth provides a buffer or impacts responses to negative shocks (including employment, interest rates, health, and separations), and how housing influences decisions around relocation, entrepreneurship, education, fertility, retirement, and marriage or separation.

Importantly, the wide coverage of these data may enable insights into the experiences of cohorts that are typically more difficult to capture in survey data, including First Nations Australians, migrants, and people living in remote or sparsely populated areas. The availability of annual data will also enable analysis of changes over time, and experiences of entrenched disadvantage.

While these data will add significant value, limitations remain that future advancements in WHAM may seek to address, particularly relating to property transactions prior to 2017, and current property prices.

3.2 Superannuation

Australia's superannuation system is significant, complex, and interacts heavily with other government policies around aged care, income support, and taxation. The inclusion of

superannuation data in WHAM will enable deeper understanding of the superannuation system and these interactions.

The superannuation components of WHAM include:

- ATO self-managed super fund data: Detailed information on incomes, expenses, assets and distributions for SMSFs. This currently covers the 2014-2024 financial years, but there is potential for a longer-term series.
- ATO member contributions statement: For each individual in each financial year, this contains information on employer and individual contribution amounts, account balances and other details. This covers 2013-2018 financial years, with potential for a longer time series.
- ATO account attribute service: Similar to the member contributions statement but at the individual and account level, with reporting from 2019 onwards.

These data will enable significant new insights into the superannuation system over time. For example, it can be used to better understand how rates of accumulation differ for different groups in Australia, patterns of drawdown and how these have changed over time. It will also allow for deeper assessment of the interactions between superannuation, other wealth, and the aged pension. Finally, these enable a more thorough evaluation of existing and proposed superannuation policies, including the first home super saver scheme, the COVID early super release program, and other taxation settings.

3.3 Firm ownership and other assets

Data on ownership of businesses, both controlling and investments, is currently very limited. Moreover, in many cases wealth is held through partnerships and trusts, meaning that it is crucial to be able to link them to individuals to get a proper picture of wealth. To fill this gap in the data landscape the WHAM pilot will bring in:

- ATO ASIC register: Allows linkage of the business to the director. For small companies, the directors and owners are often the same, so this data can inform estimates of business ownership.
- ATO annual income investment report (AIIR) dataset and dividend schedule: Provides information on who firms have provided dividend payments to, giving a picture of ownership of equity.
- ATO partnerships and trusts distribution: Tax return information on the distributions of partnerships and trusts, allowing for the linkage of the entity to its beneficial owner.

As noted above, these data will help to build up a more complete picture of household wealth, as they will allow researchers to link wealth held via companies and other entities to the owner. This can also provide a better picture of who ultimately benefits from business support payments.

These data will also provide insights relevant for competition policy and understanding business dynamism, which have important implications for productivity growth. For example, these data can provide information on interlinked directorships and common ownership of firms, which may have

implications for competition. Moreover, by enabling researchers to link other information on the wealth and income of individuals to the businesses they own, it can help provide insights into the relationship between wealth, particularly housing wealth, and firm decisions and credit constraints.

3.4 Location Spine

As evident in Figure 1, having a high-quality location spine is extremely important in bringing together many of the elements of WHAM, particularly those related to housing. For example, it is needed to identify whether houses are rented or owner-occupied, which in turn is important for many of the potential use cases for WHAM. It is also valuable in creating measures of households and household size, and therefore work that would look to assess overcrowding or housing underutilisation. And it can potentially significantly improve our understanding of business locations.

The WHAM Pilot is accelerating ABS work to continue to improve and develop a Location Spine. This will make the existing data even more powerful and will have the potential to unlock new integrations with more complex geographical information, which may not have a simple address, such as environmental data and lots. This could facilitate additional work to better understand issues ranging from housing and zoning, to environmental issues and natural disasters.

4. Case studies of how WHAM can bring policy insights

Here we document two early case studies that highlight the potential value of the WHAM pilot, and WHAM project more generally. In the first, we decompose recent changes in average household size based on housing tenure, which was not previously possible. In the second, we construct a proxy for the year that owner-occupiers purchased their house and use this to understand the income dynamics of those households that bought just before the recent monetary policy hiking cycle.

4.1 Average household size by housing tenure

Housing prices, availability and affordability reflect the combination of housing supply and housing demand. As such, understanding developments in housing demand is crucial to understanding housing markets.

A key housing market outcome is average household size (AHS), which is influenced by supply and demand. For example, on the supply-side, a loosening of planning restrictions on apartment construction would likely lead to more small dwellings, reducing AHS. On the demand-side, if Australians wanted to live in smaller households, this would reduce AHS. As discussed in Hunter (2024) and Agarwal, Bishop and Day (2023), AHS declined sharply following COVID, likely reflecting a combination of shifting preferences for space and work from home. This put upwards pressure on dwelling prices and rents.

Currently there are relatively few sources of data that allow us to understand changes in AHS. The Census provides a detailed snapshot of AHS but is only released every 5 years. The RBA has recently developed a monthly AHS measure using the microdata underlying the Labour Force Survey (LFS) (Agarwal, Bishop and Day 2023). This is timely and high-frequency, so can be used to assess higher-frequency dynamics. A similar measure is used by the e61 institute, who have demonstrated the role of decreasing fertility in reducing average household size since the 1980s (Gibbons, 2025).

Treasury have recently developed an alternative measure of AHS using the LFS that facilitates more detailed analysis of drivers in the change in the number of households and average household size, further demonstrating the importance of fewer children per household in the apparent fall in AHS (Burleigh and Nizelnik, forthcoming).

While the LFS measures are extremely powerful, the data have some shortcomings. The LFS collects limited other information, making it hard to do certain types of analysis. For example, the LFS does not provide detail on whether houses are rented or owned. Understanding whether changes are driven by renters or owners is useful because it can help us to understand what markets the additional demand is likely to appear in, particularly given the rental housing stock and owner-occupier housing stock can differ. The LFS also has very limited details on income and income support, which may be relevant in understanding what drives changes in AHS, and the effects on vulnerable people.

Moreover, the LFS measure has a small sample size, limiting the ability to look at smaller groups or geographies. This could make it harder to understand issues such as overcrowding amongst vulnerable groups, or exploit variation in local conditions to tease apart causal determinants of changes in AHS. And finally, the LFS lack a panel dimension,⁷ making it harder to understand how adjustments occurred. For example, to what extent did people achieve smaller AHS by moving geographically to areas with more vacant housing?

WHAM has the potential to overcome a number of these weaknesses. PLIDA has rich data on individuals. The new data in the WHAM pilot will provide important information on ownership and rental status. And ongoing investments in the location spine and family linkages will enable high-quality granular estimates of AHS.

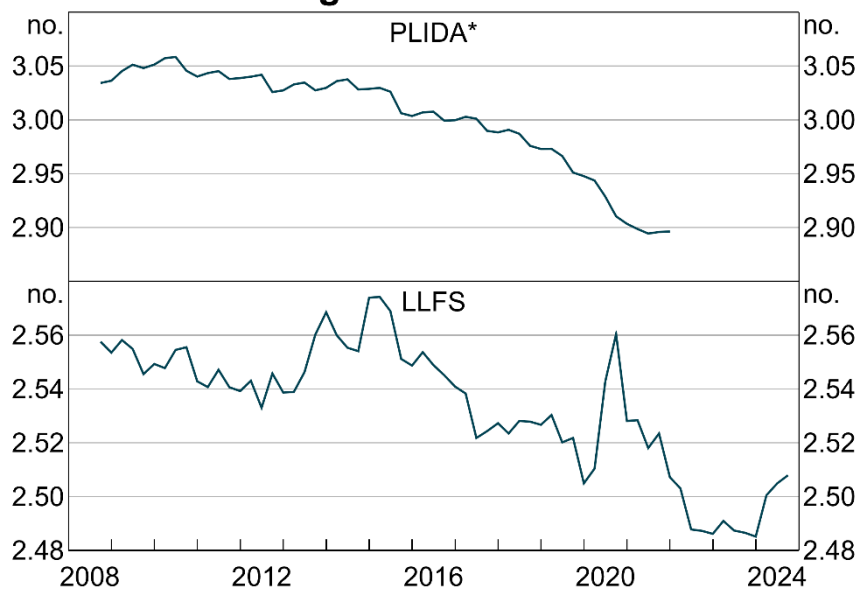
To demonstrate, Figure 2 shows a measure of AHS constructed using the existing location data in PLIDA.⁸ While it is slightly higher than the LFS (and Census) measure it follows a similar, if smoother, path. It increases slightly in the mid-2010s, before trending down in the late 2010s, similar to the LFS. While we do not see the same spike during the peak of COVID, the PLIDA measure shows a similar decline post COVID.⁹ Overall it appears that the PLIDA measure does a good job of tracking changes in AHS over time. Ongoing investment in location data, as being undertaken as part of WHAM, and work to improve methodologies, will likely enable further improvements to the measure.

⁷ The Labour Force Survey uses a rotation method, where each month one-eighth of the sample rotate out and are replaced by a new group. This means individuals remain in the sample for eight months.

⁸ More precisely, the measure estimates the average number of adults, as children who are too young to interact with the tax or transfer system are generally excluded. Further work could be done to bring them in.

⁹ It's not surprising that the PLIDA measures does not pick up the short-term spike during COVID. The spike likely reflected a number of people temporarily moving in with family or friends for caring and other reasons. Individuals may not have changed their registered addresses for tax and social services during this time. This is a fairly minor limitation, as such dynamics should be very rare.

Figure 2
Average Household Size



* Average adults per household. Excludes households with over 10 adults.

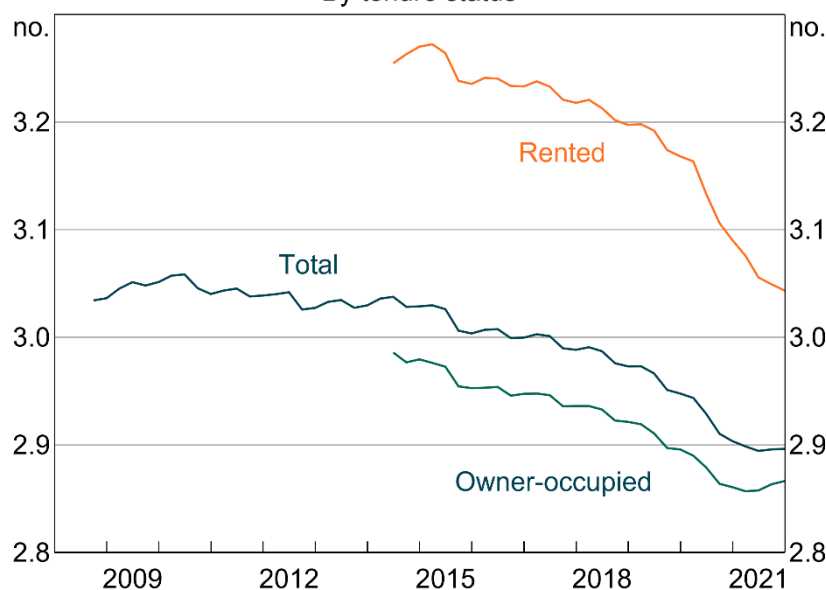
Sources: ABS; Authors' calculations.

We now decompose the PLIDA measure into rented and owned dwellings. To do so we make use of the Rental Property Schedules from 2015 to 2022 and decompose the housing stock into owner-occupied housing, and rental housing (see Appendix A for more details).

Looking at the measures, we can see that renter dwellings' AHS tends to be higher, around 3.2 in 2019, compared to a little over 2.9 for owner-occupied dwellings. Second, we can see that average household size did decline for owner-occupiers but the decline amongst renters was much sharper. This is not surprising, given renters likely have more flexibility to adjust their housing choices than owners. It may also reflect other compositional factors, such as the net outflow of foreign students if they tend to have higher AHS. Further work could assess the drivers in more detail.

As well as the specific use for understanding AHS, this case study demonstrates how WHAM can be used to identify owner-occupied and rented housing. This could facilitate policy-relevant analysis. For example, it could be used to assess to whether rental and owner-occupier markets appear to be segmented, or whether dwellings switch between uses. This has important implications for how different forms of housing supply would influence the housing market. It can also be used to construct local-level measures of renter and owner-occupier activity, to understand the effects of macroprudential policies. And it represents the first step in constructing measures of housing wealth.

Figure 3
Average Household Size*
 By tenure status



* Average adults per household. Excludes households with over 10 adults.

Sources: ABS; Authors' calculations.

4.2 Income growth by time of property purchase

One area of concern for the RBA and other agencies over the recent interest rate hiking cycle has been whether the sharp increase in interest rates, alongside high inflation, could make it difficult for borrowers to meet their mortgage obligations. This concern would be particularly acute for those who purchased a home just before rates rose, and so had large, unseasoned loans (see, for example, RBA 2024). If these households did have difficulty meeting their obligations, it would have implications for the macroeconomy and financial system, as well as for the individuals themselves.

To monitor these issues, the RBA undertakes household 'stress testing'. This makes use of the RBA's Securitisation Database. While this database has detailed information on payments, buffers and other relevant details, one shortcoming is that it only has information on incomes as at origination of the loan. As such, this stress testing tends to use aggregate measures of income gains, such as the Wage Price Index, to model income growth.¹⁰

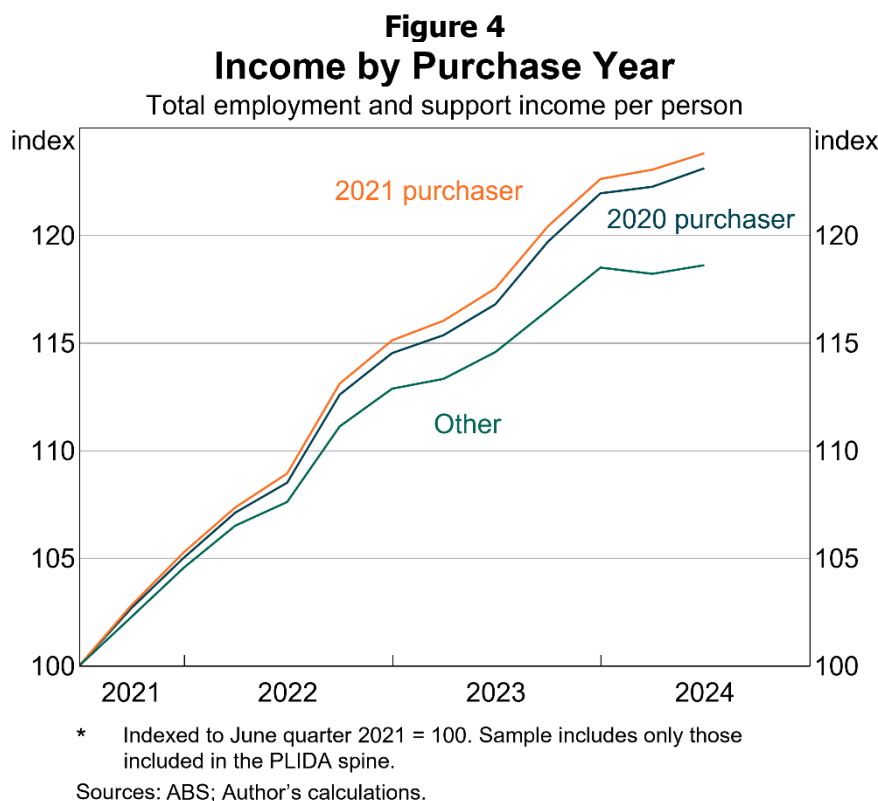
One concern with this approach is that income growth may differ substantially for different household types. For example, households who have recently bought a house, and so who likely have the largest loans and sensitivity to rising rates, might have stronger income growth, particularly if they made the purchase in part because they expected their income to rise strongly. Using an average income growth measure may therefore lead to an overstatement of risk.

To understand whether recent purchasers have had different income growth we use recently integrated rental property schedules to identify what look like owner-occupier purchases in 2019/20 an 2020/21 financial years (see Appendix B for further detail).

¹⁰ See Hughes (2024) for a detailed discussion.

This indicator of purchase is integrated with a measure of total income for the individual in each fortnight. This is constructed by combining Single Touch Payroll employment incomes and DOMINO income support payments.

The key takeaway is that households that purchased in 2019/20 and 2020/21 appear to have had stronger income growth than other households. Overall, their income growth over the period was around 5 per cent higher than others. This suggests that using aggregate measures of income growth could give a slightly conservative estimate of household stress. Still, these results should be taken with some caution as they are preliminary and further work and data will allow for further refinements.



This preliminary work demonstrates the types of household stress-related analysis that could be done. For example, work could be done to isolate first-home buyers, who may tend to be more financially fragile, or to examine investor's income growth, and how investors are affected by income shocks to their tenants.

More generally, this demonstrates how the WHAM pilot data can be used to identify changes in ownership. As discussed above, this could be used to think about how housing filters up and down the income ladder, how housing policies that affect certain groups (e.g. stamp duty changes and first homebuyer policies) influence purchases, and how different types of buyers contribute to housing market cycles.

5. The WHAM pilot significantly improves our ability to answer policy questions, but there is still work to do

WHAM represents a significant uplift in our ability to answer important policy questions, filling significant data gaps that previously prevented researchers from providing insights into key policy

challenges. Moreover, it will take Australia closer to the global frontier for administrative microdata, and has the potential to attract the best and brightest from Australia and around the world to try to answer Australian policy questions.

The WHAM pilot represents the first big step in this journey. By incorporating data on housing, superannuation, and firm ownership it will give researchers the ability to construct measures of wealth, and better understand housing markets, firm behaviour and competition, and the interactions between the three.

The WHAM pilot is ongoing, but it is envisioned that the data will start being made available for approved researchers from the second half of 2025. To facilitate effective use of these data, work is also being scoped to publish papers documenting the data, their coverage and limitations, and how they can be used to construct various concepts of interest, such as measures of wealth or the allocation of dwellings between owner-occupier and rentals. In time, some of these concepts may be formalised by the ABS to save researchers the time and effort.

While highly valuable, the pilot is only the first step. There remains significant scope to expand WHAM to deliver substantial additional value to research and policy community.

One high-value direction is to work with states to incorporate housing transaction and other data collected at a state level. This would, for example, allow for information on land values and rates, lots, as well as housing transactions pre-2017, to be incorporated. This would significantly improve WHAM, allowing for more effective analysis of earlier housing cycles and policies, and more effective measurement of wealth over time. As well as providing benefit to the broader policy community, this would provide direct value to the states themselves, allowing them to better leverage their own data in designing and advising on policy.

Work could be also done to partner with not-for profits or for-profit organisations. For example, incorporating data from social housing and homelessness services could facilitate a richer understanding of how people move in and out of homelessness. Data from insurance providers could provide deeper insights into the extent of under-insurance and its drivers, particularly in vulnerable areas.

Another key area for uplift could be additional investments measuring family links. This could take the form of joint work by the ABS and ATO to leverage work done in the ABS's PLIDA and ATO's Alife to create a best practice measure of family links. This would help to refine existing methods, and open the scope to analyse additional questions around intergenerational wealth inequality, bequests, and the role of the 'bank of mum and dad' in housing markets.

Finally, further investments in location spines will also facilitate further linkage opportunities that will enable research into other important questions. This could include environmental information that facilitates analysis of vulnerability to climate change, as well as state and local council information that helps examine the effects of zoning policies and land development.

Overall, WHAM and the WHAM pilot have the potential to deliver significant value to the Australian people, filling crucial data gaps and facilitating policy-relevant research into significant challenges

facing the nation. By continuing to invest in these data, and bringing researchers together from across government, think-tanks and academia, we can realise this potential.

Appendix A: Classifying the housing stock

To decompose the housing stock into owner-occupied and rented, we make use of the Rental Property Schedules from 2015 to 2022. We use the following algorithm to identify whether a dwelling was rented in a given month:

- If the property was rented for 52 weeks, we assume it was rented in all months.¹¹
- If it was rented for fewer than 52 weeks, and there is a rental schedule in year t and $t-1$, we assume the rented weeks occur at the start of the tax year. For example, if it's rented for 26 weeks, we assume it was rented from July to December.
- If it was rented for fewer than 52 weeks, and there is a rental schedule in year t and $t+1$, we assume those rented weeks occur at the end of the tax year.
- If it was rented for fewer than 52 weeks, and there is a rental schedule in year t and $t-1$ and $t+1$ we assume the missing weeks occurred in the middle of the tax year. For example, if it's rented for 26 weeks, we assume it was rented from July to September, and April to June.
- If it was rented for fewer than 52 weeks, and there is no rental schedule in year $t-1$ or $t+1$ we assume the missing weeks occurred at either end of the tax year. For example, if it's rented for 26 weeks, we assume it was rented from October to March.

Future work could refine the algorithm. For example, work could be done to remove non-private housing (e.g. hospitals, care homes) from the measure, as they may bias the results. Data on rental bonds and transactions could be used to refine start and end dates.¹² Work could also be done to compare our measure to the Census for crossover periods.

¹¹ Note that if two people jointly own a property, they will both fill out rental schedules. Where there is more than one rental schedule in a year, we assume any indicating the same number of rented weeks represent co-ownership, so we only count them once (i.e. take the average). If they have different numbers of weeks we assume this represents a transaction and we sum the weeks (i.e. we treat them as non-overlapping periods).

¹² We also constructed measures focusing on only houses that were rented for the full period to ensure the algorithm was not affecting the result.

Appendix B: Identifying purchases

To identify purchases, we use the following algorithm. We assume it is an owner-occupier purchase if:

- An individual moves into a new dwelling where they have not lived before and stays there until the next year. This ensures we are not picking up children moving back in with parents, or owner-occupiers who have been in place for several years.
- There is no rental schedule attached to the property in the following year. This ensures that the person is an owner-occupier, not a renter. We do not require that there is no rental schedule in the current year, as people may move into a previously rented home mid-year.
- The person did not previously rent out the property.
 - This removes the case where the person may have owned the property for several years, and only recently moved in.

References

- Agarwal, N., J Bishop and I Day. (2023).** 'A New Measure of Average Household Size' RBA Bulletin. <https://www.rba.gov.au/publications/bulletin/2023/mar/a-new-measure-of-average-household-size.html>
- Azar, J., Schmalz, M. C., & Tecu, I. (2018).** 'Anticompetitive Effects of Common Ownership'. *The Journal of Finance*, 73(4), 1513–1565. <https://doi.org/10.1111/jofi.12698>
- Australian Bureau of Statistics. (2019-20).** *Survey of Income and Housing, User Guide, Australia*. ABS. <https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/survey-income-and-housing-user-guide-australia/latest-release>.
- Australian Bureau of Statistics. (2023-24).** *Australian System of National Accounts*. ABS. <https://www.abs.gov.au/statistics/economy/national-accounts/australian-system-national-accounts/latest-release>.
- Australian Bureau of Statistics. (n.d.),** "PLIDA Privacy Statement." Australian Bureau of Statistics. 2021. <https://www.abs.gov.au/about/legislation-and-policy/privacy/privacy-abs/plida-privacy-statement>.
- Australian Taxation Office. (2025).** "ALIFE." Alife-Research.app. Research. 2025. <https://alife-research.app/info/overview>.
- Gibbins, T. (2025).** 'Unpacking average household size'. Plus 61. e61 Institute. <https://e61.in/unpacking-average-household-size/>
- Hansen, J and Rambaldi A. N. (2022),** "How do homes transfer across the income distribution? The roles of supply constraints, *IARIW General Conference*. [Hansen-Rambaldi-IARIW-2022.pdf](#).
- Hughes A (2024),** 'How the RBA Uses the Securitisation Dataset to Assess Financial Stability Risks in Mortgage Lending', *RBA Bulletin* July 2024.
- Productivity Commission. (2025).** "A-PLIDA-Nalysis: Using PLIDA for Public Policy Research and Reporting - Commission Research Paper." Pc.gov.au. 2025. <https://www.pc.gov.au/research/completed/plida>
- Reserve Bank of Australia. (2024).** *Resilience of Australian Households and Businesses. Financial Stability Review – March 2024*. <https://www.rba.gov.au/publications/fsr/2024/mar/resilience-of-australian-households-and-businesses.html>
- Summerfield, M., Garrard, B., Nesa, Mossamet., Kamath, R., Macalalad, N., Watson, N., Wilkins, R. and Wooden, M. (2024).** HILDA User Manual – Release 23. Melbourne Institute: Applied Economic and Social Research, University of Melbourne.

Copyright and Disclaimer Notice

BLADE and PLIDA Disclaimer

The following Disclaimer Notice refers to data and graphs sourced from the Australian Bureau of Statistics' BLADE (Business Longitudinal Analysis Data Environment) database. The results of these studies are based, in part, on data supplied to the ABS under the Taxation Administration Act 1953, A New Tax System (Australian Business Number) Act 1999, Australian Border Force Act 2015, Social Security (Administration) Act 1999, A New Tax System (Family Assistance) (Administration) Act 1999, Paid Parental Leave Act 2010 and/or the Student Assistance Act 1973. Such data may only be used for the purpose of administering the Census and Statistics Act 1905 or performance of functions of the ABS as set out in section 6 of the Australian Bureau of Statistics Act 1975. No individual information collected under the Census and Statistics Act 1905 is provided back to custodians for administrative or regulatory purposes. Any discussion of data limitations or weaknesses is in the context of using the data for statistical purposes and is not related to the ability of the data to support the Australian Taxation Office, Australian Business Register, Department of Social Services and/or Department of Home Affairs' core operational requirements.

Legislative requirements to ensure privacy and secrecy of these data have been followed. For access to MADIP and/or BLADE data under Section 16A of the ABS Act 1975 or enabled by section 15 of the Census and Statistics (Information Release and Access) Determination 2018, source data are de-identified and so data about specific individuals has not been viewed in conducting this analysis. In accordance with the Census and Statistics Act 1905, results have been treated where necessary to ensure that they are not likely to enable identification of a particular person or organisation.