TE TAI ŌHANGA
THE TREASURY

Analytical Note

April 2023

Trends in the household income distribution: 2007-2021

Author: Meghan Stephens

Analytical Note 23/01

JEL Classifications: D31, G51, D63, E21, E3

Interactive figures and the underlying data are provided here:

https://treasury-analytics-and-insights.github.io/analytical-note-23-01-trends-in-the-income-distribution/analytical-note-23-01-trends-in-the-income-distribution.html

Introduction

The distribution of household income in New Zealand is continually changing and is expected to significantly transform in the years ahead due to demographic trends and shifts in the labour market. This raises an important question: how can we ensure that our tax and transfer system can support future generations of New Zealanders, especially as the population ages and work patterns evolve in response to technological advancements and environmental challenges?

This note lays the groundwork for future research by presenting comprehensive data on household income distributions over the past 14 years. It goes beyond headline statistics on incomes, inequality, and poverty by using unit record data from Stats NZ to supplement existing outputs and reports, such as Perry (2019, 2022). Through visual representations, this note presents household income distributions for various types of households from 2007 to 2021, shedding light on trends and patterns. Additionally, the note delves into the impact of an aging population by creating a counterfactual distribution based on 2007 age demographics, allowing for further investigation into this important demographic shift. The detailed data and analyses presented in this note serve as a valuable foundation for future work in understanding household incomes in New Zealand.

Key Results

Real household incomes have been steadily increasing between 2007 and 2021. That is, after accounting for inflation, incomes have increased consistently since 2007. Inequality indicators have decreased over the same period if we consider after housing cost incomes but have been level if we consider before housing cost incomes.

The trends for different household types help explain these aggregate trends.

- Superannuitants:
  - Increases in superannuitant's real incomes reflect the indexation of New Zealand Superannuation. In more recent years, there also seems to be an increase in multi-adult, superannuitant households who have income in addition to New Zealand Superannuation. This could be because more people carry on working after they reach 65 or because there is an increase in the number of superannuitants who live with other people.
  - The data suggest that there were proportionally more higher-income superannuitant households in 2021 than in 2007. Many have low housing costs (and higher after-housing cost incomes) so this may have contributed to a reduction in population-level after housing cost inequality over the last 14 years.
- Households with children:
  - Before housing cost incomes increased in the first half of this period but in recent years they have increased even more rapidly.
  - Looking at higher incomes, the results suggest increases in wage rates and/or increases in the total hours worked by couples with children.
  - Increases in lower incomes are consistent regardless of beneficiary status, suggesting that they may be partly driven by increases to Working for Families tax credits.
  - The increases are muted when considering after housing cost incomes. Low-income households with children (who tend to have higher housing costs) have not seen the same income increases as superannuitants (who have lower housing costs).
- Households without children:
  - The real incomes of households without children increased less than other households over this period, but most have higher incomes than other household types.
  - There is also evidence of increases to market incomes at the bottom of the distribution.

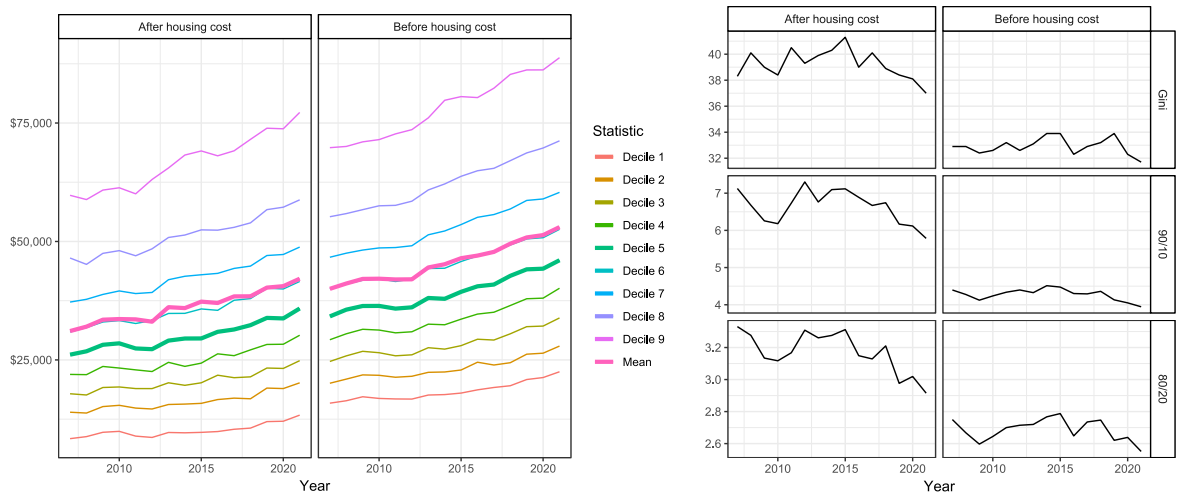
New Zealand's aging population is evident in the data, which show that an increasing proportion of New Zealanders are over 65 years old. However, by creating counterfactual distributions based on the 2007 age distribution, this analysis shows that many of these older households have higher incomes than their 2007 counterparts.

Incomes have been increasing…

Figure 1 shows trends in some commonly used summary statistics for after-housing cost (AHC) and before housing cost (BHC) incomes between 2007 and 2021. The data show that the mean and median real AHC and BHC incomes have been steadily increasing over this period. That is, after accounting for inflation, incomes have been increasing consistently since 2007.

These data also show that high incomes (eg, Decile 9) have increased more than low incomes (eg, Decile 1). This might lead to the conclusion that inequality is increasing, but the commonly used inequality metrics (Gini, 90/10, and 80/20) suggest that inequality has decreased or stayed steady. Roughly, this is because these inequality metrics are focused on proportional changes. For example, Table 1 shows that although the top decile of AHC increased more in absolute terms from 2007 to 2021, it increased at a slower rate than the bottom decile, which led to a decrease in inequality as measured by these three inequality metrics.

Figure 1: Summary statistics of household income distributions between 2007 and 2021



Household incomes are typically compared using household equivalised disposable income (HEDI), which uses equivalisation to allow for comparisons across different household compositions (Stephens M. , 2022). The figures present real incomes, that is, incomes that account for inflation.

Source: Author's calculations using data from Stats NZ.

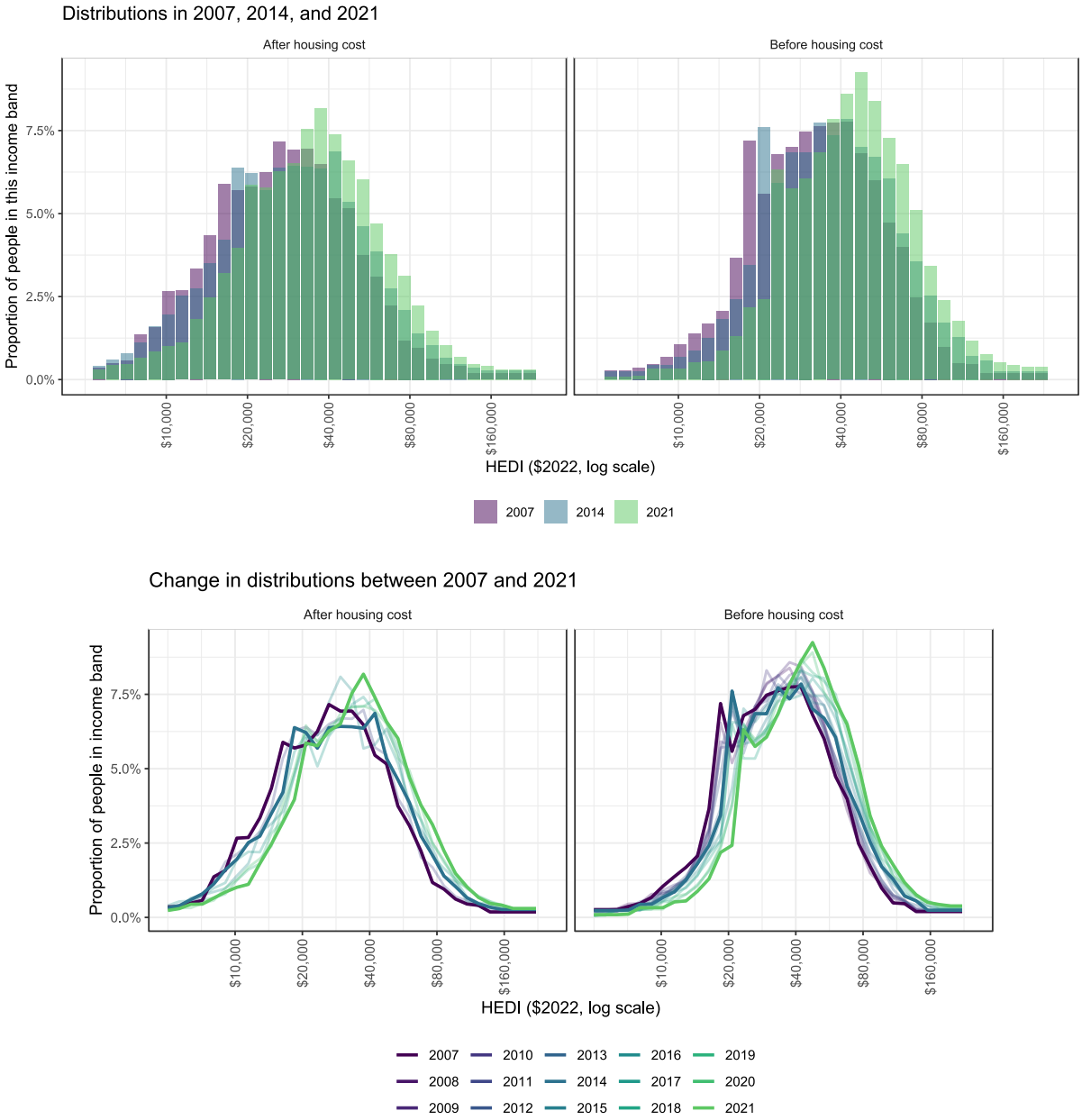
Table 1: Average yearly growth in household disposable income (HEDI) deciles

	After housing cost		Before housing cost	
Decile	**Proportion**	**Dollars**	**Proportion**	**Dollars**
1	3.4%	$355	2.5%	$473
2	2.7%	$443	2.4%	$560
3	2.4%	$499	2.3%	$656
4	2.3%	$590	2.3%	$776
5	2.3%	$696	2.1%	$846
6	2.1%	$761	2.0%	$894
7	2.0%	$829	1.9%	$978
8	1.7%	$877	1.8%	$1,144
9	1.9%	$1,249	1.7%	$1,356

Source: Author's calculations using data from Stats NZ.

Clearly summary statistics are useful, but they remove information. To provide more details, Figure 2 shows how the AHC and BHC income distributions have changed between 2007 and 2021. They demonstrate the increases in income that are evident from the summary statistics, but also show that the shape of the distribution has changed (note that these use a log-scale). For example, in 2021 fewer people have BHC incomes around the $20,000 band and more have incomes around $40,000-$60,000 compared to 2007.

Figure 2: Household income distributions between 2007 and 2021



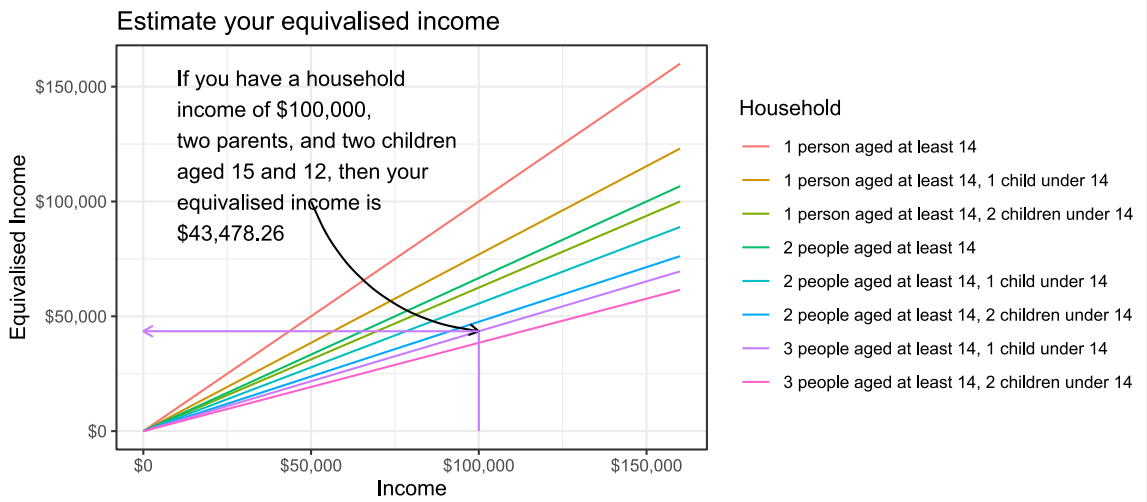
The figures present real (adjusted for inflation) household disposable equivalised incomes. Incomes are presented using a log scale because the distribution has a long tail.

Source: Author's calculations using data from Stats NZ.

Box 1. What is your household equivalised disposable income?

It can be difficult to put household equivalised disposable income (HEDI) values into context. Figure 3 shows some example equivalised incomes (see the Annex for an interactive figure and instructions for you to calculate your own HEDI).

Figure 3: Examples comparing incomes and equivalised incomes



Source: Author's calculations

But not all households have experienced the same income increases

Figure 4 shows the income distributions decomposed into different household types. This decomposition allows us to analyse important aspects of the distribution:

- the relative incomes of people in different types of households, and
- the proportion of New Zealanders who live in different types of households.

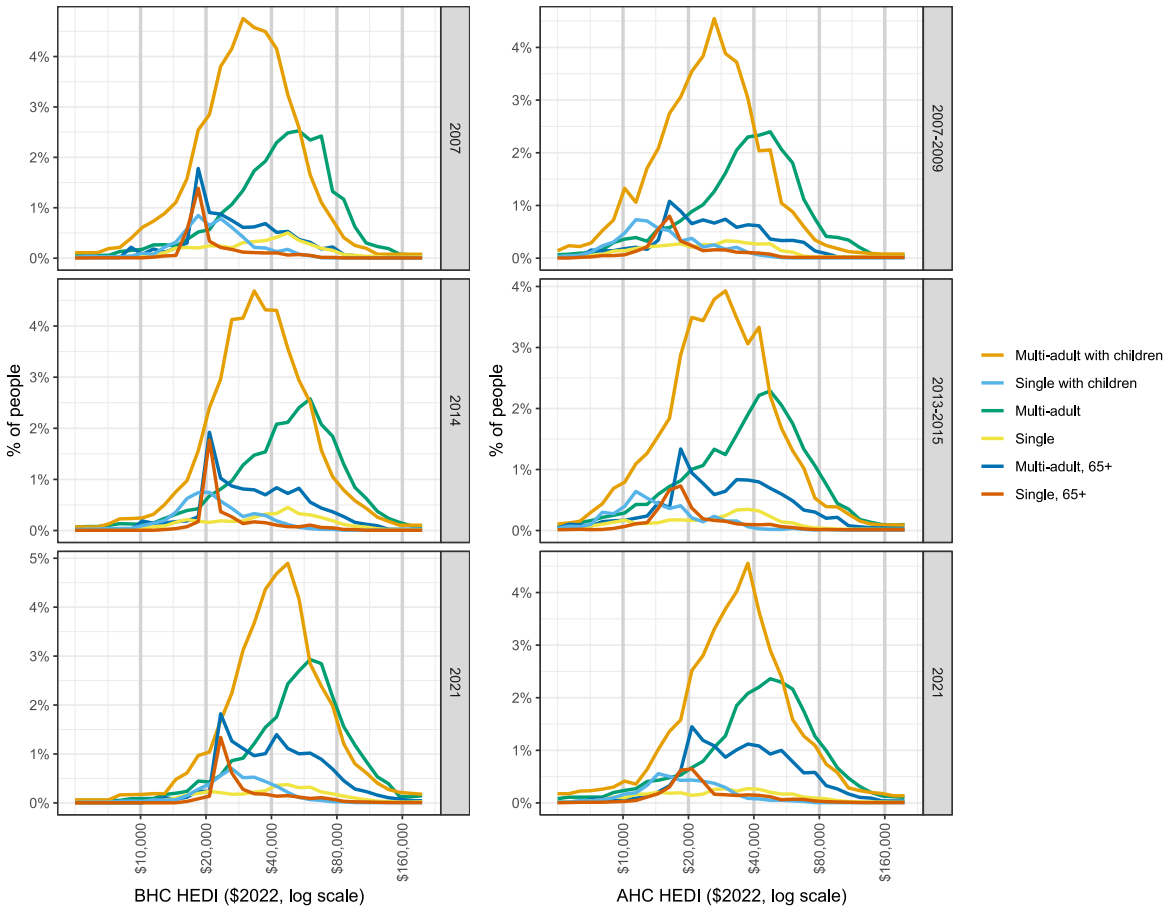
The shapes of these underlying distributions help explain patterns and trends in the total income distribution over time.

Superannuitants

The real incomes of households that contain at least one person aged 65 or older have been increasing steadily (Figure 4), reflecting the indexation of New Zealand Superannuation. In more recent years, there has also been an increase in multi-adult, 65+ households that have income from other sources. This could be because more people carry on working after they reach 65 or because there is an increase in the number of superannuitants who live with other people.

These higher-income, 65+ households are also more likely to have low housing costs, as many will have paid off mortgages on their homes. This could be leading to a decrease in after housing cost inequality.

Figure 4: Decomposition by household type



Source: Author's calculations using data from Stats NZ. Note that AHC data is pooled over three years prior to 2020 because of small samples sizes.

Households with children

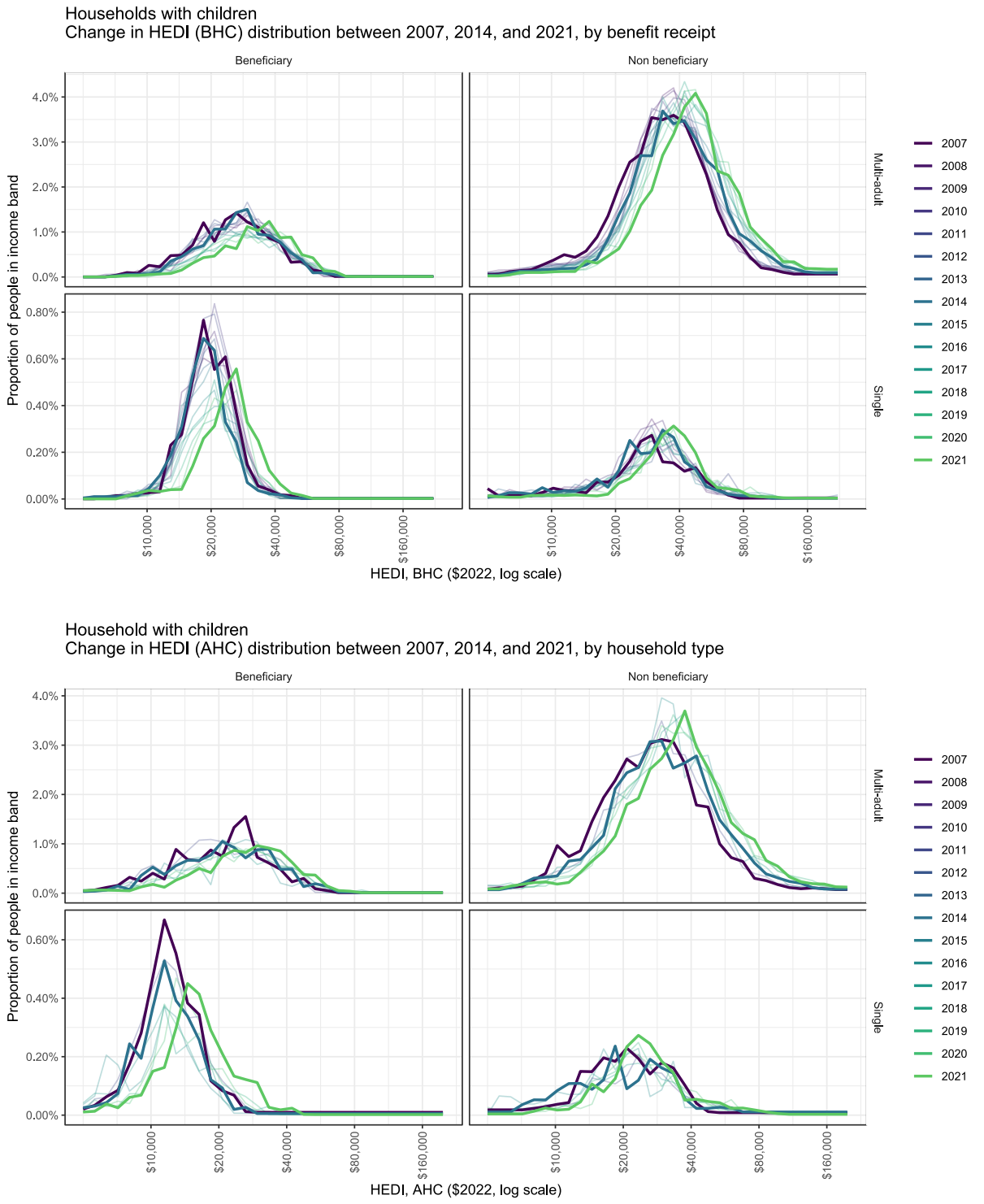
Figure 5 focuses on the incomes of people in households with children[1] in 2007, 2014, and 2021. These figures also consider if anyone in the household received a core benefit (ie, Job Seeker Support, Sole Parent Support, or Supported Living Payment).

The data presented at the top of Figure 5 show that real BHC incomes increased over this period and increased at a faster rate in recent years. The increases in income are consistent regardless of beneficiary status, suggesting that they may be driven in part by changes to Working for Families over this period. However, there are also significant increases at the top of the distribution for non-beneficiary couples, suggesting increases in wage rates and/or increases in the total hours worked by couples.

The bottom of Figure 5 shows the AHC income distributions. The increases in AHC incomes appear smaller than those in BHC incomes. In particular, Figure 4 shows that AHC income increases to low-income households with children do not appear to keep up with increases to superannuitant household AHC incomes.

[1] Here a child is defined as a person under 18 years old.

Figure 5



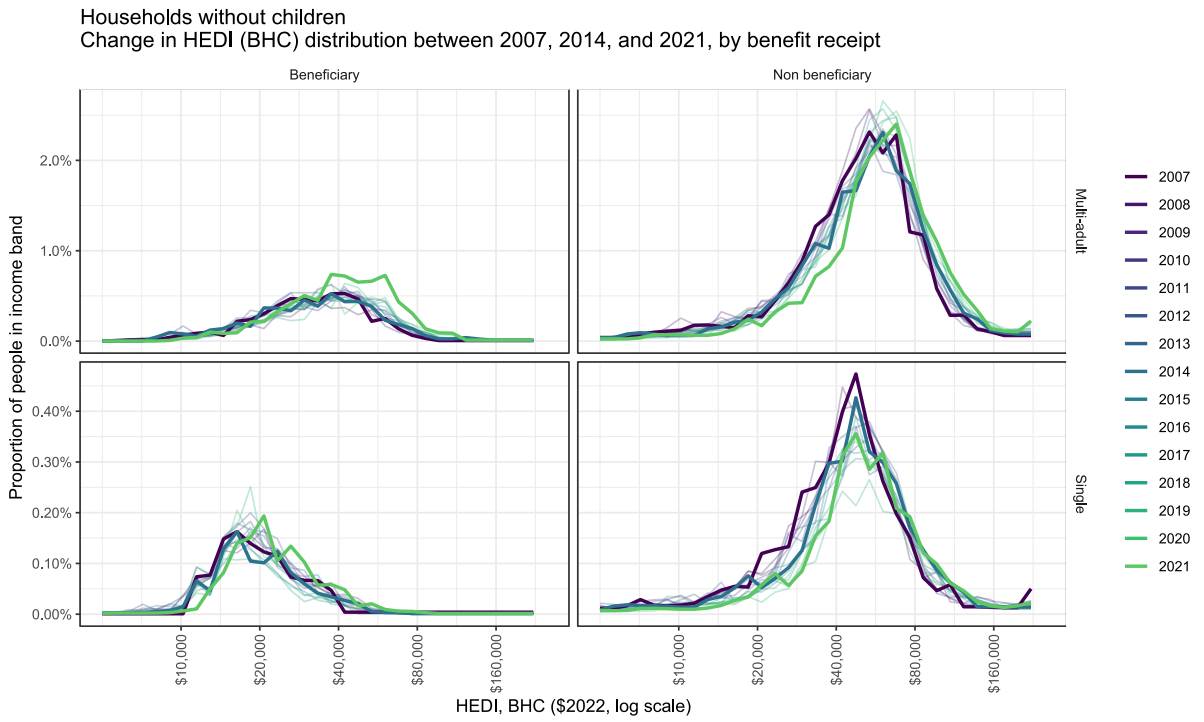
Source: Author's calculations using data from Stats NZ.

Households without children

Figure 6 focuses on the incomes of people in households without children[2] in 2007, 2014, and 2021.

The real incomes of people in households without children have increased over this period but not to the same extent as people in households with children. However, comparison with Figure 4 shows that these people (except for single beneficiaries) typically have higher household incomes than the people in other types of households.

Figure 6



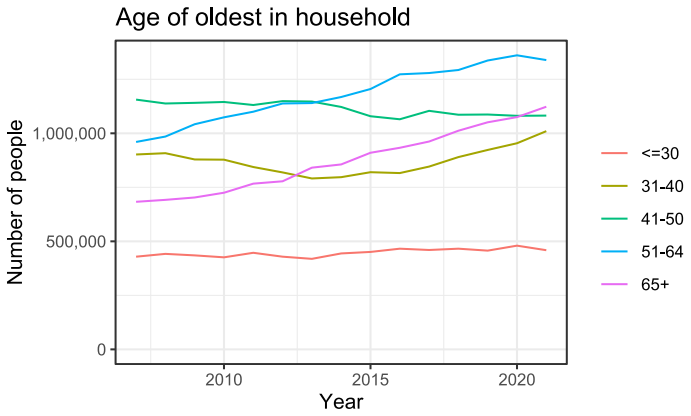
Source: Author's calculations using data from Stats NZ.

[2] These households do not contain anyone under the aged of 18 but may contain adult children living with their parents.

The income distribution also depends on the composition of the New Zealand population

Figure 7 shows trends in the number of people over time split by the age of the oldest person in the house, which reflects the aging population. Given that incomes are highly correlated with age, this change in demographics could have had a significant impact on New Zealand income distribution[3].

Figure 7



Source: Author's calculations using data from Stats NZ.

Figures 8 and 9 present modelled results that use incomes from 2021 but the age composition from 2007. The aim is to answer the question "What would New Zealand's income distribution look like if we did not have an aging population?". Note that the method used to create this counterfactual is similar to that used by Hyslop and Mare (2005) and is a simple re-weighting of the population (see Appendix for more details). As noted by Hyslop and Mare:

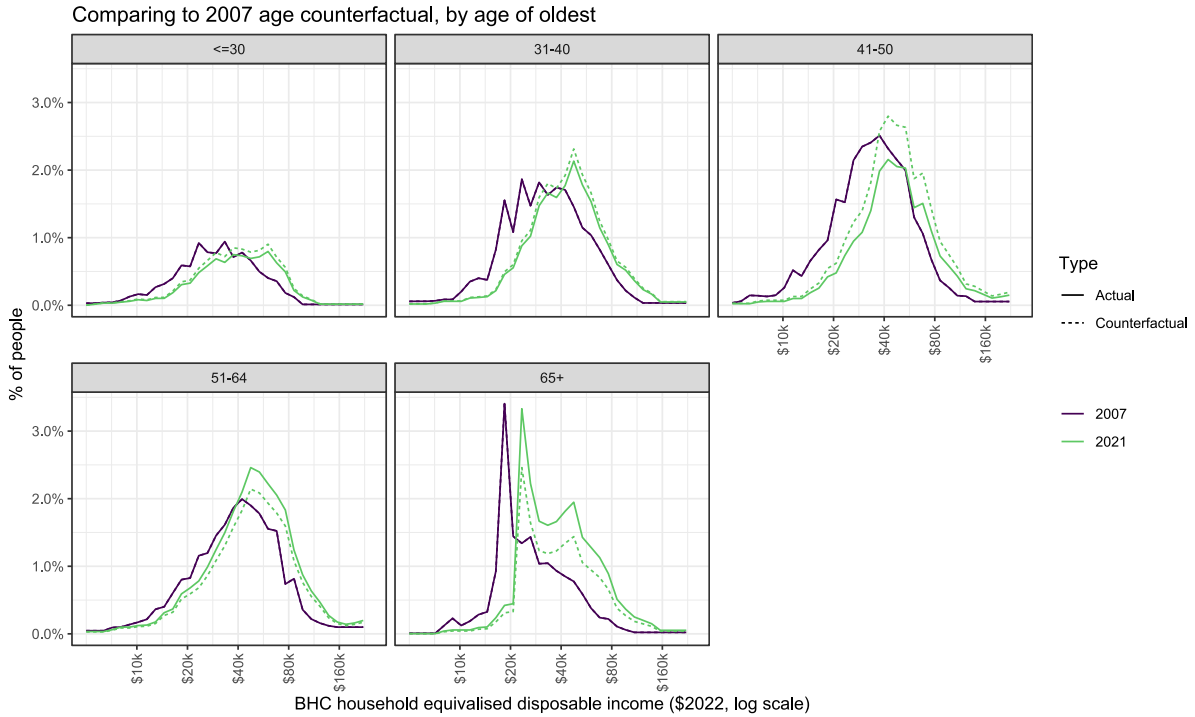
> "… the analysis provides information on the nature of the distributional changes, and the extent to which they were associated with particular sets of observable household attributes."

In Figure 8, the data show that the counterfactuals have more people in the 41-50 age group and less in the 65+ age group. It would be reasonable to assume that the increase in older households that can be seen in the data would lead to an increase in lower income people, as older households have typically had lower incomes. However, the "new" older households are relatively evenly spread across the income distribution (both BHC and AHC).

Aggregating the counterfactual distributions up to the full population (Figure 9), we see that the drop in the number of 41-50 group is mostly counteracted by the 65+ group who have similar incomes. This suggests that there haven't been significant inequality increases driven by age demographics, because the income distributions of older households have also changed.

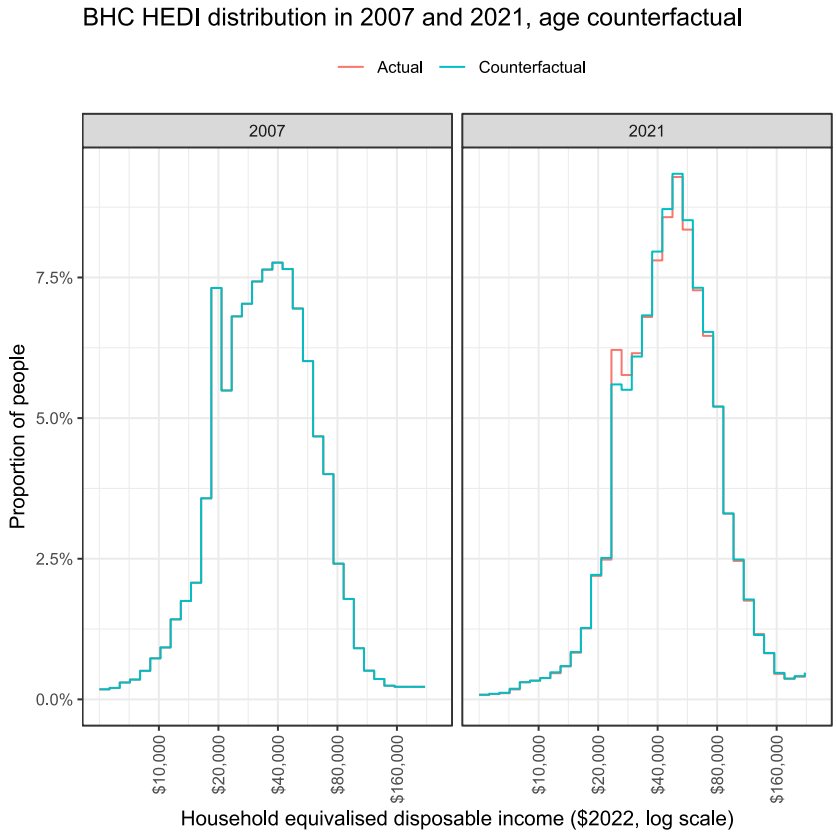
[3] Alimi, Mare and Poot (2018) and Aziz et al. (2015) also investigate the impact of an aging population on incomes.

Figure 8



Source: Author's calculations using data from Stats NZ.

Figure 9



Source: Author's calculations using data from Stats NZ.

Treasury will continue to build our understanding of household incomes

This note sets the scene for future work by providing detailed household income distributions from the last 14 years, providing information about the data underneath headline statistics on incomes, inequality, and poverty. We plan to dig deeper into the underlying linked data to understand how different data sources can impact headline statistics, investigate more counterfactuals, and decompose headline statistics to investigate drivers of change.

Acknowledgements

Many thanks to my colleagues in Treasury's Analytics and Insights team, in particular to Michael Eglinton for helping derive these modelled outputs. I am also grateful to Caroline Brooking and Chris Pooch at Stats NZ for producing and sharing their data.

References and further reading

Alimi, O., Mare, D., Poot, J. (2018). More Pensioners, Less Income Inequality? The Impact of Changing Age Composition on Inequality in Big Cities and Elsewhere, New Zealand Population Review

Aziz, O., Ball, C., Creedy, J. & Eedrah, J. (2015) The distributional impact of population ageing in New Zealand, *New Zealand Economic Papers*

Hyslop, D. & Mare, D. (2005). Understanding New Zealand's Changing IncomeDistribution, 1983–1998: A Semi-parametric Analysis. *Econometrica*, 469-495.

OECD (n.d.). *What are equivalence scales?* Retrieved from https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf

Perry, B. (2019). *Household incomes in New Zealand: Trends in indicators of inequality and hardship 1982 to 2018.* Wellington: Ministry of Social Development.

Perry, B. (2022). *Child Poverty in New Zealand.* Wellington: Ministry of New Zealand.

Stats NZ (2020). *Microdata output guide – 5th edition.* Stats NZ.

Stats NZ (2022). *Child poverty statistics: Year ended June 2021 – technical appendix*. Retrieved from https://www.stats.govt.nz/methods/child-poverty-statistics-year-ended-june-2021-technical-appendix

Stephens, M. (2022). *Insights from New Zealand child poverty data.* New Zealand Treasury.

Annex: Method, Data, and IDI Disclaimer

Detailed results

Interactive figures and the underlying data are provided here:

https://treasury-analytics-and-insights.github.io/analytical-note-23-01-trends-in-the-income-distribution/analytical-note-23-01-trends-in-the-income-distribution.html

Data

The unit record data used in this analysis were provided by Stats NZ and cover the period from 2007 to 2021. This is the same dataset that is used by Stats NZ to create child poverty and household income statistics.

For 2019–2021, the data are derived from the Household Economic Survey (HES) linked to administrative income data in the Integrated Data Infrastructure (IDI). Stats NZ increased the size of HES to around 20,000 households in 2019 as part of their work to improve Child Poverty measurement (Stats NZ, 2022).

For years prior to 2019, this analysis uses data that was derived by Stats NZ to produce a baseline for the Child Poverty reporting regime. In these years, the HES was a much smaller survey of around 3,500–5,000 households. For before housing cost incomes, the data comprise HES combined with the Household Labour Force Survey (HLFS) linked to administrative data in the IDI. After housing cost incomes are only derived from HES combined with administrative data because HLFS does not contain information on housing costs.

This means that there is a series break in 2019, due to different data sources and sample sizes. However, the data show consistent trends in the distributions and statistics of interest to this paper over this time period.

How to calculate your own Household Equivalised Disposable Income (HEDI)

To calculate your own HEDI:

- First calculate the disposable income for your household over the previous year. For each household member, add up all their income from the previous year, including income from private sources (eg, wages, self-employment, interest), core benefit income (eg, Job Seeker Support, Supported Living Payment), supplementary benefits such as Accommodation Supplement, tax credits such as Working for Families, and any other income. Then subtract any personal income tax and ACC levy paid. Your (BHC) household disposable income is the total for all people in your household. That is,

$$Household\ disposable\ income := all\ income\ minus\ all\ personal\ income\ taxes$$

- Your AHC disposable income is your BHC disposable income minus your total housing costs (including rent, mortgage interest and principal, and other housing costs such as rates). That is,

$$AHC\ disposable\ income := Household\ disposable\ income - all\ housing\ costs$$

- Next, calculate an equivalisation factor. The modified OECD scale assigns a weight of 1.0 to the first adult in the household, 0.5 to each additional adult (anyone 14 years or older) and 0.3 to each child (OECD, n.d.; Stats NZ, 2019). That is,

$$Eq\ factor := 1 + 0.5(1 - \#aged\ at\ least\ 14) + 0.3(\#aged\ less\ than\ 14)$$

  So, a household containing a couple aged 50 and 55, and three children aged 16, 10, and 3 has an equivalisation factor of $1 + 0.5(2) + 0.3(2) = 2.6$.

- Finally, your HEDI is your disposable household income (before or after housing costs) divided by your equivalisation factor.

$$HEDI := \frac{Household\ disposable\ income}{Eq\ factor}$$

Estimating the full distribution using data that can be released from the IDI

Stats NZ have strict rules for releasing IDI data to ensure that no one in the data can be identified (Stats NZ, 2020). Researchers analyse unit record data within the IDI, create summaries for output (eg, histograms), and submit these outputs for release. The Stats NZ checking team thoroughly investigate the outputs to check that the appropriate confidentiality rules have been applied. The IDI has been a great success, but the volume of requests has put some strain on the checking process. Everyone in the IDI community benefits if we minimise the number of outputs sent for checking. With this in mind, this analysis uses a method to estimate multiple summaries and measures of the distribution using only one IDI output, that is, quantiles.

Details

We can estimate full distributions using quantiles and linear interpolation. Consider the cumulative distribution function, $F(x)$, where

$$F_X(x) = P(X \leq x)$$

That is, $F(x)$ is the probability that random variable X is less than or equal to x. Quantiles are cut points that divide a distribution into equal intervals. Deciles are an example of a quantile. The first decile is a value such that 10% of the population are below this value and 90% are above. The quantiles, q, are more formally defined as

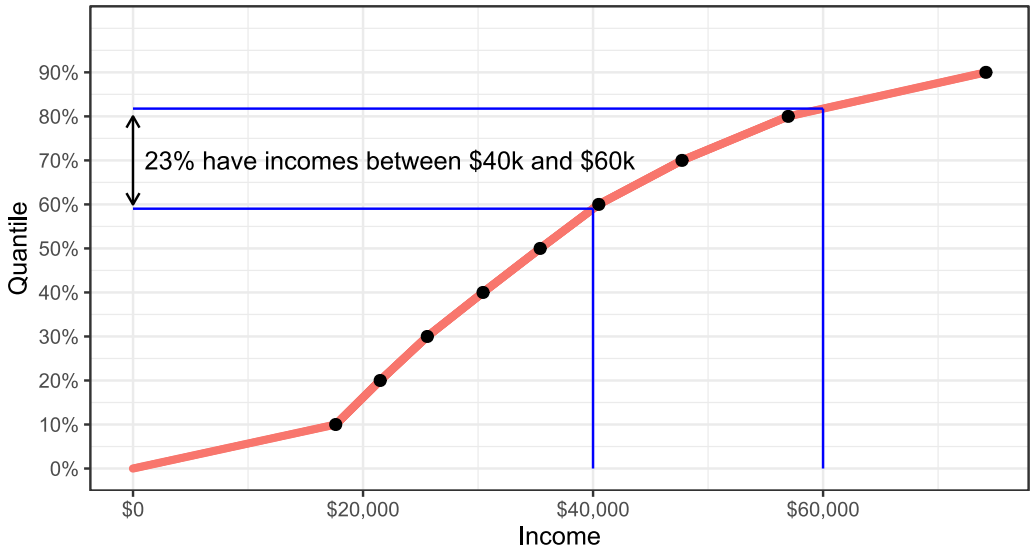
$$F^{-1}(q) = \inf\{x : F(x) > q\}$$

where inf is the infimum or greatest lower bound. To analyse the data, we pick the quantiles we are interested in, q_i, and calculate the values v_i such that $v_i = F^{-1}(q_i)$. For example, if we are interested in percentiles, $q_i = \{1\%, 2\%, 3\%, \ldots, 99\%\}$. The quantiles can be released from the IDI[4], and then we can interpolate the $\{q_i, v_i\}$ pairs outside the IDI to estimate other statistics of interest.

[4] If confidentiality requirements are met, see Stats NZ (2020).

As an illustrative example, Figure 10 shows the deciles of some hypothetical income data and the piecewise linear interpolant, which can be used to estimate different points in the distribution. To construct the histograms in this analysis we chose arbitrary incomes bands on the x-axis (Figure 10 considers the income bin $40,000 and $60,000), calculate the corresponding quantile (on the y-axis), and calculate the proportion of people in each band by taking the difference of the quantiles. Other statistics of interest can also be estimated using the interpolated quantiles. We cannot release the minimum or maximum income from the IDI, so we can either estimate the min/max or only estimate the distribution between the bottom/top quantile. We can assume the minimum income is $0 (although some people do have negative incomes), but the maximum is more difficult.

Figure 9



Illustrative example demonstrating how quantiles can be used to estimate arbitrary histograms. The black points are the deciles of a hypothetical income distribution and the red line is the piecewise linear interpolant.

Estimating a counterfactual based on a different age distribution

This analysis also constructed a counterfactual distribution based on the 2007 age distribution. The counterfactual used the incomes from each year but uniformly reweighted the population in each income band so that the proportion of the total population in each "age of oldest in household" group was the same as in the 2007 data. Let $N_{g,y}$ be the total number of people in group *g* in year *y*. To derive the counterfactual, we multiply the number of people in each income bin of the histogram for year y, household type *g* by:

$$Scaling\ factor := \frac{N_{g,2007}}{N_{g,y}}$$

This is similar to the method in Hyslop & Mare (2005) and limits the counterfactual to only change the total number in each of the "age of oldest" groups. The same re-weighting method can be used for other groups.

IDI Disclaimer

These results are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI) which is carefully managed by Stats NZ. For more information about the IDI please visit https://www.stats.govt.nz/integrated-data/. The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes and is not related to the data's ability to support Inland Revenue's core operational requirements.