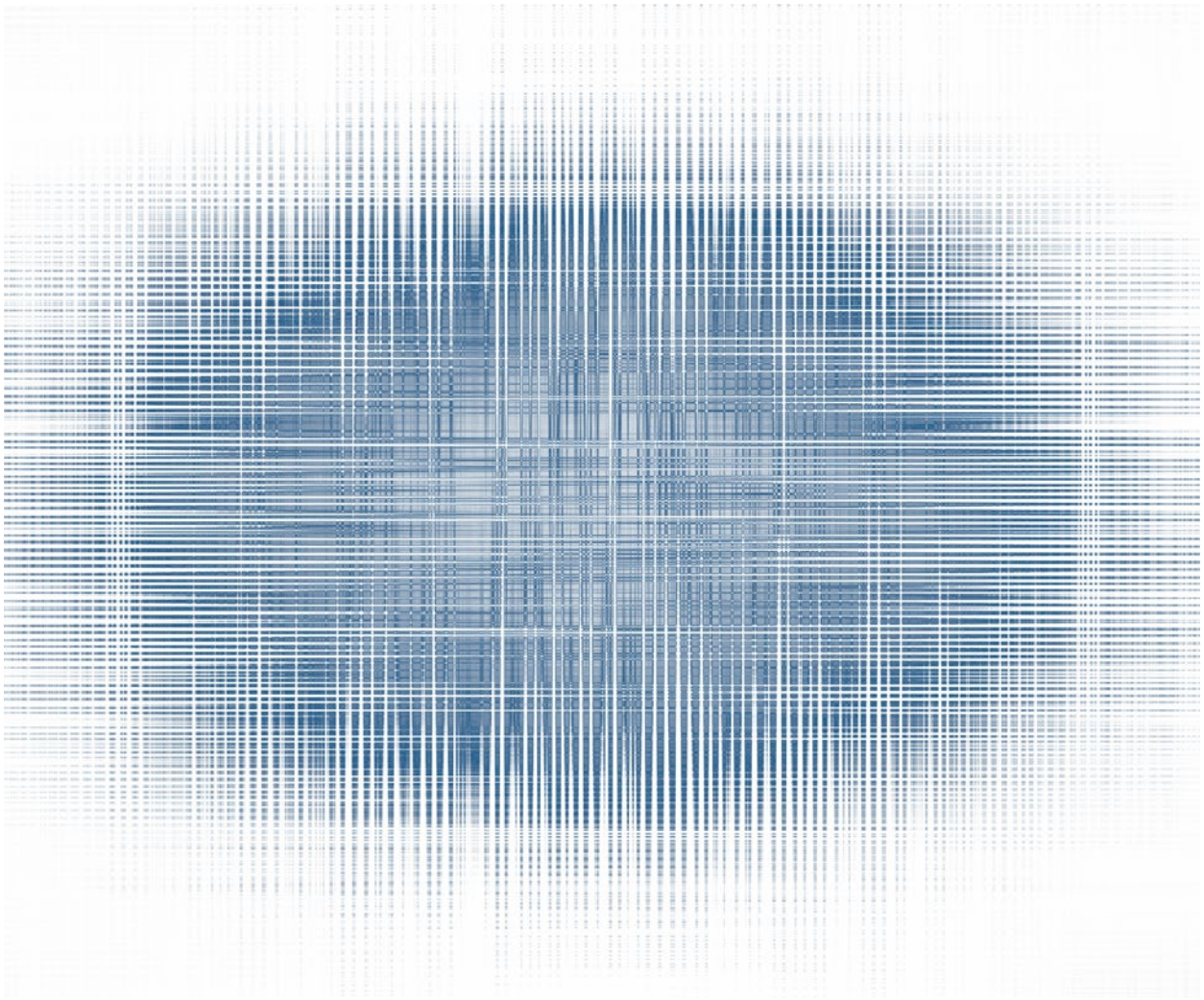
Australian National University
CENTRE FOR SOCIAL RESEARCH & METHODS



Describing the top of the income distribution in Australia

N Biddle, R Breunig and F Markham

CSRM WORKING PAPER
NO. 4/2019

Series note

The ANU Centre for Social Research & Methods (CSRM) was established in 2015 to provide national leadership in the study of Australian society. CSRM has a strategic focus on:

- development of social research methods
- analysis of social issues and policy
- training in social science methods
- providing access to social scientific data.

CSRM publications are produced to enable widespread discussion and comment, and are available for free download from the CSRM website (http://csrm.cass.anu.edu.au/research/publications).

CSRM is located within the Research School of Social Sciences in the College of Arts & Social Sciences at the Australian National University (ANU). The centre is a joint initiative between the Social Research Centre and the ANU. Its expertise includes quantitative, qualitative and experimental research methodologies; public opinion and behaviour measurement; survey design; data collection and analysis; data archiving and management; and professional education in social research methods.

As with all CSRM publications, the views expressed in this Working Paper are those of the authors and do not reflect any official CSRM position.

Professor Matthew Gray
Director, ANU Centre for Social Research & Methods
Research School of Social Sciences
College of Arts & Social Sciences
The Australian National University
June 2019

Working Paper No. 4/2019

ISSN 2209-1858
ISBN 978-1-925715-22-4

An electronic publication downloaded from http://csrm.cass.anu.edu.au/research/publications.

For a complete list of CSRM working papers, see http://csrm.cass.anu.edu.au/research/publications/working-papers.

ANU Centre for Social Research & Methods
Research School of Social Sciences
The Australian National University

Describing the top of the income distribution in Australia

N Biddle, R Breunig and F Markham

Nicholas Biddle is an Associate Professor and Associate Director of the Centre for Social Research & Methods, Research School of Social Sciences, College of Arts & Social Sciences, Australian National University. He is also a Fellow at the Tax and Transfer Policy Institute, Crawford School of Public Policy, College of Arts & Social Sciences, Australian National University.

Robert Breunig is the Director of the Tax and Transfer Policy Institute.

Francis Markham is a Research Fellow at the Centre for Aboriginal Economic Policy Research, College of Arts & Social Sciences, Australian National University.

Abstract

In this paper, we use a new source of linked Australian census, tax, social security and Medicare data to analyse the characteristics of those who were at the very top of the income distribution in 2011. The Basic Longitudinal Extract 2011 (BLE2011), from the Multi-Agency Data Integration Project (MADIP) overcomes a number of limitations of previous datasets. In addition to providing tax data for a very large proportion of the adult population, it combines census data linked at the individual level. Importantly, it has a household identifier, which allows us to calculate the distribution of equivalised household income, as well as the distribution of individual income. We show that there is quite substantial movement in and out of the top of the income distribution, depending on whether we use individual or household data. Furthermore, despite some assumptions to the contrary in the popular discourse, a much higher proportion of people at the top of the equivalised household taxable income distribution are professionals, as opposed to managers. Finally, although receipt of social security is quite low at the very top of the income distribution, a non-negligible number of people in the top 2% of the income distribution still received some form of payment or allowance; the most common payments were Carer Allowance, the Seniors Health Card, the Age Pension, and Family Tax Benefit Part B and Part A (in that order).

Acknowledgments

The data for this paper were provided through the Australian Bureau of Statistics (ABS) Virtual Data Laboratory. The authors would like to thank the ABS for the collegial approach to data access and recognise the considerable efforts to make data available in a way that maintains data privacy.

Acronyms

ANU	Australian National University
BLE2011	Basic Longitudinal Extract 2011
CEO	chief executive officer
CSRM	ANU Centre for Social Research & Methods
MADIP	Multi-Agency Data Integration Project
PIT	personal income tax
WTID	World Top Incomes Database

Contents

Tables and figures

1 Why are we interested in the top of the income distribution?

Politically, the most powerful framing of the issue of income inequality has come from the Occupy movement and the associated slogan 'We are the 99%' (Klein 2011). Sociologist Todd Gitlin (2012) attributes the origin of the slogan to widely circulated articles on alternative media websites discussing the findings of Piketty and Saez on the income share of the top 1% in the United States (Piketty & Saez 2003, Saez 2009) and calling for a movement of the 99%. Also influential was Professor Joseph Stiglitz's 2011 article 'Of the 1%, by the 1%, for the 1%' in Vanity Fair. In the opening paragraph, Stiglitz states that:

> *The upper 1 percent of Americans are now taking in nearly a quarter of the nation's income every year. In terms of wealth rather than income, the top 1 percent control 40 percent. Their lot in life has improved considerably. Twenty-five years ago, the corresponding figures were 12 percent and 33 percent.*

Academically, Thomas Piketty (2014) in *Capital in the 21st century* presented a huge amount of historical and contemporary data across a number of countries (but predominantly the United States, the United Kingdom and France) on the concentration of income and wealth. This book built on previous work by Atkinson and Piketty (2007, 2010) and on tax (and other) data for a range of countries and time periods compiled in the World Top Incomes Database (WTID).

The quantitative distinction between the 1% and the 99% diverges from the qualitative distinction between capital owners and workers found in classical political economy (e.g. Veblen, Marx and Weber). A fundamental insight of these classical approaches is that, at the population level, the owners of capital engage in the economy and society in categorically different ways to workers. In this understanding, the differences between the economic elites and the masses are not just differences in degrees of income and wealth, but also differences in social power. Simplistic categorical distinctions are empirically difficult to sustain, however, in an age when many of the highest income earners are employees (e.g. chief executive officers [CEOs]) and in which compulsory superannuation has made many of the poorest workers owners of capital (Wright 1997). As such, quantitatively drawn distinctions between the top 1% and the rest can produce similar insights to the previous qualitative distinctions between the lives of the owners of capital and labourers, but without relying on defining these empirically intertwined categories.

The focus on the top of the distribution when analysing income inequality has been criticised recently by Richard Reeves (2018), who argued that this distracts from the impacts that the tax, education and planning (among other) systems have on maintaining the position of those in the more numerous top quintile of the income distribution. Likewise, in its review of income inequality in Australia, the Productivity Commission (2018) also focused on inequality across the distribution, rather than differences between those at the very top. Echoing some of these criticisms, the WTID has recently been renamed the World Inequality Database and now has a much greater focus on inequality across the distribution, as well as the measurement and tracking of wealth inequality. We would argue though, and the data presented in this paper would support this, that inequality across the distribution and differences between the very top and the rest of the distribution are complements rather than substitutes.

The seminal article on the very top of the income distribution in Australia was written by Tony Atkinson and Andrew Leigh in 2007. They argued that 'the share of income accruing to the very top groups is of importance both because their share of the total is significant and on account of

the economic power that it conveys'. They also showed that, like in the United States, the United Kingdom, Canada and New Zealand, there was:

> *… a decline in top income shares in the three decades after the Second World War, followed by a sharp rise from the mid-1970s onwards. At the start of the twenty-first century, the income share of the richest 1 per cent of Australians was higher than it had been at any point since 1951, while the share of the richest 10 per cent was higher than it had been since 1949.*

Although we have a reasonable idea of the size and income share at the very top of the income distribution, we know very little about the characteristics of those at the top, their stability through time and how the extremely high income group (the top 1% in our framing below) compares with just the very high income group (the 2nd top percentile) or the high-income group (the 81st to 98th percentile).

One constraint on academic understanding of those at the top of the income distribution in Australia is a relative paucity of data. In a recent article building on the work of Atkinson and Leigh (2007), and Burkhauser et al. (2015), Burkhauser et al. (2018) compare and contrast the data situation in Australia with other comparable countries, as well as our knowledge of the income distribution. They summarise the situation as follows:

> *Access to contemporary unit record tax data from the entire population of tax filers in the United States and in the United Kingdom has improved the ability of researchers to capture the share of income held by top income groups. In Australia, we find that this added value from tax records data is limited by our inability to capture the top 1 per cent of the income distribution in the 1 per cent unit record tax samples the ATO has provided researchers for 2003 through to 2010. Although we find that the 2 per cent tax samples made available since then provide more accurate estimates of gross taxable income held by top income groups, they still contain censored and perturbed income variables that affect their ability to provide a complete picture of the top part of the income distribution.*

The aim of this paper is to use the type of unit record data that Burkhauser et al. (2018) are referring to, to understand the characteristics of those at the top of the income distribution. Section 2 of the paper describes the data that we use, with a focus on the income variables. The substantive empirical results are then presented across separate sections as follows:

- In Section 3, we break the population down into six groups across the income distribution, and present the average income and income bands for those groups.
- In Section 4, we look at the relationship between a person's position on the individual and household (equivalised) income distribution.
- In Section 5, we look at the demographic characteristics of the six income groups.
- In Section 6, we look at the occupation and industry of those who are employed.
- In Section 7, we look at receipt of social security.

The final section of the paper (Section 8) provides some concluding comments, and an outline of future work on the dataset.

When thinking about the concentration of income and wealth in the top 1% or indeed any share of the distribution, it is important to keep in mind the wider economic context. The past 50 years, which have seen this increased concentration of income and wealth in the top 1% (Picketty 2014), have simultaneously seen a large drop in poverty both in Australia and more dramatically across countries, and a large increase in wellbeing, irrespective of how it is measured. Australia and the world have seen dramatically improved health, increased life expectancy, large increases in education, and a decrease in global inequality as poor and rich countries have converged (at least for some poor countries).

Popular discourse seems to have an unstated counterfactual that the income and wealth that are controlled or received by the top 1% would have been generated in the same quantity even had they flowed to the rest of the distribution. This is a very important policy question, but not one that is answerable by the data that we have access to. While our paper investigates the

characteristics of those in the top 1% in Australia, the paper has nothing to say about whether the current income distribution is optimal using other metrics (such as subjective wellbeing).

2 Data – the Multi-Agency Data Integration Project and the Basic Longitudinal Extract 2011

To understand the characteristics of those at the top of the income distribution in Australia, it is necessary to develop and use new datasets with a large sample of individuals that include sociodemographic and household information alongside a person's income. Ideally, and this will be the focus of a subsequent paper, this dataset would have longitudinal information across the distribution.

2.1 Describing the Basic Longitudinal Extract 2011

A recently developed dataset that meets these criteria is the Multi-Agency Data Integration Project (MADIP) Basic Longitudinal Extract 2011 (BLE2011). According to the summary provided by the Australian Bureau of Statistics (2018), the BLE2011 is built around a full cohort of the Australian population in 2011[1] and:

> *… includes key demographic, social, health care, government payment and income information for this population over the period 2011–2016. The microdata product contains approximately 22.5 million records and 122 data items; 74 of these data items have information for multiple years to enable longitudinal analysis.*

Four sources of data in the BLE2011 have been linked at the individual level:

- the Medicare Enrolments Database and Medicare Benefits Schedule data, which include information on the number of services used, benefits paid and fees charged
- personal income tax (PIT) data, which include information on wages and salaries; government allowances, pensions and payments; total income (summation of the previous, as well as other forms of income); and taxable income
- Social Security and Related Information data, which include information on whether a person was receiving payments at a September snapshot of 28 separate payments, and whether their partner was receiving any of the payments
- a subset of data from the 2011 Census of Population and Housing (2011 Census), which includes a household identifier, ethnicity, country of birth, education levels and participation, employment status (including hours worked, industry and occupation), disability status and income.

In addition to data from the individual datasets, a small number of derived variables combine information from multiple datasets. These are age, sex, Indigenous status and geographic location as of 2011. The latter is available at the level of Statistical Area Level 1, a very detailed level of geography.[2]

Table 1 summarises the years and data available in the BLE2011.

In a recent paper analysing the validity of the dataset for understanding income dynamics in Australia (Biddle et al. 2019), we concluded that 'the BLE2011 has the potential to shed new light on the determinants, dynamics, and distribution of income in Australia. However, analysis of the dataset should be carried out with caution and taking into account some of the limitations'. Some of the limitations identified were that the BLE2011 has tax data only on slightly over half of the sample and that around one-third of the sample reported income in PIT data either above or below their reported census income category (in a roughly even split). There are, however, consistent predictors for being outside that range, suggesting that these differences are often systematic rather than errors. Finally, those with linked census information on the BLE2011 have similar, but not completely identical, characteristics to those on the full 2011 Census.

2.2 Income variables in the MADIP

A number of income variables are available or can be calculated from the BLE2011. Some of these are available only for a single cross-section, with others available for 6 financial years. These are listed below, based on source and availability:

- Census based
  - Total income bands. Available for 2011 only, based on the question 'What is the total of all wages/salaries, government benefits, pensions, allowances and other income the person usually receives?' Results are available in 12 categories, including one for negative income and another for zero income.

Table 1 Datasets and years available in the MADIP Basic Longitudinal Extract 2011

Year	Census	Medicare	PIT	SSRI
2010	na	na	2010–11 financial year	na
2011	2011 Cross-sectional	2011 calendar year services and benefits	2011–12 financial year	September 2011 payment type
2012	na	2012 calendar year services and benefits	2012–13 financial year	September 2012 payment type
2013	na	2013 calendar year services and benefits	2013–14 financial year	September 2013 payment type
2014	na	2014 calendar year services and benefits	2014–15 financial year	September 2014 payment type
2015	na	2015 calendar year services and benefits	2015–16 financial year	September 2015 payment type
2016	na	2016 calendar year services and benefits	na	September 2016 payment type

MADIP = Multi-Agency Data Integration Project; na = not applicable; PIT = personal income tax; SSRI = Social Security and Related Information

- PIT based – available for all financial years in $1000 increments (including negative income and zero income categories)
  - Salary and wages. Top-coded at $250 000 or more, with no individuals in the negative income category. Corresponds to Item 1 (sum of labels C, D, E, F and G) on the Australian Taxation Office's (ATO's) Tax Return for Individuals.
  - Australian Government allowances and payments. Top-coded at $20 000 or more, with no individuals in the negative income category. Corresponds to Item 5, label A, on the ATO's Tax Return for Individuals.
  - Australian Government pensions and allowances, with no individuals in the negative income category. Top-coded at $20 000 or more. Corresponds to Item 6, label B on the ATO's Tax Return for Individuals.
  - Total income (or loss). Top-coded at $250 000 or more. Corresponds to the value at the bottom of page 3 on the ATO's Tax Return for Individuals.
  - Taxable income (or loss). Top-coded at $250 000 or more. Corresponds to the value stated at the bottom of page 4 on the ATO's Tax Return for Individuals.

The analysis presented in this paper does not focus on the census-based income. Because we are most interested in distinctions among those at the top of the income distribution, we assign those who do not have PIT data present for a particular year a value of zero for taxable income.

We aggregate income across the household, based on the census household identifier in 2011. To calculate equivalised income,[3] we assume the second and subsequent adults in the household cost 0.5 times the amount of the first adult, and children 15 years and under cost 0.3 times the amount of the first adult. One limitation of the BLE2011 is that we only have a household identifier for 2011. We are unable, therefore, to capture changes in access to economic resources through time that are due to changes in household size or composition.

In our analysis, we focus on salaries and wages for individuals and households (equivalised), as well as taxable income for individuals and households (equivalised). That is, four income measures in total.

3 Segmenting the income distribution

For the four variables described above, we break the adult population into five groups across the income distributions. Given the focus of this paper, we segment the top part of the income distribution more finely than we do the bottom or the middle, with the latter a focus of future papers using the BLE2011. These groups are:

- low income – the bottom 20% of the income distribution, including those with zero income
- middle income – the middle 60% of the income distribution, or the 21st to 80th percentiles; we separate this group into the low–middle and the middle–high part of the income distribution, with two groups of 30 percentiles on either side of the median
- high income – the 81st percentile to the 98th percentile of the income distribution
- very high income – the 99th percentile of the income distribution
- extremely high income – the 100th percentile, or the top 1% of the income distribution.

We assign an income value of zero to those who did not have a PIT record, rather than excluding them from our analysis. However, because we were only able to calculate equivalised household income for those with a linked 2011 Census record, those not linked to the census were excluded from the analysis of equivalised income but included in the analysis of individual income. Finally, although they are included in the calculation of equivalised household income, we excluded children from our percentile calculations.

With the above inclusions and exclusions in mind, 177 000–181 000 individuals were in each percentile for the individual income measures, and 107 000–108 000 individuals were in each percentile for our equivalised household measures. Table 2 gives the mean income values for each of the six groups for the four income measures used in our analysis, as well as the number of people in that group.

For the first three of the four income types, the bottom quintile of the income distribution on the PIT system is made up entirely of those with zero taxable income. A small number of adults in the bottom quintile of the equivalised household taxable income distribution have a non-zero income but they make up a very small minority, with the average income in the quintile still very low ($255). We will describe this group in Section 4.

Looking at the middle part of the income distribution (the 2nd, 3rd and 4th quintiles), taxable income is about 5.6% higher than salary income for individuals, with the gap a fair bit greater for equivalised income (13.8%). Finally, for both salary income and taxable income in this group, equivalised household income is higher than individual income. This shows that the assumed level of sharing of resources across children in the household is more than counterbalanced by the pooling of income across households.

The PIT data in the BLE2011 are not ideal for looking at the average income of those at the bottom, and to a lesser extent the middle part, of the income distribution. In Biddle et al. (2019), we showed that for the first three income bands in the census ($1 to $20 799 per year) between 50% and 60% of the sample were not in the tax system. For the next income band ($20 800–31 199) more than one-third of respondents were not in the tax system. However, at the other end of the distribution, the vast majority of people (more than 95%) of those in upper income bands can be linked to records in the tax system. The last three groups in Table 2 are therefore much better captured in the BLE2011 than the rest of the distribution (hence the focus in this paper).

Looking first at the high-income (but not very or extremely high income) group, we have information on around 3.2 million people for individual income and around 2 million for equivalised income. These groups had an average

Table 2 **Mean income and number of people by income type and income group, 2010–11 financial year**

	Individual wage/ salary		Equivalised wage/ salary		Individual taxable income		Equivalised taxable income	
Income group	Mean (A$2016)	*n*	Mean (A$2016)	*n*	Mean (A$2016)	*n*	Mean (A$2016)	*n*
Low income	0	8 577 263	0	3 088 614	0	6 196 502	255	2 174 967
Low–middle income	2 713	572 964	14 991	2 349 559	12 300	3 052 143	22 564	3 261 431
Middle–high income	32 741	5 504 133	45 639	3 251 548	42 382	5 442 956	52 114	3 254 100
High income	86 183	3 236 292	86 171	1 956 921	93 931	3 193 226	95 501	1 955 086
Very high income	163 468	176 835	141 331	107 117	195 800	181 494	172 715	107 807
Extremely high incomea	236 576	180 161	189 803	108 044	265 781	181 327	241 933	108 412
Total	29 166	18 247 648	35 712	10 861 803	35 725	18 247 648	43 758	10 861 803

a The mean income for this group is an underestimate due to the top-coding of income in the BLE2011.

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

salary income of around $86 000 (regardless of whether it was measured as individual or equivalised household), an average individual taxable income of around $94 000 and a slightly higher equivalised household taxable income of around $96 000.

The second last group in the table, which we have labelled as the very high income group, have a much larger gap between salary income ($162 000 for individuals and $141 000 equivalised) and taxable income ($196 000 for individuals and $173 000 equivalised) than the high-income and middle-income groups. Although we do not show it in Table 2, only a small share of this additional income comes from transfer payments. Rather, a much larger percentage comes from other sources, including interest payments, dividends, rental income, income or loss from a business, capital gains and annuity/superannuation income.

Data in BLE2011 are top-coded at $250 000 per year, or $274 237 when converted to 2016 dollars. The lower bound for the top 1% of the income distribution in the PIT data from BLE2011 ranges from $185 562 (in 2016 dollars) for individual salary income to $192 191 for equivalised taxable income. That means that a non-negligible proportion of the extremely high income group (33.4% for individual salary income and 67.3% for individual taxable income) are top-coded and we do not know their true income. Given the distribution of incomes within this group, the mean income for the extremely high income group calculated from the data in the BLE2011 is therefore an underestimate of the population value.

4 Relationship between household and individual income

One of the main benefits of the BLE2011 for understanding income and its distribution in Australia is the ability to incorporate census household information to calculate someone's position on the distribution for both household and individual income. Although other datasets do this (in particular, Household, Income and Labour Dynamics in Australia, and the Survey of Income and Housing), the much larger BLE2011 sample produces a major reduction in sampling error relative to other datasets.

There are a number of potential reasons why a person might have a very different position on the income distribution based on their household compared with their individual income (and vice versa). Someone might have a higher position in terms of household income compared with individual income if their partner or other household members contribute a significant amount to the household budget. On the other hand, they may have a lower position if they need to share their own income across a number of people in their household who do not bring in an income themselves.

In Table 3, we cross tabulate a person's position on the income distribution based on their individual income against their position on the distribution based on their equivalised household income. Given the focus of this paper (and the comparative advantage of the dataset), we collapse the first three groups into one, and focus on variation across and within the top quintile of

Table 3 Comparison of individual and equivalised household income distributions – total

Individual salary income	Equivalised household salary income				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	7 900 070	798 183	25 927	17 445	8 741 625
High	786 566	1 054 393	52 923	28 169	1 922 051
Very high	2 948	62 272	11 484	20 513	97 217
Extremely high	0[a]	42 073	16 783	41 917	100 773
Total	8 689 584	1 956 921	107 117	108 044	10 861 666
Individual taxable income	**Equivalised household taxable income**				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	7 945 570	800 711	23 429	10 751	8 780 461
High	744 189	1 068 418	45 469	22 878	1 880 954
Very high	722	52 828	20 314	25 914	99 778
Extremely high	0[a]	33 129	18 595	48 869	100 593
Total	8 690 481	1 955 086	107 807	108 412	10 861 786

a Set to 0 to avoid the privacy risk of disclosing the very small number of individuals in this table cell

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

the distribution. We present the results separately for salary income and taxable income.

Looking first at differences in the composition of the top quintile, there were large differences in both directions. Around 18.6% of those who were in the top quintile based on their individual salary income were outside that group based on their equivalised household salary income (the blue-shaded cells divided by the blue and white cells in the top half of the table). This is slightly higher than the proportion in the bottom half of the table, with 17.9% of people being in the bottom four quintiles for equivalised household taxable income but in the top quintile for taxable individual income. There were much higher proportions in the opposite direction (the pink shaded cells divided by the pink and white cells), or those who were in the top quintile of the distribution for equivalised income but not individual income (38.7% for salary income and 38.5% for taxable income).

Not surprisingly, the vast majority of those who were not in the top quintile based on their equivalised income but who were in the top quintile based on their individual income had high individual income, as opposed to very or extremely high income (i.e. they were in the 81st to 98th percentile, rather than the 99th or 100th percentile). For wages and salaries, 2.1% of those who moved from the bottom 80% of the distribution on individual income into the top 20% for equivalised income moved into the top 1%, with an even smaller proportion doing so for taxable income (1.3%).

Finally, looking at those in the top 1% of the income distribution based on their individual salary income, a negligible proportion were outside the top quintile, and more than half (58.4%) were in the 81st to 99th percentile based on their equivalised household salary income. There was less variation by taxable income, with 51.4% of the top 1% of the taxable individual income being in the 81st to 99th percentile based on their equivalised taxable income.

The above percentages are very different for males and females (Tables 4 and 5). Focusing on taxable income, around one-fifth of males (20.1%) were in the top quintile for individual

Table 4 Comparison of individual and equivalised household income distributions – males

Individual salary income	Equivalised household salary income				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	3 411 327	254 006	6 272	3 922	3 675 527
High	561 303	636 976	26 968	10 107	1 235 354
Very high	2 791	55 737	8 761	13 364	80 653
Extremely high	0[a]	39 273	14 786	31 819	85 878
Total	3 975 421	985 992	56 787	59 212	5 077 412
Individual taxable income	**Equivalised household taxable income**				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	3 422 431	271 060	6 990	3 563	3 704 044
High	551 820	643 253	20 669	8 111	1 223 853
Very high	679	44 899	13 079	13 434	72 091
Extremely high	0[a]	30 381	15 303	31 857	77 541
Total	3 974 930	989 593	56 041	56 965	5 077 529

a Set to 0 to avoid the privacy risk of disclosing the very small number of individuals in this table cell

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

Table 5 Comparison of individual and equivalised household income distributions – females

Individual salary income	Equivalised household salary income				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	4 488 743	544 177	19 655	13 523	5 066 098
High	225 263	417 417	25 955	18 062	686 697
Very high	157	6 535	2 723	7 149	16 564
Extremely high	0[a]	2 800	1 997	10 098	14 895
Total	4 714 163	970 929	50 330	48 832	5 784 254
Individual taxable income	**Equivalised household taxable income**				
	Low–middle	High	Very high	Extremely high	Total
Low–middle	4 523 139	529 651	16 439	7 188	5 076 417
High	192 369	425 165	24 800	14 767	657 101
Very high	43	7 929	7 235	12 480	27 687
Extremely high	0[a]	2 748	3 292	17 012	23 052
Total	4 715 551	965 493	51 766	51 447	5 784 257

a Set to 0 to avoid the privacy risk of disclosing the very small number of individuals in this table cell

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

income but the bottom 80% of the distribution for equivalised income. For females, this percentage falls to 13.6%. Even more dramatically, more than half of females were in the top quintile for equivalised income (51.8%) but not in the top quintile for individual income, more than double the percentage for males (25.5%). Finally, there was less movement within the top quintile across equivalised and individual income for females than for males. Specifically, 26.2% of females who were in the top 1% of the taxable individual income were in the 81st to 99th per centile based on their equivalised taxable income, compared with 58.9% for males.

5 Characteristics of those across the income distribution

In addition to a person's household context, a range of demographic and geographic characteristics in the census allow us to test for differences across the income distribution. Although census data are not as rich as data from sample surveys, this is compensated for by a much greater sample size. In Tables 6–7, we look at eight characteristics of individuals for each of the six income groups and each of the four income measures:

- average age
- proportion female
- proportion born overseas
- proportion who have completed year 12
- proportion with a degree
- proportion who live in a house that is owned or being purchased
- proportion employed
- average percentage of the person's income from wages and salaries.

Beginning with the three lower points on the income distribution, those with low (mostly zero) income tend to be much older than all other income groups, are far more likely to be female (at least with regard to taxable income) and are much more likely to have been born overseas. These individuals also have lower levels of education and are less likely to own their own home. Not surprisingly, they are far less likely to be employed and, for those who have an income, wages and salaries make up a very small share of their overall income.

The middle three income quintiles have some interesting differences. Those below the median (low–middle income) are slightly older, more likely to be female and more likely to have been born overseas than those above the median, but not in the high-income category. They are, however, far less likely to be employed, and have a very low percentage of their income coming from wages and salaries (only around half of their income for those in the low–middle income group). Many of these individuals are likely to be retirees.

Comparison between those in the top income quintile and these other three groups is all that is usually possible with survey data. It is sometimes possible to split the top income quintile by half and look at the top two deciles separately, but these comparisons tend to be affected by large standard errors. The BLE2011, on the other hand, allows us to look within the top of the income distribution and compare those who have high income and those with very or extremely high income. Some differences clearly emerge.

Compared with those with high incomes only, those with very high incomes (and even more so those with extremely high incomes) are older, more likely to be male and somewhat more likely to have completed year 12. In fact, they are far more likely to have a degree as their highest qualification. Interestingly though, when we look at equivalised household taxable income, those in the very high and extremely high income groups (the top 2% of the income distribution) are somewhat less likely to be employed.

Perhaps the biggest difference between those in the top 1% of the income distribution and those in the next highest income percentile (based on equivalised household taxable income) is the percentage of that person's income that comes from wages and salaries. Although both are lower than for the rest of the high-income group, those in the top 1% of the income distribution receive roughly 10 percentage points less of their income from wages and salaries than those in the rest of the top 2%. Indeed, they receive far less than half of their income from wages and salaries and a lower percentage than the low–middle income group.

Table 6 **Census-based characteristics across the income distribution – individual and equivalised household salary income**

Characteristic	Low income	Low–middle income	Middle–high income	High income	Very high income	Extremely high income
	Individual salary income					
Age	49.2	34.2	36.5	39.9	41.6	43.8
Female	0.527	0.598	0.579	0.349	0.167	0.149
Born overseas	0.324	0.226	0.257	0.280	0.310	0.335
Completed year 12	0.360	0.555	0.568	0.668	0.772	0.877
Has a degree	0.114	0.182	0.187	0.384	0.547	0.711
Home owner	0.739	0.723	0.731	0.813	0.860	0.890
Employed	0.222	0.536	0.884	0.977	0.980	0.980
Salary as percentage of taxable income	0.0	46.4	83.0	82.2	72.9	79.9
	Equivalised household salary income					
Age	58.1	39.7	38.7	38.9	40.0	42.0
Female	0.553	0.554	0.524	0.496	0.470	0.452
Born overseas	0.330	0.291	0.274	0.267	0.305	0.329
Completed year 12	0.336	0.474	0.541	0.663	0.788	0.843
Has a degree	0.108	0.145	0.199	0.351	0.540	0.638
Home owner	0.721	0.681	0.769	0.842	0.866	0.861
Employed	0.177	0.589	0.781	0.889	0.896	0.919
Salary as percentage of taxable income	0.0	56.6	79.1	79.6	73.5	74.8

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

Table 7 **Census-based characteristics across the income distribution – individual and equivalised household taxable income**

Characteristic	Low income	Low–middle income	Middle–high income	High income	Very high income	Extremely high income
Individual taxable income						
Age	47.2	40.0	40.7	41.8	44.9	47.8
Female	0.551	0.585	0.524	0.342	0.270	0.227
Born overseas	0.330	0.266	0.276	0.278	0.296	0.308
Completed year 12	0.318	0.522	0.541	0.669	0.786	0.849
Has a degree	0.081	0.157	0.198	0.395	0.564	0.680
Home owner	0.685	0.766	0.752	0.837	0.897	0.922
Employed	0.146	0.539	0.853	0.940	0.891	0.884
Salary as percentage of taxable income	7.6	45.5	70.5	73.0	54.0	57.4
Equivalised household taxable income						
Age	57.7	42.5	40.2	40.5	43.5	46.8
Female	0.570	0.550	0.516	0.494	0.480	0.475
Born overseas	0.336	0.304	0.271	0.266	0.288	0.284
Completed year 12	0.298	0.456	0.545	0.667	0.770	0.778
Has a degree	0.082	0.136	0.209	0.366	0.529	0.550
Home owner	0.658	0.690	0.797	0.863	0.900	0.919
Employed	0.119	0.537	0.782	0.858	0.808	0.794
Salary as percentage of taxable income	17.0	51.6	72.0	71.8	55.9	45.6

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

6 Occupation and industry at the top of the income distribution

6.1 Managers and professionals at the top of the income distribution

In discussing the historical change in the share of income for the top of the income distribution, Atkinson and Leigh (2007) conclude with the point that 'The rapid rise in Australian CEO salaries during the 1990s suggests that much of this recent increase was caused by higher executive pay, possibly driven by the internationalisation of the market for CEOs'. However, this assumes that most of those at the very top of the income distribution are CEOs. To test for this, Figures 1 and 2 give the proportion of people in the six income groups who are identified (through the census) as being either a manager or a professional.[4]

Looking at individual wages and salary income, the most common occupation category for the very high and extremely high income groups is managers (the stereotypical CEO). However, less than half of both income groups are managers, or 36.1% for the second highest income percentile and 44.7% for the top income percentile. By comparison, a little over one-third (35.1%) of the very high income group were employed

Figure 1 **Percentage of population who are managers or professionals, by income group (individual, wages/salaries)**

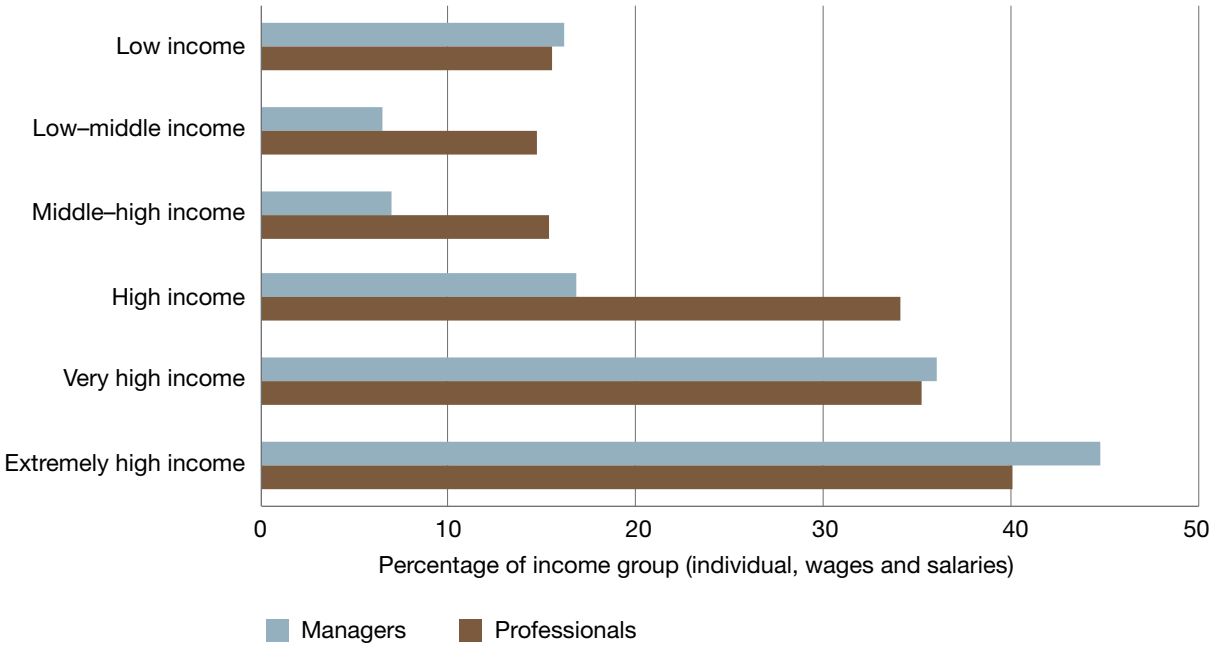
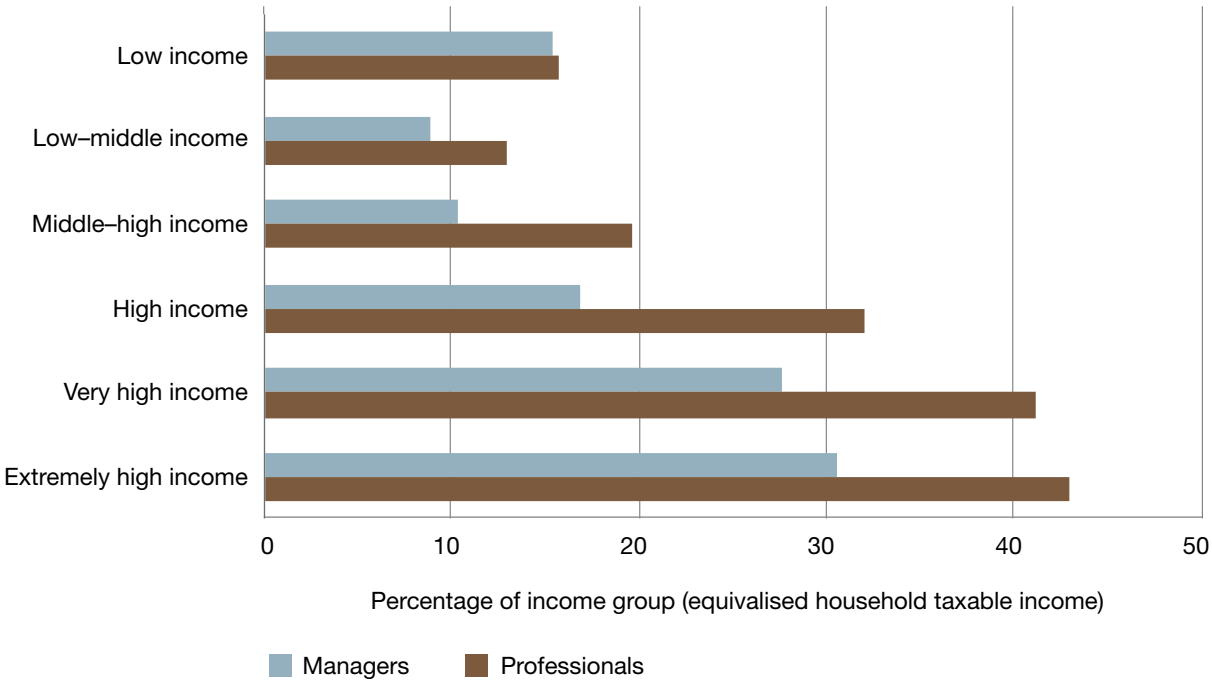


Figure 2 **Percentage of population who are managers or professionals, by income group (equivalised household taxable income)**



as professionals, alongside around two-fifths (40.0%) of the extremely high income group.

Professionals are more dominant, however, when we look at equivalised household taxable income. For this income type, less than one-third of the two top income categories are managers (27.6% for the very high income group and 30.6% for the extremely high income group), falling to around one-sixth (17.0%) of the high-income group. By comparison, the largest occupation category for those at the top of the equivalised household taxable income group is professionals, with 32.1%, 41.2% and 43.0% of the top three income groups, respectively (high, very high and extremely high).

Some of the narrative around the top of the income distribution is that they are all CEOs. This is true to a certain extent, and, when we look at a more disaggregated measure of occupation (the four-digit level), CEOs and managing directors are the most common occupation at the very top of the income distribution. However, as shown in Table 8, the top 20 occupations in the top 1% of the taxable household income distribution are equally dominated by professionals (blue shading) and managers (red shading).

Data from the BLE2011 have shown (for the first time we think in Australia) that the top of the household income distribution is as likely to be academics, lawyers and doctors as it is to be CEOs or other managers. It is likely that there is a clear hierarchy within the top 1% of the income distribution and that CEOs are the richest of the extremely rich. However, taken as a group, the top 1% is more occupationally diverse than we often assume.

6.2 Industries that contain the top 1% of the income distribution

A person's occupation is what they do as an individual while they are working. The industry they work in, however, refers more to what their company does. So, for example, a person can drive a truck for a living. But they can do that in many different industries, including, most obviously, transport, but also mining, construction or defence. As shown by Sullivan (2010), both industry and occupation are important for explaining human capital and other outcomes. Table 9 shows that there is also variation in the probability of being at the top of the income

Table 8 **Top 20 occupations in the top 1% of the equivalised household taxable income distribution**

Occupation	Individual, wage/ salary rank	Occupation	Taxable equivalised rank
Chief executives and managing directors	1	Chief executives and managing directors	1
Advertising, public relations and sales managers	2	General practitioners and resident medical officers	2
General managers	3	Solicitors	3
Finance managers	4	Accountants	4
General practitioners and resident medical officers	5	Advertising, public relations and sales managers	5
ICT managers	6	General managers	6
Financial brokers	7	Finance managers	7
Solicitors	8	General clerks	8
Production managers	9	Retail managers	9
Financial investment advisers and managers	10	Financial investment advisers and managers	10
Contract, program and project administrators	11	Construction managers	11
Engineering managers	12	Office managers	12
Human resource managers	13	Other medical practitioners	13
Financial dealers	14	Real estate sales agents	14
Construction managers	15	ICT managers	15
Other hospitality, retail and service managers	16	Management and organisation analysts	16
Other medical practitioners	17	Human resource managers	17
Other specialist managers	18	Sales assistants (general)	18
Management and organisation analysts	19	Contract, program and project administrators	19
ICT sales professionals	20	Dental practitioners	20

distribution across the 19 industries at the most aggregated (2-digit) level of the Australian and New Zealand Standard Industrial Classification.[5]

The last line of Table 9 gives the percentages for all those for whom we have industry information (from the census). Given this does not include anyone who is not employed, it is not surprising that the values for the bottom row in the first two columns are greater than 20% and 1%, respectively. To identify which industries are disproportionately present at the top of the income distribution, it is useful to compare the specific industries with this last line.

Nine industries have a relatively high percentage of people at the top of the income distribution. Those with the biggest share in the top 20% are (in order) financial and insurance services; electricity, gas, water and waste services; professional, scientific and technical services; and public administration and safety. All of these industries have more than 40% of their workers in the top income quintile.

Table 9 **Presence in the top of the equivalised household taxable income distribution, by industry (2-digit)**

Industry	Top 20%	Top 1%	Top 1% for those in top 20%
Agriculture, forestry and fishing	14.2	1.0	6.8
Mining	39.0	1.9	4.9
Manufacturing	28.3	1.0	3.7
Electricity, gas, water and waste services	44.6	1.2	2.7
Construction	26.4	1.1	4.2
Wholesale trade	30.4	1.8	5.8
Retail trade	18.8	0.7	3.8
Accommodation and food services	14.3	0.5	3.2
Transport, postal and warehousing	26.5	0.7	2.7
Information media and telecommunications	39.1	1.8	4.5
Financial and insurance services	46.9	3.9	8.3
Rental, hiring and real estate services	32.1	2.5	7.7
Professional, scientific and technical services	43.3	3.4	7.8
Administrative and support services	20.8	1.0	4.6
Public administration and safety	40.7	0.8	2.0
Education and training	35.8	0.9	2.4
Health care and social assistance	27.3	1.7	6.2
Arts and recreation services	23.8	0.9	3.8
Other services	19.3	0.6	2.9
All those with industry information	29.2	1.4	4.6

Source: Customised data from the MADIP Basic Longitudinal Extract 2011

When we look at the very top of the income distribution (the top 1%), the story is quite different. Two of those industries (financial and insurance services; and professional, scientific and technical services) are still overrepresented. However, the other two (electricity, gas, water and waste services; and public administration and safety) are not. Instead, the industry with the next highest percentage at the extreme of the income distribution is rental, hiring and real estate services, despite having a very similar share in the top 20% as all industries combined.

The availability of linked census and tax data changes our understanding of the top of the income distribution from what we observe from survey data with much smaller sample sizes.

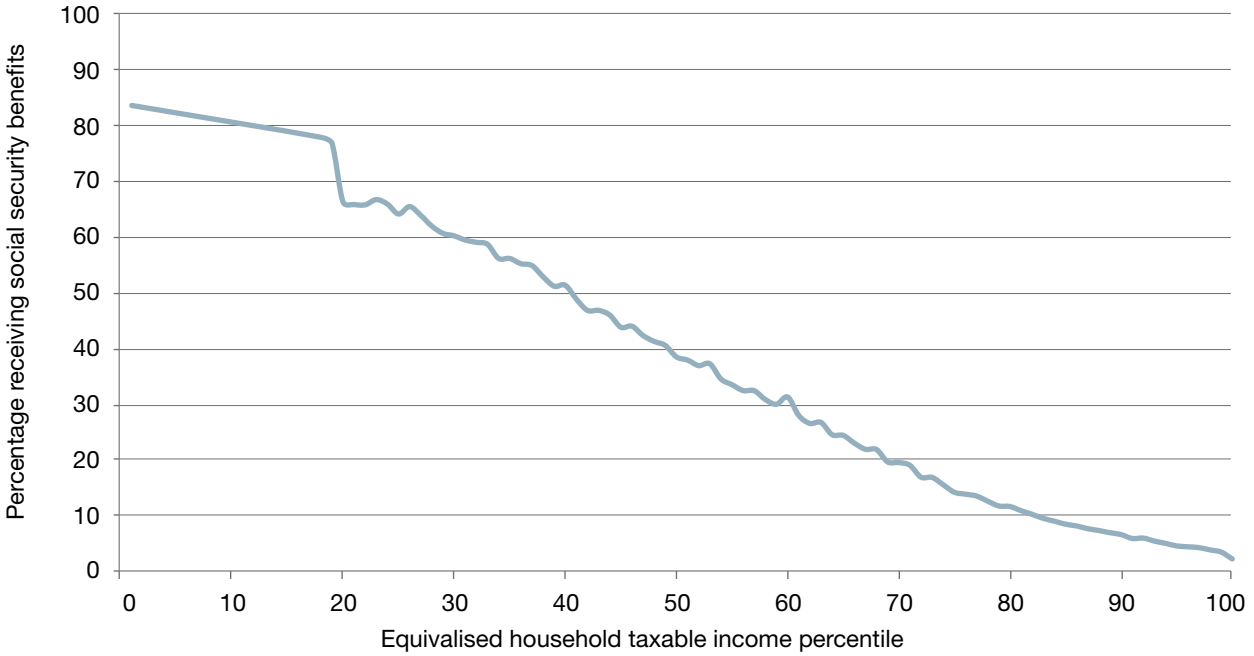
7 Social security receipt among the high-income group

A final variable that we analyse on the BLE2011 is whether or not the individual was receiving some form of social security payment in September 2011. When this is plotted for each percentile of the equivalised household taxable income distribution, the results show a fairly consistent decrease in receipt of social security payment across the income distribution (Figure 3). The proportion declines to around half by the 40th income percentile, and by the high-income cutoff used in this paper (the 81st percentile) it had declined to about 10.9%, or a little over 1 in 10. There was, however, still a decline across the remainder of the income distribution, with that group highlighted in Figure 4. For the very high income group (99th percentile), around 3.5% were in receipt of some form of social security, whereas for the extremely high income group this had declined to 2.3%.

Although these percentages are small, just over 6000 individuals in the top two income percentiles in the BLE2011 (based on 2010–11 taxable income) were receiving some form of social security in September 2011, most commonly Carer Allowance. This was followed by the Seniors Health Card, the Age Pension, and Family Tax Benefit Part B and Part A (in that order).

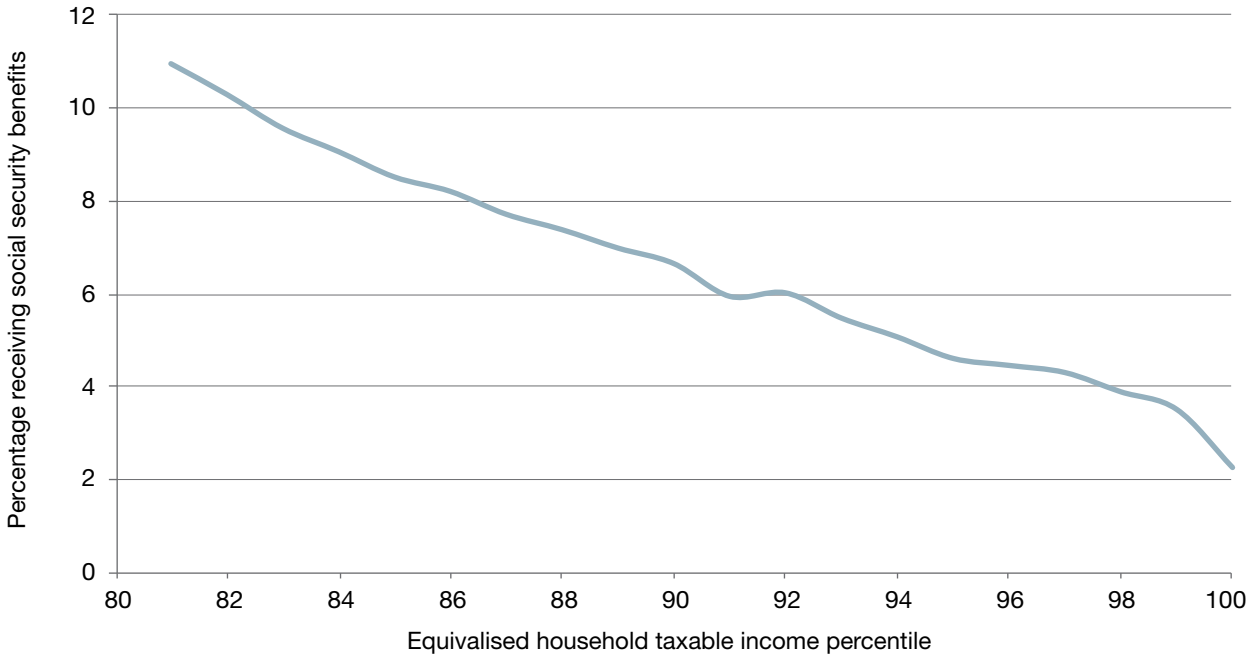
The most common payment types are quite similar for the top of the individual and the household (equivalised) income distribution, but are more concentrated in a few payment types. For the former, the most common payments and those that were received by at least 10% of the group are Carer Allowance, Family Tax Benefit Part B and Part A, and the Seniors Health Card.

Figure 3 **Percentage of population receiving a social security payment in September 2011, by 2010–11 equivalised household taxable income**



Note: Since the bottom 17 percentiles all reported zero income in 2011, they are combined.

Figure 4 **Percentage of high-income population receiving a social security payment in September 2011, by 2010–11 equivalised household taxable income**



8 Concluding comments

For all the focus on people at the top 1% of the income distribution in popular discussion, we actually know very little about them. We know that they have very high income (obviously), that they tend to have high wealth, and that their share of both is growing in most English-speaking countries. One reason why we know very little about them is that the tax data we have for analysis do not include many background characteristics. The survey data that we have used to analyse income distributions more generally have very few people at the very top of the income distribution.

In this paper, we have used a new data source that overcomes both of these limitations. The MADIP BLE2011 not only has tax data for a very large proportion of the adult population, it also has census data linked at the individual level. Importantly, it has a household identifier, which allows us to calculate the distribution of equivalised household income, as well as the distribution of individual income. The BLE2011 is not perfect. A large proportion of the dataset was not linked to a census record, and there is some top-coding at the very top of the income distribution. These limitations aside, it is far and away the richest dataset available in Australia to look at the characteristics of those at the top of the income distribution.

We have only scratched the surface of what can be done with this dataset. However, we have been able to show in this paper that:

- there is quite substantial movement in and out of the top of the income distribution depending on whether we use individual or household data
- those in the top 1% of the distribution are quite different to those in the next income percentile, who are different still from those in the rest of the top income quartile (particularly in terms of their share of income from wages and salaries)
- a much higher proportion of the top of the equivalised household taxable income distribution are professionals as opposed to managers
- financial and insurance services, as well as professional, scientific and technical services, are the industries where workers are most likely to be in the top 1% of the income distribution
- receipt of social security is quite low at the very top of the income distribution
- a non-negligible number of people in the top 2% of the income distribution still receive some form of payment or allowance, with the most common payments being Carer Allowance, the Seniors Health Card, the Age Pension, and Family Tax Benefit Part B and Part A (in that order).

A final strength of the BLE2011 that will be used in future work is the longitudinal nature of the dataset. Cross-sectionally though, the BLE2011 provides a very useful resource to understand all points on the income distribution.

Notes

1. Although we do not use the dataset in this paper, a comparison dataset (BLE2016) uses a cohort taken backwards from 2016. In our view, the BLE2011 is useful to understand the predictors of income change, whereas the BLE2016 is useful for understanding the census-based characteristics predicted by income and income change. Taken together, they can give a more complete picture of income and income dynamics than we currently have.

2. www.abs.gov.au/websitedbs/D3310114.nsf/home/Australian+Standard+Geographical+Classification+(ASGC)

3. According to the Australian Bureau of Statistics definition, 'Equivalised household income is total household income adjusted by the application of an equivalence scale to facilitate comparison of income levels between households of differing size and composition, reflecting the requirement of a larger household to have a higher level of income to achieve the same standard of living as a smaller household'. www.abs.gov.au/ausstats/abs@.nsf/0/A390E2529EC00DFECA25720A0076F6C6?opendocument

4. Unfortunately, the census only includes occupation and industry for those who are currently employed, and for the person's current employment. No question asks those who are not currently employed what their usual occupation is (unlike some surveys).

5. www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/1292.0.55.002Main+Features12006?OpenDocument

References

Atkinson AB & Leigh A (2007). The distribution of top incomes in Australia. *Economic Record* 83(262):247–261.

——— & Piketty T (2007). *Top incomes over the twentieth century: a contrast between continental European and English-speaking countries*, Oxford University Press, Oxford.

——— & Piketty T (eds) (2010). *Top incomes: a global perspective*, Oxford University Press, New York.

Australian Bureau of Statistics (2018). *Microdata: Multi-Agency Data Integration Projec*t, cat. no. 1700.0, ABS, Canberrra, www.abs.gov.au/ausstats/abs@.nsf/Lookup/1700.0main+features110Australia.

Biddle N, Breunig R, Markham F & Wokker C (2019). *Introducing the longitudinal MADIP and its role in understanding income dynamics in Australia*, CSRM Methods Paper 1, Centre for Social Research & Methods, Australian National University, Canberra.

Burkhauser RV, Hahn MH & Wilkins R (2015). Measuring top incomes using tax record data: a cautionary tale from Australia. *Journal of Economic Inequality* 13(2):181–205.

——— , Hahn MH & Wilkins R (2018). Transitioning from an historical to a contemporary use of tax record data for measuring top incomes in Australia. *Economic Papers: A journal of applied economics and policy* 37(2):113–145.

Gitlin T (2012). *Occupy nation: the roots, the spirit, and the promise of Occupy Wall Street*, HarperCollins Publishers, New York.

Klein E (2011). Who are the 99 percent? *Washington Post*, 4 October, https://www.washingtonpost.com/blogs/ezra-klein/post/who-are-the-99-percent/2011/08/25/gIQAt87jKL_blog.html?utm_term=.a82390befdfd.

Piketty T (2014). *Capital in the twenty-first century*, Harvard University Press.

——— & Saez E (2003). Income inequality in the United States, 1913–1998. *Quarterly Journal of Economics* 118(1):1–41.

Productivity Commission (2018). *Rising inequality? A stocktake of the evidence*, Commission Research Paper, Productivity Commission, Canberra.

Reeves RV (2018). *Dream hoarders: how the American upper middle class is leaving everyone else in the dust, why that is a problem, and what to do about it*, Brookings Institution Press, Washington, DC.

Saez E (2009). *Striking it richer: the evolution of top incomes in the United States (update with 2007 estimates)*, http://elsa.berkeley.edu/~saez/saez-UStopincomes-2007.pdf.

Stiglitz J (2011). Of the 1%, by the 1%, for the 1%. *Vanity Fair*, May, https://www.vanityfair.com/news/2011/05/top-one-percent-201105.

Sullivan P (2010). Empirical evidence on occupation and industry specific human capital. *Labour Economics* 17(3):567–580.

Wright E (1997). *Class counts: comparative studies in class analysis*, Cambridge University Press, Cambridge.

CENTRE FOR SOCIAL
RESEARCH & METHODS

+61 2 6125 1279
csrm.comms@anu.edu.au

The Australian National University
Canberra ACT 2601 Australia

www.anu.edu.au