Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC Satellite meeting at the ISI Regional Statistics Conference on *"Is the household sector in Asia overleveraged: what do the data say?"*

Kuala Lumpur, Malaysia, 15 November 2014

The development of databases linking micro and macro data – an Australian perspective[1]

Giancarlo La Cava, Reserve Bank of Australia

[1] This presentatio was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.

The Development of Databases Linking Micro and Macro Data: An Australian Perspective

Gianni La Cava

Reserve Bank of Australia

Irving Fisher Committee Satellite Seminar
November 15, 2014

Disclaimer: *The views expressed in this presentation are those of the author and not necessarily those of the Reserve Bank of Australia.*

Questions

- What are the challenges in linking national accounts data with micro data on household balance sheets, income and consumption?
- How has the Australian Bureau of Statistics (ABS) dealt with these challenges?

Questions

- What are the challenges in linking national accounts data with micro data on household balance sheets, income and consumption?
- How has the Australian Bureau of Statistics (ABS) dealt with these challenges?
- What does the distribution look like? Is the picture very different when consistent with the national accounts?

Questions

- What are the challenges in linking national accounts data with micro data on household balance sheets, income and consumption?
- How has the Australian Bureau of Statistics (ABS) dealt with these challenges?
- What does the distribution look like? Is the picture very different when consistent with the national accounts?
- Some uses and limitations of the data:
  - Why did the household saving rate rise sharply in the global financial crisis?
  - How has inequality evolved in Australia over the past decade?

Take-Home Messages

- Linking micro and macro data is difficult
- The ABS should be commended for taking on the challenge

Take-Home Messages

- Linking micro and macro data is difficult
- The ABS should be commended for taking on the challenge
- The new matched dataset allows for simultaneous analysis from both 'top-down' and 'bottom-up' perspectives

Take-Home Messages

- Linking micro and macro data is difficult
- The ABS should be commended for taking on the challenge
- The new matched dataset allows for simultaneous analysis from both ‘top-down’ and ‘bottom-up’ perspectives
- But there is little new information about distributional changes over time
- Not clear that we gain much over separate analysis of national accounts and unmatched survey data

The Challenges of Matching Micro and Macro Data: The Australian Solution

- Adjustments to existing micro data
- Estimating missing data items
- Irregular cross-sectional surveys
- Out-of-scope households

Data Availability

Table: Availability of Australian Micro and Macro Data

Year	**HH-level Income** Survey of Income and Housing (SIH)	**HH-level Wealth** Survey of Income and Housing (SIH)	**HH-level Consumption** Household Expenditure Survey (HES)	**Aggregate Income, Wealth and Consumption** National Accounts (ASNA)	**HH-level Income, Wealth and Consumption** Matched Micro-Macro Data
2003/04	X	X	X	X	X
2004/05				X	
2005/06	X	X		X	X
2006/07				X	
2007/08	X			X	X
2008/09				X	
2009/10	X	X	X	X	X
2010/11				X	
2011/12	X	X		X	X

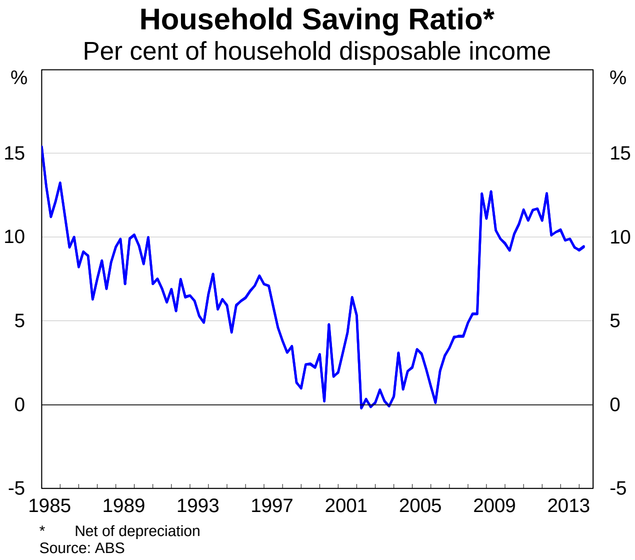
Sources: ABS; authors' calculations

Practical Uses and Limitations of Matched Data: A Central Banker's Perspective

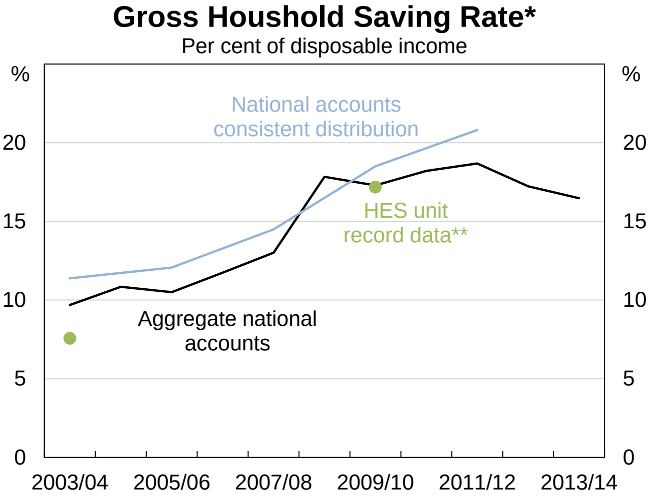
I consider two examples:

1. The rise in household saving in the global financial crisis
2. The evolution of household economic inequality in Australia

The Rise in Household Saving in Australia

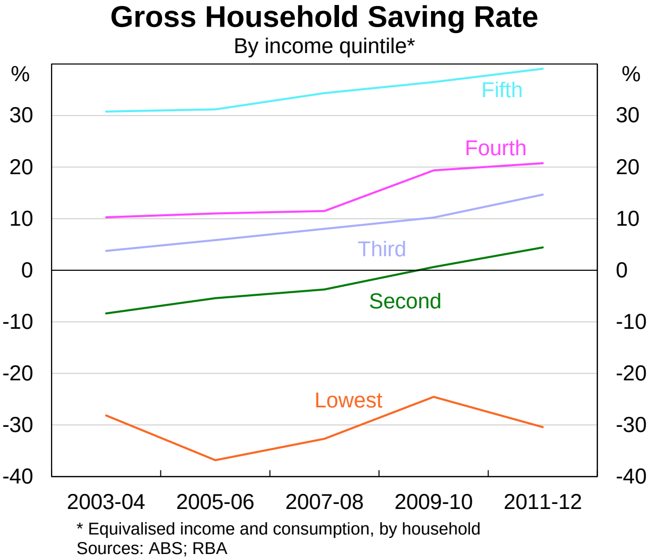


The Rise in Household Saving in Australia

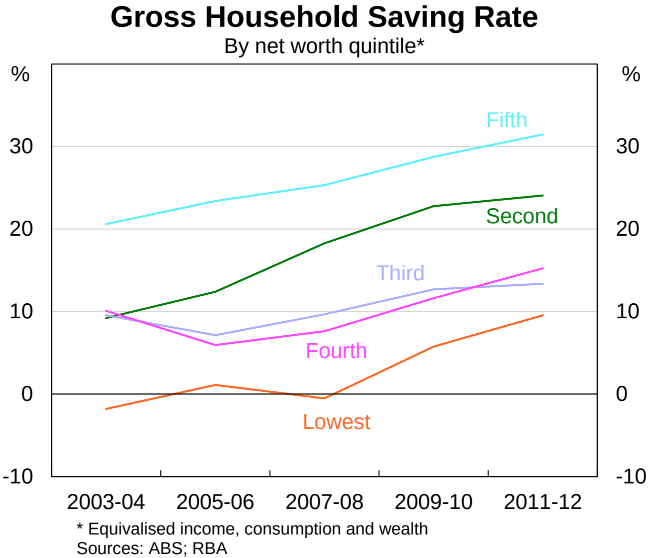


* Gross of consumption of fixed capital
Sources: ABS; RBA

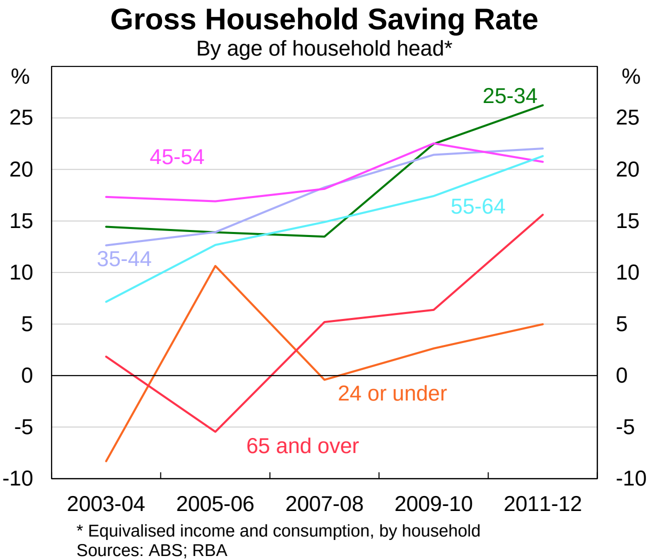
Household Saving by Income Quintile



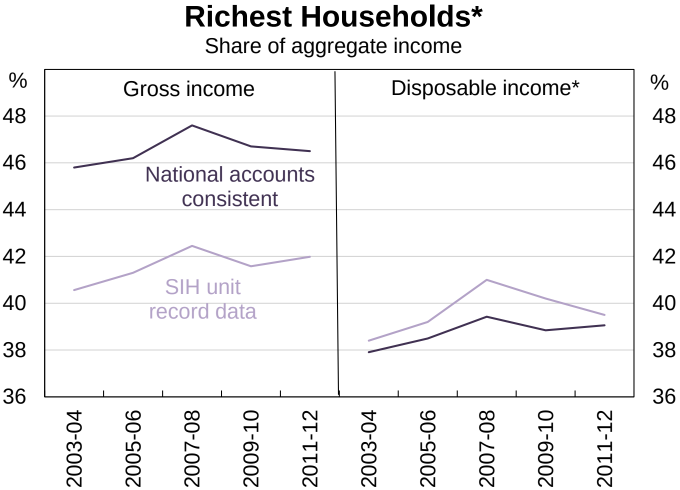
Household Saving by Net Wealth Quintile



Household Saving by Age Group



Income Inequality

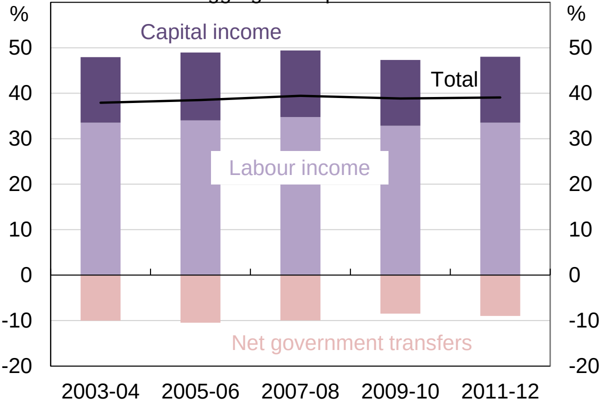


*Top income quintile
**Disposable income is gross income plus transfers less taxes
Source: ABS

Income Inequality

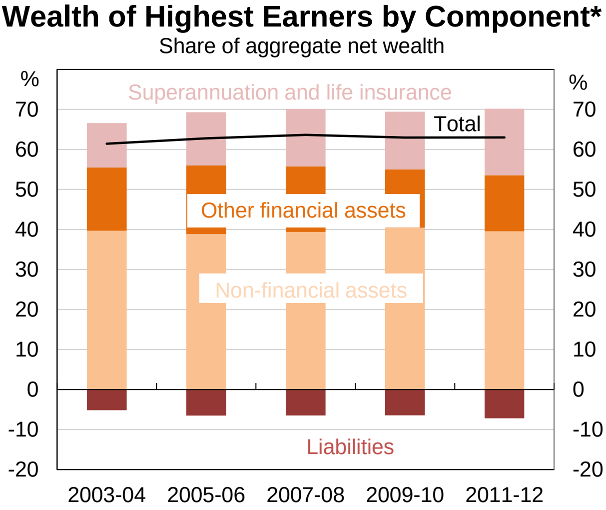
Income of Richest Households by Component*

Share of aggregate disposable income**



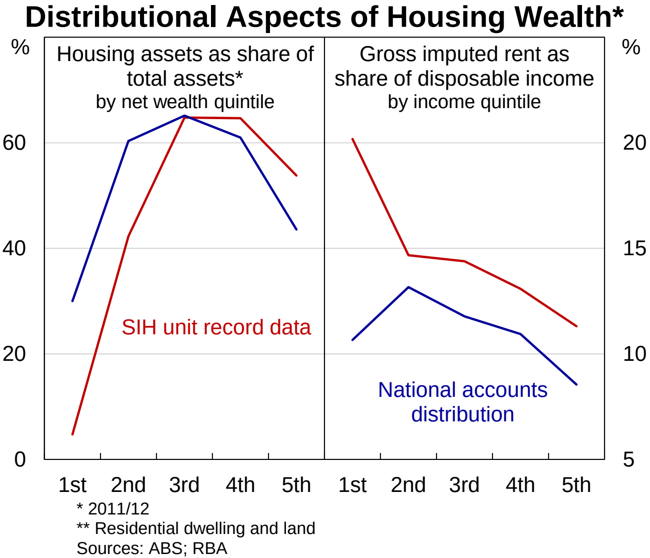
*Top income quintile
**Based on national accounts consistent data
Source: ABS

Wealth Inequality

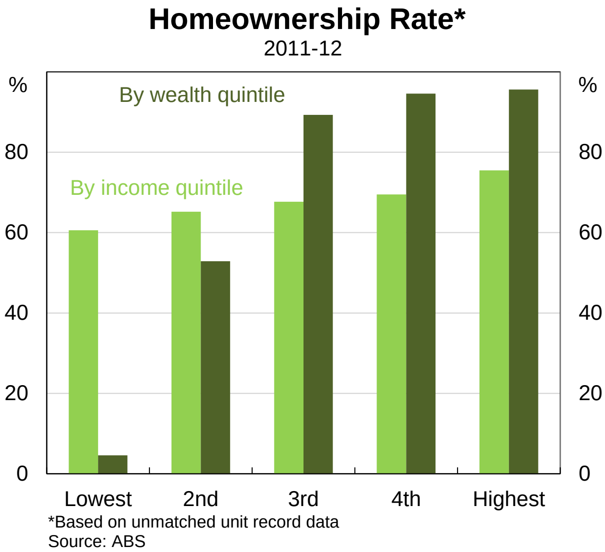


*Top wealth quintile
Source: ABS

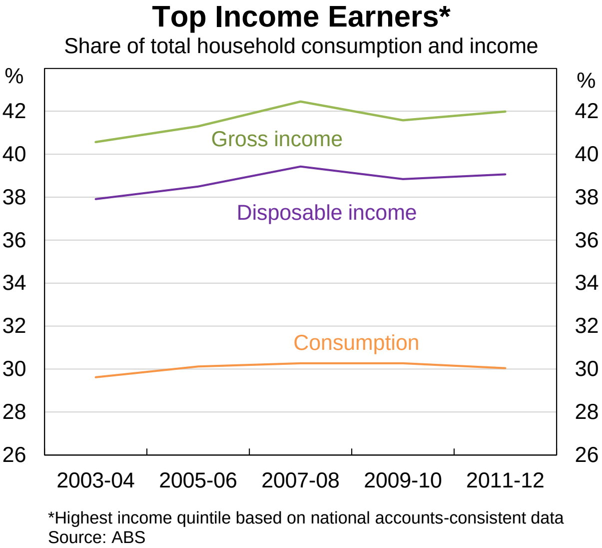
Housing Prices and Inequality



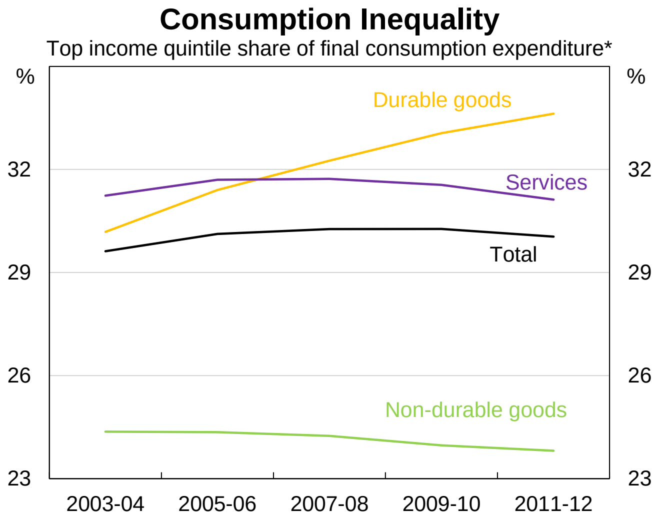
Housing Prices and Inequality



Consumption Inequality



Consumption Inequality



* Equivalised, by household
Sources: ABS; RBA

Summary

Strengths of matched micro-macro data:

- Consistency
- Complementarity
- Adaptability

Summary

Strengths of matched micro-macro data:

- Consistency
- Complementarity
- Adaptability

Weaknesses of matched micro-macro data:

- Relative contribution
- Lack of longitudinal information
- Timeliness

Irving Fisher Committee on
Central Bank Statistics

BANK FOR INTERNATIONAL SETTLEMENTS

IFC Satellite meeting at the ISI Regional Statistics Conference on *"Is the household sector in Asia overleveraged: what do the data say?"*

Kuala Lumpur, Malaysia, 15 November 2014

The development of databases linking micro and macro data – an Australian perspective[1]

Giancarlo La Cava, Reserve Bank of Australia

[1] This paper was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS or the central banks and other institutions represented at the meeting.

THE DEVELOPMENT OF DATABASES LINKING MICRO AND MACRO DATA – AN AUSTRALIAN PERSPECTIVE[1]

The Australian Bureau of Statistics (ABS) has, for the first time, recently released household survey data that are `matched' to national accounts aggregates. In this paper I discuss the conceptual and practical challenges in linking micro and macro household data and outline how the ABS has dealt with these challenges. I also consider the relative merits of the ABS matched data based on two practical examples of interest to Australian policymakers: the rise in household saving during the global financial crisis and the rise in household economic inequality over the past decade.

The new matched data are useful in that they are consistent with the national accounts and complement existing unit record data. In effect, users can simultaneously analyse the data from both `top-down' and `bottom-up' perspectives. However, the matched data do not appear to provide any new information about distributional changes over time and so are unlikely to provide many insights over and above the separate analysis of national accounts and unmatched survey data.

Introduction

The national accounts typically provide the most comprehensive information on the size and structure of the economy. However the national accounts rarely provide information on the distribution of household income, wealth and consumption. Distributional issues are very important to understanding living standards, and have been gaining more focus over recent years. These issues are also central to targeting, and improving the efficiency of, economic policies.

The Australian Bureau of Statistics (ABS) publishes both high-quality aggregate (macro) and distributional data (micro) on household economic well-being. The macro estimates are published in the Australian System of National Accounts (ASNA). The national accounts present estimates for the household sector as a whole but do not provide any information about how available resources are distributed. Instead, distributional estimates are derived from two household surveys – the ABS Survey of Income and Housing (SIH) and the Household Expenditure Survey (HES).

Due to differences in concepts, definitions and statistical practices, micro data may yield results that diverge from national accounts aggregates, and therefore distributional measures created using micro data sources may not be consistent with the aggregate national accounts estimates. To address this, the ABS has recently published results that integrate the ABS micro and macro sources. In essence, they have produced distributional data on household income, consumption and wealth that

1 The views expressed in this paper are those of the author and do not necessarily reflect the views of the Reserve Bank of Australia. This is a preliminary draft so please do not quote.

are consistent with the Australian System of National Accounts (ASNA) concepts and aggregates. I will refer to these distributional data as `matched micro-macro data'.

This matching allows researchers to undertake both top-down and bottom-up analyses on a consistent basis. But the matching also presents some conceptual and practical challenges. I am a central banker by trade and so, in assessing the relative merits of matched micro-macro data, I will take the perspective of a data user and, more specifically, a macroeconomist with a keen interest in micro data. I will address the following questions:

- What are some of the challenges in linking national accounts data with micro data on household balance sheets, income and consumption?
- How has the Australian Bureau of Statistics dealt with these challenges?
- What does the Australian distribution of household wealth, income and consumption look like? Is the picture very different when re-formulated to be consistent with the national accounts? Why or why not?
- What are some of the uses and limitations of these data (from the perspective of a central banker)? I consider two separate examples:
    - Why did the aggregate household saving rate rise sharply in the global financial crisis?
    - How has household economic inequality evolved in Australia over the past decade?

I will now discuss the challenges involved in the matching process, and how the ABS has dealt with them. And then I'll discuss some concrete examples of how the data are useful to policymakers. Along the way, I'll point out some of the pros and cons that I perceive with the matched data.

The Challenges of Matching Micro and Macro Data: The Australian Solution

Coverage Rates

To construct the matched micro-macro estimates, the ABS first compares the household national accounts estimates (macro) to the corresponding household survey estimates (micro).[2] More specifically, the ABS calculates a "coverage rate" for each income, consumption or wealth component. The coverage rate is the ratio of the value of the aggregated micro estimate (from the unmatched unit record data) to the corresponding value of the aggregate macro estimate from the national accounts.

2 My discussion of the challenges in linking micro and macro data borrows heavily from the ABS' recent publication entitled "Australian National Accounts: Distribution of Household Income, Consumption and Wealth, 2003-04 to 2011-12".

It is instructive to consider how coverage rates vary across different components of household income, spending and wealth. Fesseau, Wolff and Mattonetti (2013) provide estimated coverage rates as calculated by experts at various statistical agencies around the world.[3] Some specific estimates of Australian coverage rates can be obtained from an ABS (2013) information paper.

The cross-country estimates of income coverage rates suggest that the aggregated micro data underestimate disposable income as measured by the national accounts. The average coverage rate across all income sources is about 85 per cent. Across income components, the household survey and national accounts data are reasonably well aligned for wages and salaries (with a coverage rate of 93 per cent) and employees' social contributions (94 per cent). In contrast, coverage rates are low for dividend income (53 per cent) and income from self-employment (67 per cent).

The micro and macro estimates of household consumption are typically less closely aligned than household income, with an overall coverage rate of about 75 per cent. Across components, the micro and macro data appear to be relatively well-aligned for spending that is highly visible and frequently observed, such as actual rents paid (98 per cent). In contrast, the aggregate micro data significantly understate the national accounts estimates of spending on tobacco (40 per cent) and alcohol (51 per cent). According to the statistical agencies, the low coverage rates for alcohol and tobacco primarily reflects survey under-reporting.

For household wealth, the overall coverage rate is lower than both income and consumption at about 70 per cent. The most closely aligned wealth components include the value of mortgage debt (103 per cent) and holdings of equities (91 per cent). Notably, the coverage rate on the stock of equities is much higher than on the associated flow of dividend income, on average. In contrast, coverage rates are low for accounts receivable (37 per cent) and intangible fixed assets (48 per cent).

These differences in coverage rates across income, consumption and wealth items are interesting. It would also be informative to consider how the coverage rates vary over time for a given survey. However, these coverage issues are beyond the scope of this paper and I would refer the interested reader to Fesseau et al (2013). I will briefly discuss later some issues in comparing micro and macro estimates of housing wealth.

Adjustments to the Micro Data

Based on the estimated coverage rates, the macro and/or micro estimates for some items are adjusted to derive the most relevant common scope for comparison. To do this, the corresponding micro household items are divided into several household groups: main source of income;

3 It should be kept in mind that the coverage rate is not a measure of the (relative or absolute) quality of the micro estimates; the compilation methods followed by macro data producers may have different degrees of reliability and macro estimates are often subject to revisions. The estimated coverage rates can also differ a lot across countries. Furthermore, the coverage rate for some items is zero because the data are not collected at the micro level.

equivalised income quintiles; household composition (single parent versus couple households etc); age of household reference person; and equivalised net worth quintiles.

The Australian System of National Accounts (ASNA) household components and aggregates were distributed to the various household groups in several different ways depending on the estimated coverage rates:

- directly using the distribution of the equivalent micro component when the coverage rate was considered adequate (e.g. social assistance benefits);
- indirectly by a related micro distribution when there was no direct micro distribution information for the national accounts item (e.g. non-life insurance claims were distributed using the micro distribution for total insurance premiums paid);
- indirectly by creating a micro distribution ('synthesised') based on a related micro distribution (e.g. a synthesised micro distribution was created for financial intermediation services indirectly measured for consumer loans); and
- by the corresponding aggregate distribution for income, consumption, or wealth when micro distributions either directly or indirectly were not available (e.g. retained earnings on foreign investments).

Irregular Cross-sectional Surveys

This matching approach is suitable for the periods in which both micro and macro data are readily available. But, unlike the national accounts, the ABS does not conduct household cross-section surveys every year, nor does it collect information on household income, consumption and wealth in each survey. Over the past decade, the ABS has typically collected income and wealth data (from the SIH) every two years and consumption data (from the HES) every six years (Table 1).[4] In other words, the income and wealth data are more `complete' than the consumption data.

4 The ABS also conducted the Census in 2006 and 2011, which collects individual data for basically the entire population of Australia.

Table 1: Availability of Australian Micro and Macro Data					
	Household-level Income Survey of Income and Housing (SIH)	**Household-level Wealth** Survey of Income and Housing (SIH)	**Household-level Consumption** Household Expenditure Survey (HES)	**Aggregate Income, Wealth and Consumption** National Accounts (ASNA)	**Household-level Income, Wealth and Consumption** Matched micro-macro dataset
2003/04	X	X	X	X	X
2004/05				X	
2005/06	X	X		X	X
2006/07				X	
2007/08	X			X	X
2008/09				X	
2009/10	X	X	X	X	X
2010/11				X	
2011/12	X	X		X	X

This requires the ABS to `fill the gaps' in the survey data. The ABS investigated two options to model the distributional household indicators for the years that the source micro data were not available: 1) use the nearest available source data for the missing years and 2) linearly interpolate (or extrapolate) the data for the missing years. The ABS chose the second option given that the time in between surveys is quite long, especially for the consumption data collected in the HES.

The ABS outlines a simple example to explain the methodology used to linearly interpolate and extrapolate missing values. Let point A be a known value (V_A) at a known point in time (T_A) and let point B be another known value (V_B) at another known point in time (T_B) where $T_B > T_A$. Let point X be an unknown value (V_X) at a known point in time (T_X) that we want to estimate. Linear interpolation allows for the value of point X to be calculated based on the formula:

$$Vx = \frac{(Tx - Ta)\,Vb + (Tb - Tx)Va}{Tb - Ta}$$

To see how this applies, consider the hypothetical ABS example in Table 2 in which we have information on average clothing and footwear spending for each (equivalised) household income quintile in 2003/04 and 2009/10 but not in 2005/06.

Table 2: Household Spending on Clothing and Footwear by (Equivalised) Income Quintile Australian dollars						
Financial Year	**Numeric Year**	**1st Quintile**	**2nd Quintile**	**3rd Quintile**	**4th Quintile**	**5th Quintile**
2003/04	2003	1333	1950	2528	3305	4423
2005/06	**2005**	**1512**	**2089**	**2736**	**16962**	**5238**
2009/10	2009	1871	2366	3152	4277	6869

For the first quintile, using the formula, we would obtain an estimated average spending on clothing and footwear of $1512 for 2005/06. This is shown in the 2005/06 row of the table and below:

$$Vx = \frac{(2005 - 2003) * 1871 + (2009 - 2005) * 1333}{(2009 - 2003)} = 1512$$

This process is repeated to generate data for every other income quintile (and for other household groups, such as wealth quintiles and age groups) for all the survey years in which the consumption data are missing.

Estimating Missing Data Items

A key challenge in linking micro and macro data is dealing with missing components. Micro data sources often do not provide information for a number of national accounts components (Fesseau et al (2013)). Some examples include:

- Employers' imputed social contributions
- Financial Intermediation Services Indirectly Measured (FISIM)
- Reinvested earnings on foreign direct investment
- Property income attributed to insurance policy holders
- Social Transfers in Kind
- Imputed rents for owner-occupied housing

These components can be missing in micro sources for both conceptual and practical reasons. For example, the compilers of micro statistics may consider that some national accounts components that are useful for describing the economy as a whole are not relevant when the focus is the economic behaviour of households (e.g. FISIM). On the other hand, some components of national accounts aggregates may be missing for practical reasons because the information is difficult to collect or impute (e.g. imputed rent for owner-occupier households).

In the Australian case, the earlier household surveys (2003/04 and 2005/06) also did not include some balance sheet items that were collected in later surveys (2009/10 and 2011/12). This implies that the aggregated data from the earlier surveys underestimate the size of household balance sheets relative to later surveys. This is particularly problematic in the case of missing income and wealth data as these data are used to derive the income and wealth quintiles. The missing data items were estimated by applying `factors' to the earlier survey data.

To take an example, there is distributional information on Social Transfers In Kind (STIK) in the 2009/10 and 2011/12 surveys but not the 2005/06 or 2007/08 surveys. The later surveys can be used to estimate the relative contribution of income (and spending) that is due to STIK for each income quintile in the earlier surveys. In 2011/12 these transfers made up a relatively large share of income for poor households (in the bottom quintile) at 44 per cent and a low share of income for rich households (in the top quintile) at 8 per cent. These factors can be then applied to the earlier survey data depending on the relative position of each household in the income distribution. The ABS methodology paper is not clear on how these factors are calculated (or applied) but presumably the income of each household is scaled up by the proportion of STIK depending on their quintile. Regardless, the ABS indicates that relatively few items are affected by this procedure.[5]

Out of Scope Households

There is an additional practical challenge in incorporating information on people that are typically out of scope for micro surveys, such as those living in very remote communities and those living in non-private dwellings (e.g. prisons, hospitals, nursing homes). These people were excluded from the ASNA estimates and distributed separately using data from the 2006 and 2011 ABS Census of Population and Housing. These distributions were then added to the ASNA distributions based on the micro surveys to obtain the final distribution of the ASNA household balance sheet estimates.

Practical Uses and Limitations of Matched Data: A Central Banker's Perspective

Despite being a macroeconomist I am a strong advocate for micro data and analysis. In fact, I believe the future of applied macroeconomic research *is* micro data. So it's interesting to consider the types of policy questions that we can address with the matched micro-macro data. Based on my perspective as a macroeconomist I will now run through some examples that demonstrate the benefits and limitations of the matched distributional estimates.

5 It would seem possible to use more advanced econometric techniques to allow for changes in the distribution over time. For instance, there is detailed information on the characteristics of individual households in each survey year (e.g. income, education, age, migrant status). It should be possible to replace missing estimates of, say, household spending for each individual household using their observed characteristics and matching techniques. The resulting distribution could then be adjusted to be consistent with the national accounts aggregates.

The Rise in Household Saving in the Global Financial Crisis

Australia experienced a sharp rise in the household saving rate during the global financial crisis (Figure 1). In fact, the saving rate more than doubled within the space of just six months and has basically remained at an elevated level since then. Similar sharp increases in household saving occurred in other advanced economies ([Mody, Ohnsorge and Sandri (2012)]). Household saving has largely remained elevated since that time. This raises an obvious question: what caused the saving rate to rise? The distributional data are clearly suited to addressing this type of question as they allow us to dig deeper into the national accounts estimates, and examine *which* households are responsible for the change in saving behaviour.

Figure 1

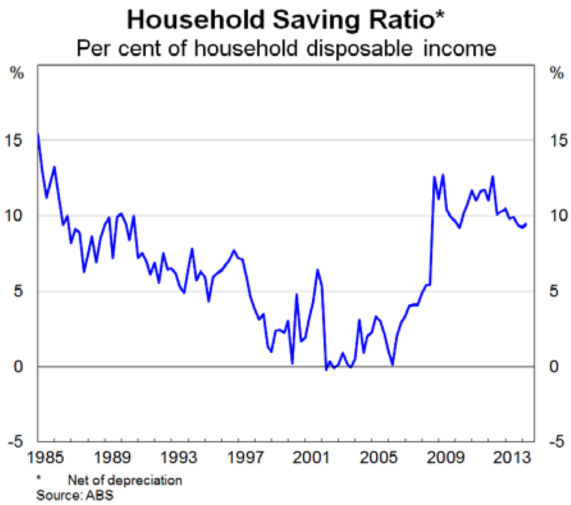
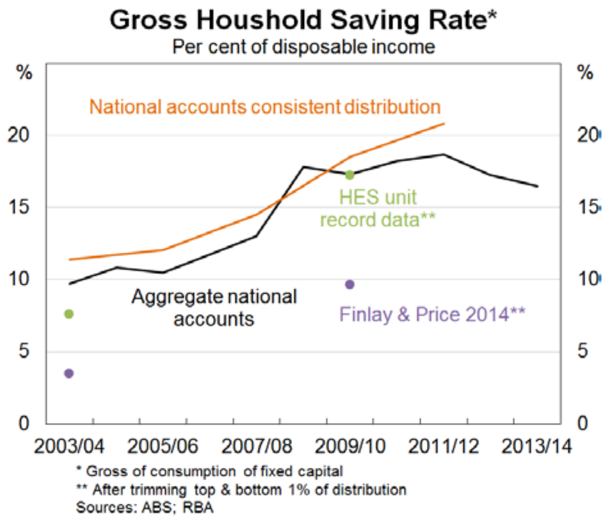


Figure 2



To begin, it is worth noting that the household saving rate implied by the matched data lines up slightly better with the national accounts than the estimate based on the unmatched unit record data (Figure 2). This is not surprising given the data are specifically designed to be consistent with the national accounts. More importantly, the *trends* in the matched and unmatched data are essentially the same, presumably due to the linear interpolation method. And this particular research question is about explaining the *change over time* in the saving rate. For this practical purpose, it appears that the matched data do not contribute much *over and above* the (aggregated) unmatched unit record data because they do not really allow for distributional changes over time.

Looking more closely at the disaggregated matched data, there is very wide dispersion in saving rates across income groups; the richest households save more than 30 per cent of their disposable income while the poorest households spend about 130 per cent of their disposable income, on average (Figure 3).

The matched data also indicate that the rise in saving over the decade was very broad-based; saving rates rose at a similar rate in each income quintile (Figure 3) wealth quintile (Figure 4) and age group (Figure 5).

Figure 3

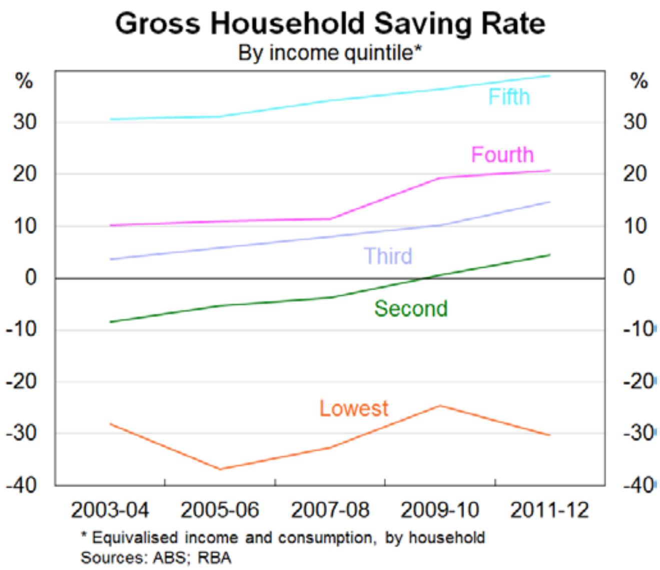


Figure 4

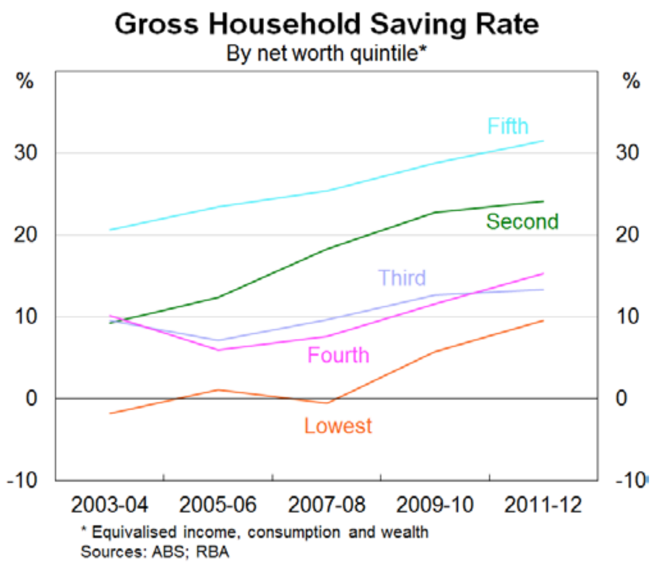
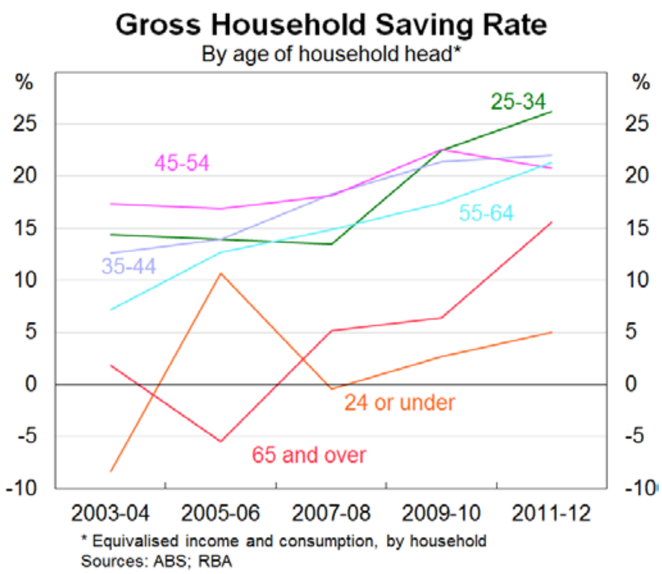


Figure 5



These decompositions illustrate a key limitation of the matched data in that they are only currently available on a univariate basis. More specifically, there is information on spending by income group *or* age group but not income *and* age group. And there is no information on the distribution of household spending and income across other important dimensions, such as the level of education of the household head. In contrast, the unit record data, even if unmatched, provide flexibility in allowing the user to explore the data in a multivariate fashion. Moreover, the unit record data can be used to estimate econometric models that control for a wide range of observed household characteristics that might explain household saving (see Finlay and Price (2014) for a more detailed analysis of the Australian case).

Taking this argument one step further, to properly model household saving behaviour, ideally we would have longitudinal household information (with the same households sampled each period) rather than repeated cross-sectional surveys (with different households sampled each period). There are significant benefits to having panel data on households as opposed to the repeated cross-sections available in the matched ABS data. For example, by allowing us to track the same individuals over time, panel data would allow us to control for unobservable characteristics that do not vary with time (e.g. a person's level of risk aversion). Therefore, as a user, I would argue that statistical authorities should focus more on building longitudinal panel datasets than on matching cross-sectional surveys to national accounts. I'll discuss this point in more detail in the next example.

In summary, the household saving rate example suggests that the matched data are consistent with the national accounts as the two aggregate estimates line up very well at a point in time. But the matched data do not really provide any new information about distributional changes over time. For this, we still need the unmatched unit record data. In other words, the matched data are a complement rather than a substitute for the unmatched unit record data.

The Evolution of Household Economic Inequality in Australia

The matched dataset can also be used to explore various dimensions of household economic inequality, including inequality in household income, wealth and consumption. I will consider each type of inequality in turn.

Income Inequality

Similar to several other advanced economies, Australia has experienced an increase in income inequality over the past decade (Piketty (2013)). However, the rise in inequality has been relatively modest compared to countries such as the United States. We can construct estimates of inequality using the share of aggregate income that flows to the richest households. According to the matched data, the share of total gross income accounted for by the richest households (in the top quintile) rose by 1.4 percentage points from about 40.6 per cent in 2003/04 to 42.0 per cent in 2011/12 (Figure 6, left-hand-side panel). The unmatched data suggest a relatively similar trend in inequality, although the level of inequality is lower by 5 percentage points on average. The estimates of disposable income inequality (which takes into account the effect of government taxes and transfers) display a similar time-series pattern (Figure 6, right-hand-side panel). Although, somewhat surprisingly, the matched estimates for disposable income inequality are lower than the corresponding unmatched estimates (which is the opposite pattern to gross income inequality).[6]

6 The inclusion of social transfers in kind (STIK) significantly affects the level of measured inequality but it appears to have little impact on the trend. When we adjust the income estimates to account of STIK, the share of disposable income accounted for by the top earners rose from about 33 per cent in 2003/04 to 34.5 per cent in 2011/12

Figure 6

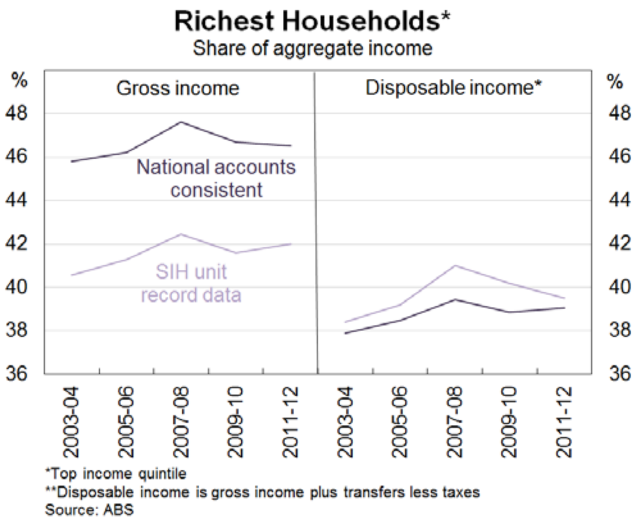
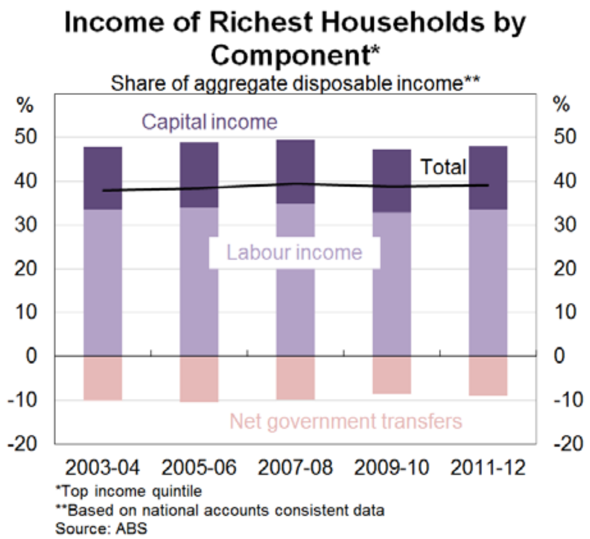


Figure 7



As in the saving rate example, the very similar trends in the unmatched and matched data should not be too surprising given the interpolation method used by the ABS. But, saying that, there are differences; for instance, between 2009/10 and 2011/12, the matched estimates suggest that disposable income inequality was rising, while the unmatched estimates imply that it was falling. The fact that there are some differences in the trends for disposable income inequality implies that there might be new information in the matched data.[7] But what factors are driving the differences?

The matched data allow us to decompose total gross disposable income into different components (e.g. labour and capital income, as well as net government transfers). This allows us to calculate, for aggregate disposable income, the individual contributions of the income components for rich households. This decomposition implies that the rise in income inequality between 2003/04 and 2007/08 was fully explained by an increase in labour income inequality (Figure 7). For the decade as a whole, a decline in income taxes on rich households appears to explain the slight rise in overall inequality in household disposable income.

We could perform a similar decomposition with the unmatched data and this might help to shed light on the different trends in inequality. But I suspect the slightly different trends between the matched and unmatched inequality estimates are due to different income definitions underpinning the quintile decompositions. More precisely, in any given survey year, a household in the top income quintile based on the matched estimates might be in a lower quintile based on the unmatched data due to, say, the inclusion of certain income items (e.g. FISIM) in only the matched estimates. However, as in the saving rate example, there is a limit to our ability to determine whether this is the case given the matched data are only available on a univariate basis.

7 It is likely that matched data would be very useful for understanding household spending patterns in the United States. Several papers document the fact that aggregate measures of expenditure from the US Consumer Expenditure Survey (CES) do a poor job of reproducing the level of expenditure in the national accounts data (e.g. Garner and Maki (2004), Attanasio, Hurst and Pistaferri (2012)).

Wealth inequality has also risen in Australia over the past decade. The matched data indicate that the share of total net wealth accounted for by the wealthiest households rose from 61.4 per cent in 2003/04 to about 63 per cent in 2011/12 (Figure 8). The rise in wealth inequality actually occurred over the first half of the decade, with wealth inequality falling slightly since 2007/08.

As in most advanced economies, housing wealth is the largest component of aggregate household wealth in Australia. Somewhat surprisingly though, changes in housing prices do not appear to be the main determinants of changes in inequality over the past decade. Instead, the matched data suggest that the rise in wealth inequality has been due to a rise in inequality in financial asset holdings and, in particular, superannuation and insurance reserves (Figure 8).

Figure 8

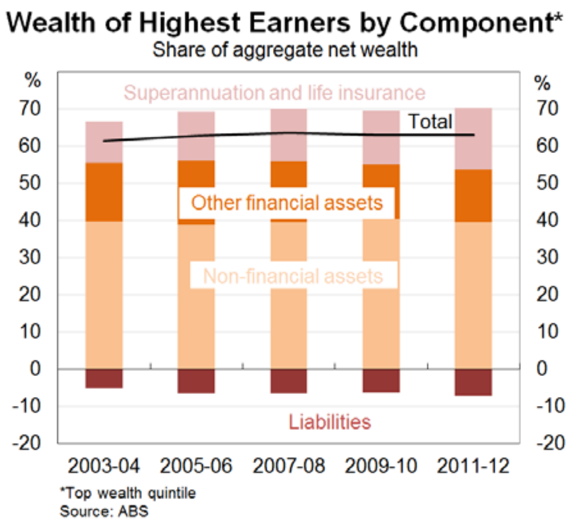
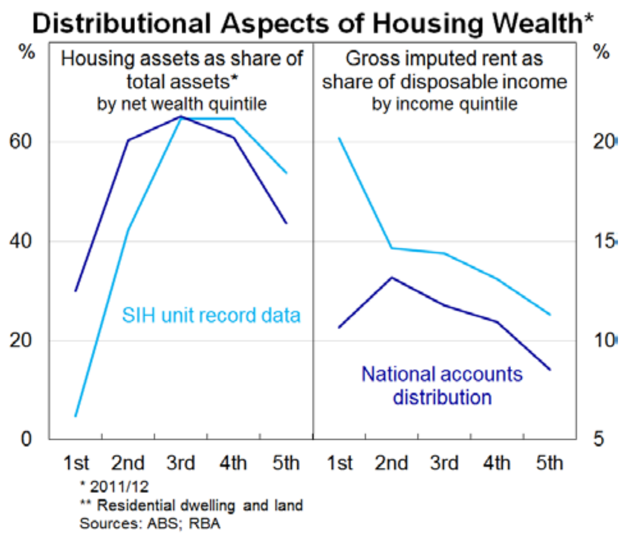


Figure 9



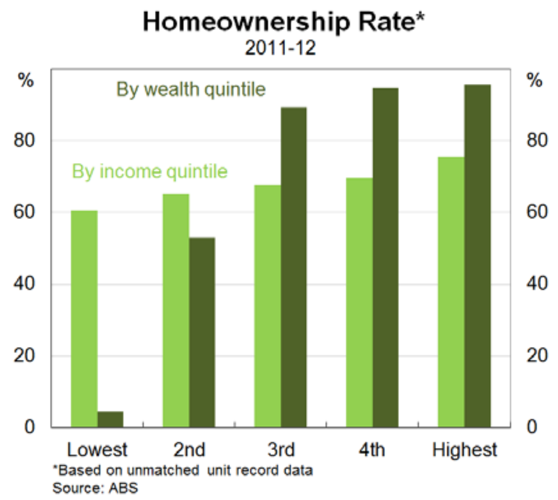
But that is not to say that housing does not matter for wealth inequality. According to the matched data, housing assets comprise a relatively small share of total assets for low-wealth households (at about 30 per cent), but close to 70 per cent of total assets for households in the middle wealth quintiles (Figure 9, left-hand side panel). The picture is very different when we look at the role of housing in income inequality. In this case, the imputed rent from housing makes up a relatively large share of income for low-income households (Figure 9, right-hand panel).

Interestingly, these differences in the income and wealth distributions suggest that changes in housing prices can have very different effects on income and wealth inequality. In particular, a rise in housing prices will typically cause wealth inequality to rise but income inequality to fall, all other things being equal.

This is because low-*income* households are not

Figure 10

the same as low-*wealth* households. Low-income households comprise a relatively large share of older retired households that own their own homes. Low-wealth households include a high share of young households that are credit-constrained and rent. In effect, the homeownership rate is much higher amongst low-income households than amongst low-wealth households (Figure 10). This difference in homeownership rates affects the sensitivity of inequality to housing prices.



This again demonstrates some of the limited capability of the matched dataset – the information about different homeownership rates within income and wealth quintiles (and hence the underlying cause of changes in inequality) can only be gleaned from the unmatched unit record data.

Given the potential sensitivity of the distributions of wealth and income to housing prices it is worthwhile considering how the ABS constructs its estimates of housing wealth on both a micro and macro basis. To obtain a macro estimate of the total value of the dwelling stock in Australia, the ABS uses information from home sale prices (with econometric adjustments made for the value of homes that are not sold). In other words, the macro estimate of housing wealth is based on the market value of the dwelling stock. In contrast, micro estimates of housing wealth are typically taken from household surveys and are based on self-reported assessments by surveyed homeowners.

The discrepancies between the micro and macro estimates might be quite informative about the state of the economy and, in some cases, macroeconomists might find it useful to preserve these differences. For instance, there may be differences between how much individual homeowners believe their homes are worth (as implied by the survey responses) and how much the market thinks they are worth (as implied by market valuations). In fact, Windsor, La Cava and Hansen (2014) use household-level panel data to show that the differences between homeowner beliefs and market valuations (`home valuation differences') are useful for predicting household consumption, leverage and wealth portfolio allocations. By attempting to reconcile these separate estimates we may be `throwing away' valuable information about the state of the housing market. In effect, we should again think of the matched and unmatched estimates as being complementary.

Consumption Inequality

Inequality studies typically focus on income and, to a lesser extent, wealth inequality. But there are several reasons why it is useful to also examine consumption inequality. First, some economists consider consumption to be a more appropriate measure of household wellbeing than income. If some households smooth temporary fluctuations in income by borrowing and saving, then income will be more variable than expenditure at a point in time and hence income will overstate the level of inequality in household welfare. Second, estimates of inequality based on consumption can be a useful cross-check if income estimates are affected by measurement error. Third, examining changes in the distribution of consumption relative to income can shed light on household saving and borrowing patterns (Beech, Dollman, Finlay and La Cava (2014)).

Unfortunately, the ABS matched dataset does not provide enough information on the household consumption distribution to be able to split households into consumption quintiles. However, the ABS does provide consumption data divided by income quintiles. So we can proxy for consumption inequality by looking at the share of total household expenditure accounted for by the richest households (i.e. consumption can be divided by *income* quintiles).

These estimates indicate that consumption inequality is lower than income and wealth inequality (Figure 11). This suggests that households are able to smooth fluctuations in income through borrowing and saving. Moreover, consumption inequality has been unchanged over the past decade; the richest households accounted for about 30 per cent of household consumption each year.

Figure 11

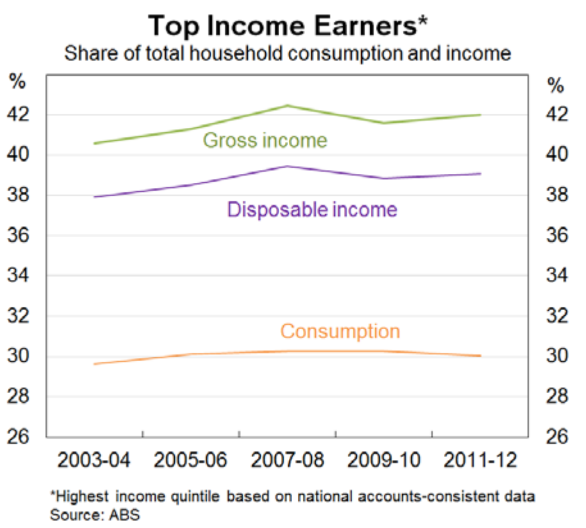
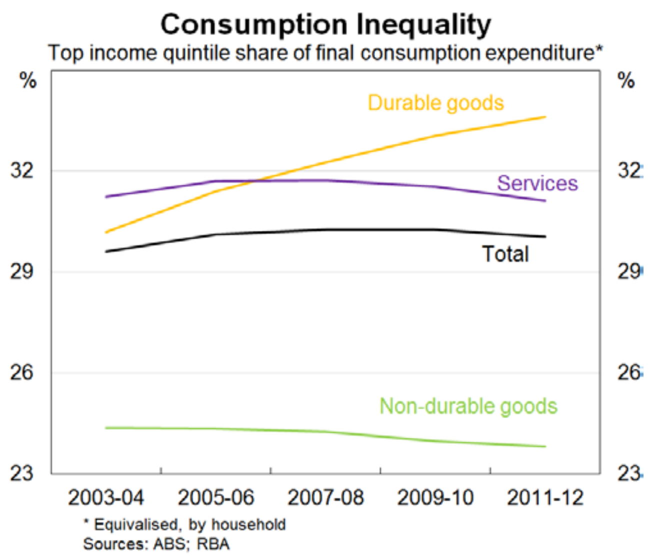


Figure 12



The lack of a trend in the overall consumption inequality estimates disguises different trends in the underlying spending components. For instance, the matched estimates imply that the richest households are increasingly accounting for a larger share of durable goods spending (especially on clothing and footwear) while services and non-durable goods spending have become more equally distributed across the household population (Figure 12).

Ideally, we would be able to examine the joint behaviour of the distributions of income and consumption in the matched dataset. But the data do not allow this. This is similar to the previous example – the general lack of longitudinal information makes it difficult to properly model the determinants of household economic inequality. Again, panel data are really needed to model income dynamics and be able to distinguish between permanent and temporary income shocks (e.g. Blundell (2014)). So, as a data user, I would again argue that statistical authorities should focus less on matching surveys to national accounts and more on building databases that exploit existing longitudinal administrative data (e.g. tax records, medical records).

Summary

The matched data produced by the ABS has a number of advantages over previously separate distributional and aggregate datasets:

Strengths

1) Consistency

The main strength of the matched data is its consistency along several dimensions. First, the data are *consistent with aggregate benchmarks*; the distributional data are benchmarked to the national accounts enabling users to interpret household distributional data within the broader context of published estimates on the Australian economy. Second, there is a *consistent time-series* with linear interpolation (and other modelling techniques) implemented where data were missing. Third, the data are *consistent with international standards*; the dataset is based on work undertaken in an Organisation for Economic Cooperation and Development and Eurostat expert group for measuring disparities in the national accounts. As a result, the ABS data are comparable with any time-series analysis on the distribution of household income, consumption and wealth in a national accounts framework performed by members of the expert group.

2) Complementarity

The construction of matched datasets allows users to separately undertake both top-down and bottom-up approaches, while also allowing them to directly reconcile the two sets of estimates through the matched data. To achieve this, it will be important to ensure users can always access the unit record data. The matched data are a complement (and not a substitute) for unmatched unit record data.

3) Adaptability

The existing matched dataset can be extended with future data points and revised with new source data, enabling a more accurate and longer time series. The time-series is based on a methodology formulated through an international expert group, and new source data from micro surveys or the national accounts can be used to easily revise and update the current data with future data points.

Weaknesses

The ABS has identified some limitations of the current matching approach – the long time between surveys and the lack of data on remote communities. From the perspective of a data user, I believe there are a few other limitations of the matched data:

1) Relative Contribution

In the Australian case at least, we already have micro data that is reasonably consistent with the national accounts. For example, the aggregate household saving rate follows basically the same trend in the national accounts and aggregated unit record data even *prior to* matching. And, in cases where there are different trends in the matched and unmatched data, such as disposable income inequality, there is currently insufficient information in the matched dataset to determine the factors that might be driving these differences.

2) Lack of Longitudinal Information

There are significant benefits to having panel data on households as opposed to repeated cross-sections (as represented by the matched ABS data). I would argue that statistical agencies should focus more on developing databases based on existing longitudinal administrative data and less on databases that link micro and macro data.

3) Timeliness

The matched Australian data currently has a two-year publication lag, with the latest available data for 2011/12. As a result, the matched data enable us to better analyse past events, but the data are less useful for understanding recent developments in the economy. So it might be worth considering whether it makes sense to collect a more narrow set of micro data to increase timeliness. Again, existing administrative datasets, such as tax records, are reasonably timely and could be useful for this purpose.

Gianni La Cava
Reserve Bank of Australia
30 October 2014

References

Attanasio O, E Hurst and L Pistaferri (2012), 'The Evolution of Income, Consumption and Leisure Inequality in the US, 1980-2010', National Bureau of Economic Research (NBER) Working Paper No 17982

Australian Bureau of Statistics (2013), 'Information Paper: Australian National Accounts, Distribution of Household Income, Consumption and Wealth, 2009-10', Cat No 5204.0.55.009, Canberra.

Australian Bureau of Statistics (2014), 'Australian National Accounts: Distribution of Household Income, Consumption and Wealth, 2003-04 to 2011-12', Cat No 5204.0.55.011, Canberra.

Beech A, R Dollman, R Finlay and G La Cava (2014), 'The Distribution of Household Spending in Australia', RBA Bulletin, March 2014

Blundell R (2014), Income Dynamics and Life-cycle Inequality: Mechanisms and Controversies, The Economic Journal, 124 (576), May 2014, pp 289-318.

Fesseau M, F Wolff and ML Mattonetti (2013), 'A Cross-Country Comparison of Household Income, Consumption and Wealth Between Micro Sources and National Accounts Aggregates,' November 2013

Finlay R and F Price (2014), 'Household Saving in Australia', RBA Research Discussion Paper No 2014-03.

Garner TL and A Maki (2004), 'The Gap Between Macro and Micro Economic Statistics: Estimation of the Misreporting Model using Micro-data Sets Derived from the Consumer Expenditure Survey', mimeo

Mody A, F Ohnsorge and D Sandri (2012), 'Precautionary Saving in the Great Recession', IMF Economic Review, 60, pp 114-138

Piketty T (2013), 'Capital in the Twenty-First Century', Harvard University Press

Windsor C, G La Cava and J Hansen (2014), 'Home Price Beliefs in Australia', RBA Research Discussion Paper No 2014-04.