

Distributional National Accounts in New Zealand

By Zahra Soleimaninajafabadi

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Masters of Commerce.

Victoria University of Wellington

2021

Disclaimer: The results in this thesis are not official statistics. They have been created for research purposes from the Integrated Data Infrastructure (IDI), managed by Statistics New Zealand. The opinions, findings, recommendations, and conclusions expressed in this study are those of the author(s), not Statistics NZ. Access to the anonymised data used in this study was provided by Statistics NZ under the security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business, or organisation, and the results in this [report, paper] have been confidentialised to protect these groups from identification and to keep their data safe. Careful consideration has been given to the privacy, security, and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz/NZ (2020).

The results are based in part on tax data supplied by Inland Revenue to Statistics NZ under the Tax Administration Act 1994. This tax data must be used only for statistical purposes, and no individual information may be published or disclosed in any other form, or provided to Inland Revenue for administrative or regulatory purposes. Any person who has had access to the unit record data has certified that they have been shown, have read, and have understood section 81 of the Tax Administration Act 1994, which relates to secrecy. 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.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Motivation | 1 |
| 1.2 | Problem Statement | 1 |
| 1.3 | Summary of Contributions | 3 |
| 1.4 | Thesis Outline | 3 |
| | List of Figures | 1 |
| | List of Tables | 1 |
| 2 | Literature Survey | 5 |
| 2.1 | Introduction | 5 |
| 2.2 | A General Literature on DINA | 5 |
| 2.2.1 | Inequality in New Zealand | 7 |
| 2.2.2 | DINA in New Zealand | 10 |
| 3 | Data and Methodology | 13 |
| 3.1 | Data | 13 |
| 3.1.1 | Micro Data | 13 |
| 3.1.2 | Macro data | 15 |
| 3.1.3 | Data Issues | 16 |
| 3.2 | Methodology (How to Construct DINA) | 18 |
| 3.2.1 | Overview of the approach | 18 |
| 3.2.2 | The Income Concepts | 19 |
| 3.2.3 | Factor Income | 20 |
| 3.2.4 | National Income | 20 |
| 3.2.5 | DINA | 20 |
| 4 | Results and Discussions | 27 |
| 4.1 | Introduction | 27 |
| 4.2 | DINA constructed by HES database | 27 |
| 4.3 | DINA constructed by IR database | 28 |
| 4.3.1 | Distributional Factor Income: | 28 |
| 4.3.2 | Distributional National Income | 30 |

| | | |
|----------|---|-----------|
| 5 | Conclusions and Future work | 33 |
| 5.1 | Conclusion | 33 |
| 5.2 | Future directions | 33 |
| | Appendices | 35 |
| A | Classification of income sources based on HES database | 37 |
| B | R code | 41 |

1 Introduction

1.1 Motivation

Even a quick consideration of income inequality studies reveals the fact that the adverse effects of this phenomenon are too serious to be ignored. For instance, *Hsieh and Pugh* (1993) show that more unequal societies face the higher rates of homicide and crime and *Fajnzylber et al.* (2002) provide evidence that inequality affects health, longevity and quality of social relations in the society. Therefore, policy makers should have special attention to this matter as one of their most important obligations. But, the primary concern is having reliable measurements of income inequality. The best decisions or policies for any economic problem are made when a comprehensive and precise vision as well as accurate indicators to measure or compare that matter are available. For this aim either accessing to reliable sources of data or corroborated method for processing data are critical. Also, comparability of indicators is another point that makes choosing proper data set and methodology more substantial.

1.2 Problem Statement

Considering the importance of measuring inequality, establishing a time series of different group income shares to reflect the distribution of income inequality and its changes over time has been attracting the attention of many researchers, economists and policy makers across the world, specifically in the Organisation for Economic Cooperation and Development (OECD) countries. Due to this fact, there has been several efforts for measuring income inequality, finding the share of different income groups from economic growth, and monitoring the change of income inequality specially across the top

income percentiles. Particularly, it became even more attractive to focus on during and after the period 1984–96, that New Zealand experienced a wide range of economic and social policy reform (*Evans et al.*, 1996). To obtain this aim, researchers mostly have used tax income data and information from surveys such as Household Economic Survey to determine Gini coefficient or top income shares. For example, *Podder and Chatterjee* (2002) examines the trends of household income inequality in New Zealand using Household Economic Survey (HES) data in unit record form which indicates a steady upward trend in income inequality in New Zealand. Although these endeavours shed to a light over this matter in New Zealand, some limitations with their data bases and methodologies has left this picture still vague in some aspects. Relying merely on micro-data (tax and survey data) causes a gap between National Accounts and these studies, that leads the result to be inconsistent with economic total aggregations and growth. Therefore, there is no satisfactory answer for estimating the share of different income groups (bottom 50 percentiles, middle class, top 10 percentiles and top 1 percentile) from economic growth. Moreover, we are interested in knowing if changes in income inequality is because of changing in capital-labour ratio in national income or changes in distribution of income sources (wage and salary or capital income) and returns to capital. 'How does the participation of female as labour force shape the trends in income distribution' and 'how are the distributions of income trends reshaped by interfering the government with using tax and transfers tools' are both noticeable questions that there are no clear responses for them with mentioned previous studies. New studies came by new methods to overcome all these blind points.

Distributional National Accounts (DINA) is an idea suggested by *Alvaredo et al.* (2016) that provides decomposition of changes in income by different income groups using a scale factor that is constructed by both micro-data and national accounts sides, also by gender factor to cover the inconsistency problem as well as female contribution problem. The World Inequality Database (WID.world) *Database* (2020) aims to provide open and convenient access to the most extensive available database of the world distribution of income and wealth, both within countries and between countries, focuses on this methodology to make comparable and consistent series for the countries across

the world.

1.3 Summary of Contributions

The contribution of the present study is estimating Distributional National Accounts for New Zealand to investigate income inequality changes in an specific period of time (2000 to 2018). For this purpose, Inland Revenue tax data are combined with National Income data. A second version is created using the Household Economic Survey data in place of the Inland Revenue tax data. Our series display the evolution of distribution of National Accounts (either Factor Income or National Income) using two separate data bases (IR data and HES data) so our results show how results are different using different database as well. We find that the gap between high and low incomes has changed neither over period 2006-2015, nor in a wider period of time (2000-2018). The income share of the bottom 50 percentiles from National Income has been between 9 and 13 percent, share of middle class (50 to 90th percentile) has been between 34 and 39 percent, and share of top 10 percent has been between 50 and 55 percent. The share of the top 1 percent from National Income has increased drastically in years 2005, 2011, and 2016 while share of percentiles 90th-99th has dropped in those years.

1.4 Thesis Outline

The rest of this paper is organised as follows. Section 2 describes some previous studies which focus on measuring inequality with different approaches for either New Zealand or other countries. In section 3, data sources and methodology are described. Section 4 includes results and main findings. Section 5 presents concluding comments and research perspectives.

2 Literature Survey

2.1 Introduction

Growing interest in either measuring income and wealth inequality or developing efficient methodologies, for this measurement specially for developed countries, gave rise to a flourishing literature in this regard in recent years. In particular, studies for estimating the share of income for top income groups has increased significantly in last two decades. We review some efforts which have been made around the world using same approach to measure inequality. They all used the same method we used to estimate distributional national accounts. Results of all the efforts below are available on WID.world database for their methodology, data and other aspects of their measurements are consistent with this database's approach.

2.2 A General Literature on DINA

Piketty et al. (2018) combined tax, survey, and national accounts data to measure the distribution of national income in the United States since 1913. Their work comprises estimating not only the distribution of pretax national income, but also distribution of national income after taking into account the government intervention. Their calculations display that average pretax real national income per adult increased 60 percent from 1980 to 2014, but they show that it decreased for the bottom 50 percent of the distribution around 16,000 dollar a year. The pretax income of the middle class has grown 40 percent since 1980. In this attempt, they also estimate the share of either capital or labour income for different income groups. According to their findings, upsurge of top incomes was first a labour income phenomenon but it has changed to capital

income phenomenon since 2000. This study reveals that the government's role is not substantial in decreasing the income inequality in mentioned years. Their study has another sector which make it even more valuable. They broke down DINA regarding to the gender of individuals and their results cover the share of female income from total income in each percentile. According to that share of females has increased in last 20 years that has caused reduction in inequality among adults.

In France, in DINA fields there are two separate studies for pretax distributional national accounts and post tax distributional national accounts. *Garbinti et al.* (2018) focused on DINA before government intervention with capital vs. labour income, age and gender breakdown. As other cooperation in WID, they combined national accounts, tax and survey data in a consistent way to build homogeneous annual series on the distribution of national income by percentiles for years 1900 to 2014. They present one substantial result that displays taking advantage of DINA methodology. They compare their series with previous tax-base findings in long-run inequality series in France. Their series show higher inequality levels for the recent decades. They indicate that although a sharp drop in the concentration of wealth and capital income seemed to cause a decline in inequality, a rising part of capital income increases it considerably. Their detailed breakdowns by age and gender show that gender inequality in labour income declined in recent decades, although it is not substantial among top labour incomes. They also take advantage of DINA results to compare inequality between the U.S and France. They point out that average pretax income among bottom 50 percent adults is 20 percent larger in France than in the U.S while national income per capita is 30 percent smaller in France. *Bozio et al.* (2018) worked on building DINA series in this effort considering government role in France. They discuss different ways of measuring tax progressivity then use their new series to understand the impact of taxes and transfers on redistribution of income. Their findings display the fact that taxes and transfers reduce total income inequality by 23 percent on average over the 1990-2018 period. For this, both upper-end and lower-end redistribution are applied. They explain that increasing redistribution in France tax system in considered period of time, which is from reductions in non-contributive social security contributions for

the bottom 50 percent of individuals and tax increases for the top 10 percent, has caused a relatively constant level of disposable income inequality.

WID.world provides DINA series for all countries which have built them and updates them constantly.

2.2.1 Inequality in New Zealand

In this sector we summarise studies on inequality in New Zealand. This literature largely focuses on estimating share of income for top income percentiles and most of them using Gini coefficient as an index to talk about inequality. Some of the studies carry out regression approach in this regard, using the income shares of a specified percentile of the distribution as dependent variable or several variables have been used as the dependent variable.

Podder and Chatterjee (2002) examined the trends of household income inequality in New Zealand over the period 1984–96. They investigate the manner in which the national income is divided up amongst different groups in society after implementing a wide range of economic and social policy reform in New Zealand. They also measure the contributions of the different sources of personal income to the overall inequality decomposing income inequality by income components. In this effort unit record data from the Household Expenditure and Income Surveys and the Household Economic Survey (HES) are used. Their results clearly show that inequality of household incomes in New Zealand has been on the rise over the mentioned period as indicated by the rising value of the Gini coefficient over the period (over 14 percent increase in 12 years from 1984 to 1996). They note there is a decline in the shares of the bottom eight deciles of the households, the share of the ninth decile has remained steady, the share of the top decile has increased significantly. To capture the changes in the relative income shares, they observe how the ratios of the shares of the top and the bottom percentiles for example have altered over time. The top 5 percent of income earners received over seven times as much as the bottom 10 percent in 1983/84. For the next part of the study they consider seven components of total income. Wages and Salaries, income from self-employment, income from investment, personal superannuation, na-

tional superannuation, government cash transfer, and other incomes. Then note that income from wages and salary also has a concentration coefficient higher than the Gini coefficient of total income which implies that this income is more unevenly distributed (in favour of the higher income groups) than total income. This uneven distribution affects the overall inequality strongly.

Dixon(1996) used Household Economic Survey (HES) database to measure the distribution of individual earnings and investigate long-run changes in the earnings structure between 1984 and 1995. This study's results include aggregate earnings inequality, the gender earnings gap and shifts in relative earnings by level of educational attainment. This study take advantage of regressing variables. To explore the changes in the independent influence of each measured attribute on earnings, a simple linear earnings equation is estimated for each annual dataset. They indicate that the overall increase in inequality in mentioned decade was very small, and caused by a rise in the relative earnings of workers at the top of the distribution, rather than a decline in the relative earnings of low wage workers. In addition, there has been a substantial reduction in gender differentials over the decade, reflecting an upward shift of all levels of the female distribution and some down-ward movement of the male distribution. They also found an increase in earnings inequality among males with post-school qualifications.

Winkelmann and Winkelmann (1998) examined inequality from a labour market perspective in an econometric framework. They examine the labour market outcomes, in terms of employment and incomes, for immigrants in New Zealand with the help of data primarily from the 5- or 10-yearly population censuses between 1981 and 1996. This study looks at the income differentials within immigrant groups, and between specified categories of immigrants and native-born New Zealanders. They conduct a cohort analysis of immigrants' relative incomes and obtained results from regressions using all employed individuals aged 15-64 for whom income data are available. Their analysis answers the question of how much of the difference in incomes between immigrants and natives remains after we control for hours of work, gender, and productive characteristics. They display evidence for a substantial income disadvantage of arriving immigrants relative to natives after we account for differences in qualification levels

and other personal characteristics.

Bakker and Creedy (1999) analysed the effects of macroeconomic variables on the personal distribution of income over time. They use a conditional lognormal-exponential mixture maximum likelihood estimates to model the complete distribution of income in each year. This method has been used to make New Zealand income distribution for wage and salary earners over the period 1985 to 1994. They came to the conclusion that either increasing the unemployment rate and reductions in the rate of GDP growth have driven the increase in inequality in that period.

Easton et al. (2013) and *Easton et al.* (2014) estimated the Gini index for the entire adult population from 1926 to 2013 as well as top income shares using census data. By including individuals with zero or very low, non-wage income (i.e. not only taxpayers) the Gini coefficients increase in this period. Easton et al, found that income share of the top decile fall between the 2001 and 2006 censuses, reversing in the 2013 census to be broadly unchanged over the period.

Ball and Creedy (2016) used HES data and survey calibration method, first, for analysing annual income and expenditure inequality in New Zealand over 1983-2013, and second, for comparing the inequality of market incomes with that of disposable incomes to investigate the extent of redistribution through the tax and benefit system. They aimed to describe components of inequality as well and to reach this purpose a decomposition method is used involving five sets of variables (age/gender structure, labour force participation, household type, housing tenure type and occupancy rate) along with the sample itself. They found an increase in the inequality of market and disposable income per adult equivalent person from the late 1980s to the early 1990s. They also point out that inequality changes are influenced by a range of factors associated with the structure of the population, which are expected to change over the relevant period. In addition, there has been an increase in female labour force participation as well as the increase in participation among older males over the period. Therefore, with a constant demographic and labour force structure, the inequality of expenditure displayed a ‘flatter’ profile over the period.

Creedy et al. (2017) estimated the Gini index between 1935 and 2014, using tabulated data on personal taxable incomes from Inland Revenue for estimates between 1981 and 2014. They find that the Gini index is relatively constant. In addition, the authors investigate differences in the Gini index of males and females. They find that while overall income inequality is unchanged, income inequality among females has declined since 2000 and is lower than that of males over the entire period.

2.2.2 DINA in New Zealand

NZ (2018) followed the methodology recommended by the OECD Expert Group on Disparities in a National Accounts framework (EGDNA) to distribute the national accounts values across different household groups. To reach this purpose they use Household Economic Survey (HES) data as well as the National Accounts (more specifically the Household Sector Accounts). The results plot the ratio to the average of disposable income and adjusted disposable income by quintile for 2016. They show that the disposable income of households classified to the highest quintile is 2 times the average, while the lowest quintile is 30 percent of the average. The similar ratios for adjusted disposable income are 1.8 times and 50 percent respectively, illustrating how the addition of social transfers in kind reduces the income disparities. It also points out that the consumption disparity measure is much lower than the income measure. While the disposable income of households in the highest quintile is 2 times the average, the level of final consumption expenditure for the same group of households is only 1.4 times the average. For the lowest quintile, similar figures are 30 percent and 70 percent respectively. Using a consumption ratio significantly lowers the measured level of inequality.

There are some valuable attempts to investigate the evolution of the income distribution focusing on top income shares applying Atkinson-Leigh method¹. These studies measure top income shares using tabulated data on taxable income published on Inland Revenue's website(www.ird.govt.nz) *Revenue* (2020). These type of estimates have been

¹(For more details about this method see Appendix 8A of Atkinson and Leigh (2007). *Atkinson and Leigh* (2007)

repeatedly updated in the World Inequality Database by Atkinson and Leigh (2007, 2008) *Atkinson and Leigh* (2008), Atkinson (2012) *Zealand et al.* (2013), Alvaredo and Atkinson (2013) *Alvaredo et al.* (2013), Alvaredo and Atkinson (2014) *Alvaredo et al.* (2014), Alvaredo (2017) *Alvaredo et al.* (2017). In general speaking, they determine that top income shares in NZ did not greatly change between 2000 and 2016.

The closest effort to our study in terms of methodology to measure distributional national accounts in New Zealand is *Kergozou* (2017). Apart from using Atkinson-Leigh method to measure the evolution of top income shares, she also combines tabulated income tax and national accounts data to build simplified distributional national accounts for New Zealand between 2009 and 2016 and basic simplified distributional national accounts between 2000 and 2008. DINA series which she has built display no significant increase in inequality in NZ over 2000-2008 as well as 2009-2016. The results indicate that both the top and the bottom income groups benefited from economic growth in more recent years (2011-2016).

3 Data and Methodology

In this section we describe the concepts, data sources and main steps of the methodology that we use to construct income distribution series. We use two main types of data, national accounts and fiscal data (income tax returns). A third type (survey data) is used to construct an alternative income distribution series. We describe our data sources for the period we can use micro-files of income tax returns in New Zealand, explain the methodology, and present our income concepts.

3.1 Data

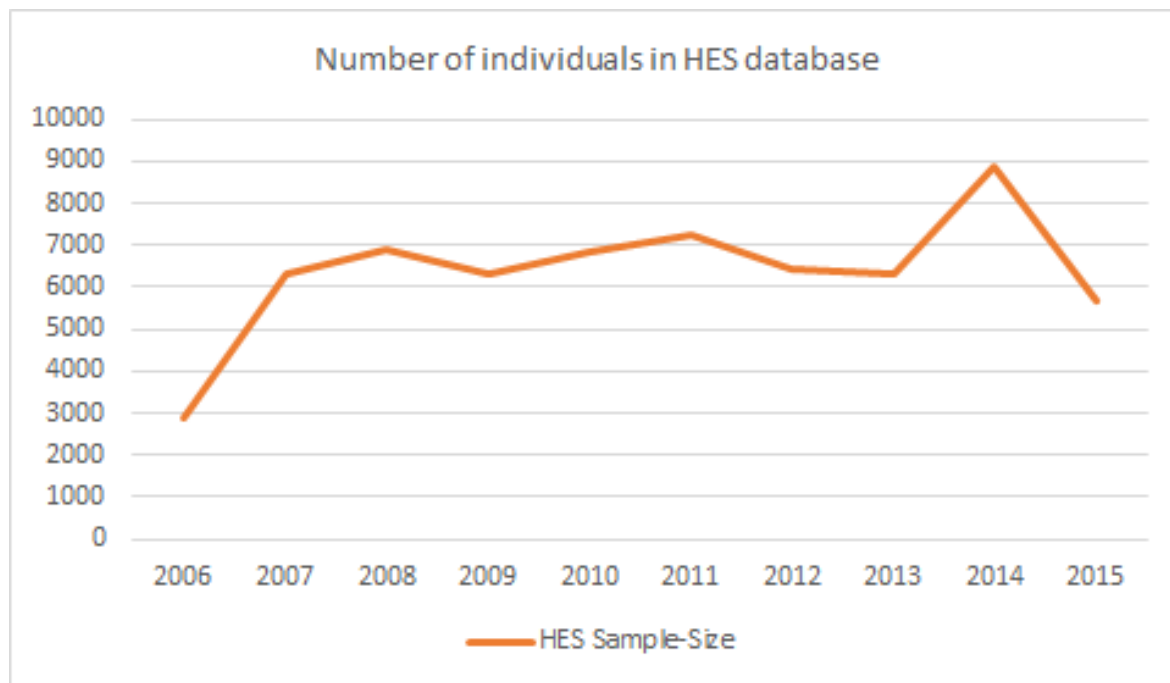
3.1.1 Micro Data

Household Economic Survey(HES) Tables

One of the databases we used for distributing national accounts is one of the survey based databases of New Zealand's Integrated Data Infrastructure (IDI). The Household Economic Survey (HES) collects information on household income, savings, and expenditure, as well as demographic information on individuals and households. The primary objective of HES (Income) is to provide indicators on how personal and household income, housing costs, and living standards changed over time ?. To estimate the income shares in this part, we have used table hes-clean.hes-income from refresh IDI-Clean-20190420. Data from the HES containing source of income is available from years 2006 to 2015. Reporting year for the HES is 30 June, that is different from reporting year of our other databases that is 31 March for taxable incomes and national accounts data. This causes an inconsistency in comparing results from these three sources. A table containing all income classifications from Stat NZ' HES data dictionary has been

attached to Appendix A. Figure 3.1 shows the sample size of HES database in each year.

Figure 3.1



Inland Revenue’s Taxable Income Data

Inland Revenue (IR) is the most extensive source of income information in the IDI. Inland Revenue is the New Zealand government’s revenue collection agency. This database includes tax information of all taxpayers in New Zealand consists of more than 16 million rows and 100 million observations. The raw IR data sets in the IDI contain information about income from four main IR income tax return forms:

1. Employer Monthly Schedule (EMS) provides gross earnings where PAYE (Pay As You Earn)¹ is deducted at source. The EMS consists of all wage and salary earners, withholding payments, government transfer payments, and payments from ACC (Accident Compensation Corporation)². It includes categories for

¹Employees earning a wage or salary are taxed directly from their pay. This is known as PAYE (pay as you earn).

²The Accident Compensation Corporation is the New Zealand Crown entity responsible for administering the country’s no-fault accidental injury compensation scheme, commonly referred to as the ACC scheme.

government benefits, student allowances, paid parental leave, and New Zealand Superannuation payments. The EMS is filed monthly by the employers and provides pay details of employees who work for them.

2. IR3 for self-employment (filed annually by sole traders) which includes non-zero partnership, self-employment, or shareholder salary income, as well as rental income.
3. IR4S filed by companies includes remuneration income paid to shareholders, directors, and relatives of shareholders (filed annually).
4. IR20 (formerly IR7) for partnership and look-through companies (filed annually).

In March 2014, Statistics NZ introduced derived tables. Table `data.income-tax-yr` has been used in this study. This table is comprised of all records in the Employer Monthly Schedule (EMS), plus additional records from the IR3, IR4S and IR20 tax forms. It also orders the monthly data into tax years, as the first records begin in April 1999. Thus, when referring to month 1 of this tax year, this means the month of April (e.g. year=2000, refers to the period April 1999 to March 2000). Data from IR tables is available since 1999 onward (*Info-Share*, 2020).

Figure 3.2 shows the size of IR database in each year.

3.1.2 Macro data

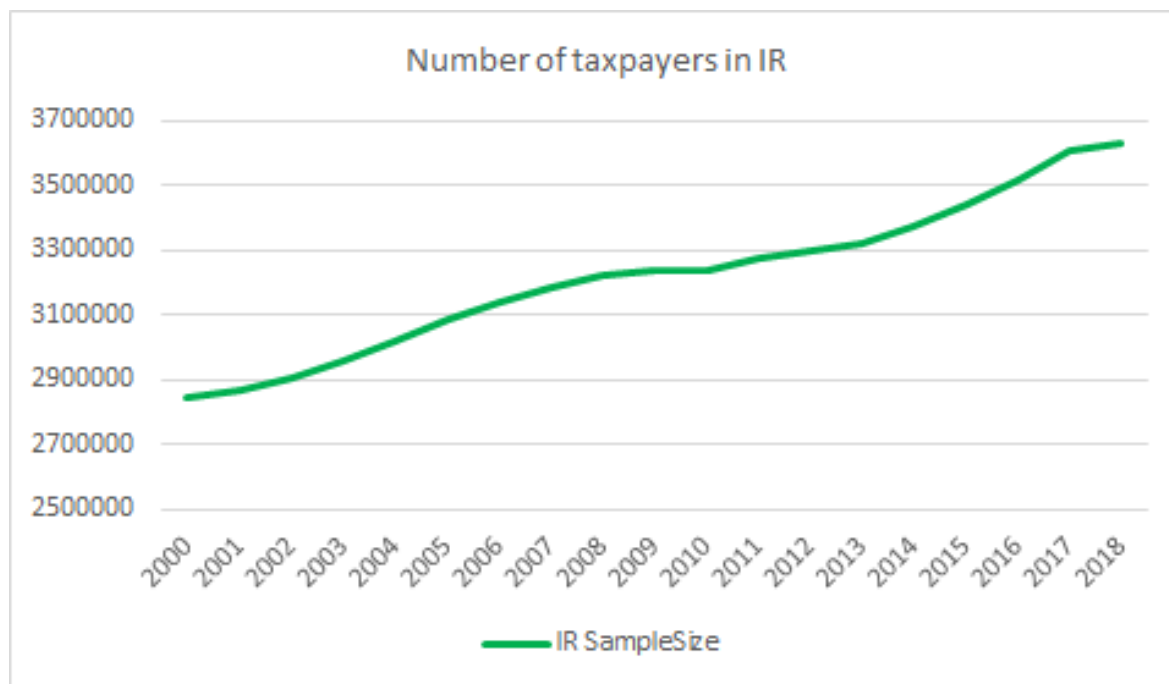
National Accounts

In this study all macroeconomic total amounts such as aggregate gross domestic products, compensation of employees, total capital income, national income, and all components to build factor Income and net national income are from Statistics New Zealand data bases available in Economic indicators tab at *infoshare* website³. National accounts available in this data base are calculated following SNA2008⁴ – SNE table

³<http://archive.stats.govt.nz/infoshare/>

⁴The System of National Accounts (SNA) is the internationally agreed standard set of recommendations on how to compile measures of economic activity. The SNA describes a coherent, consistent and integrated set of macroeconomic accounts in the context of a set of internationally agreed con-

Figure 3.2



Series, National disposable income account, Nominal, Actual, Total (Annual-March), and table Population Estimate – DPE, Estimated Resident Population by Age and Sex (1991+) (Annual-March). Figure 3.4 visualizes net national income (NNI), net labour income, and net capital income based on National Accounts in New Zealand.

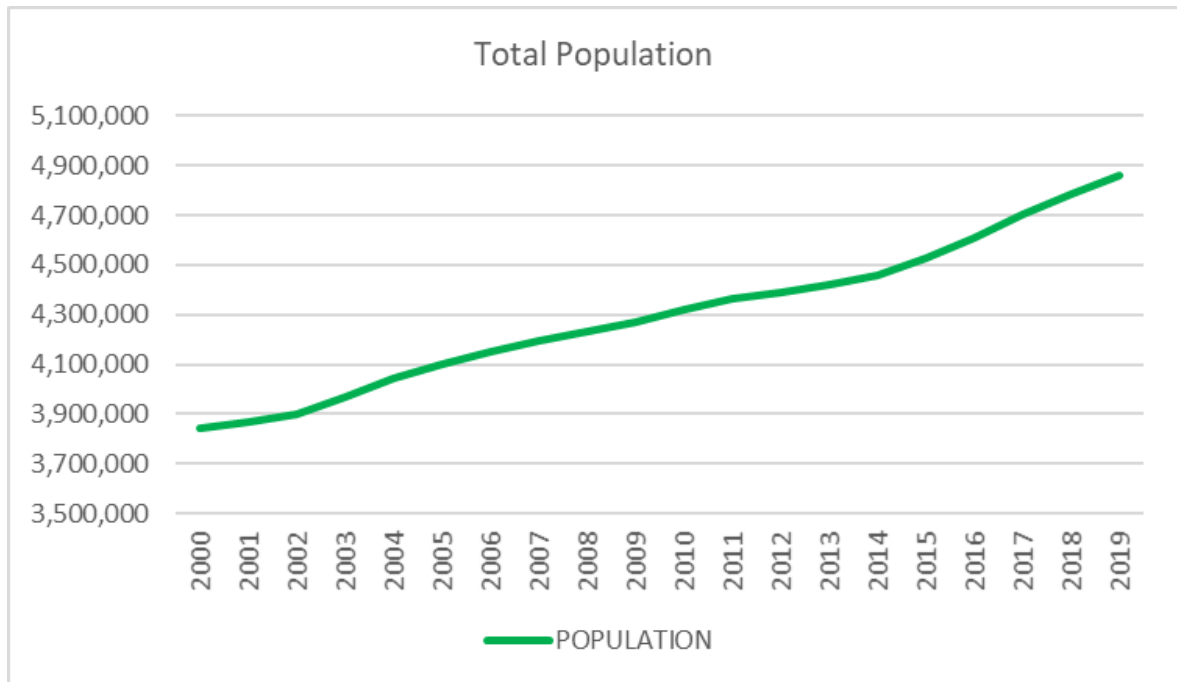
Figures 3.3 and 3.4 are based on National Accounts data in New Zealand.

3.1.3 Data Issues

As there are various issues in terms of all three databases used in this study, we view our attempt an initial draft to construct distributional national accounts for New Zealand. It will be always possible to improve findings once more reliable databases, more knowledge to impute national account components, and improved methodologies

cepts, definitions, classifications and accounting rules. In addition, the SNA provides an overview of economic processes, recording how production is distributed among consumers, businesses, government and foreign nations. It shows how income originating in production, modified by taxes and transfers, flows to these groups and how they allocate these flows to consumption, saving and investment. Consequently, the national accounts are one of the building blocks of macroeconomic statistics forming a basis for economic analysis and policy formulation. The SNA is intended for use by all countries, having been designed to accommodate the needs of countries at different stages of economic development. It also provides an overarching framework for standards in other domains of economic statistics, facilitating the integration of these statistical systems to achieve consistency with the national accounts.

Figure 3.3



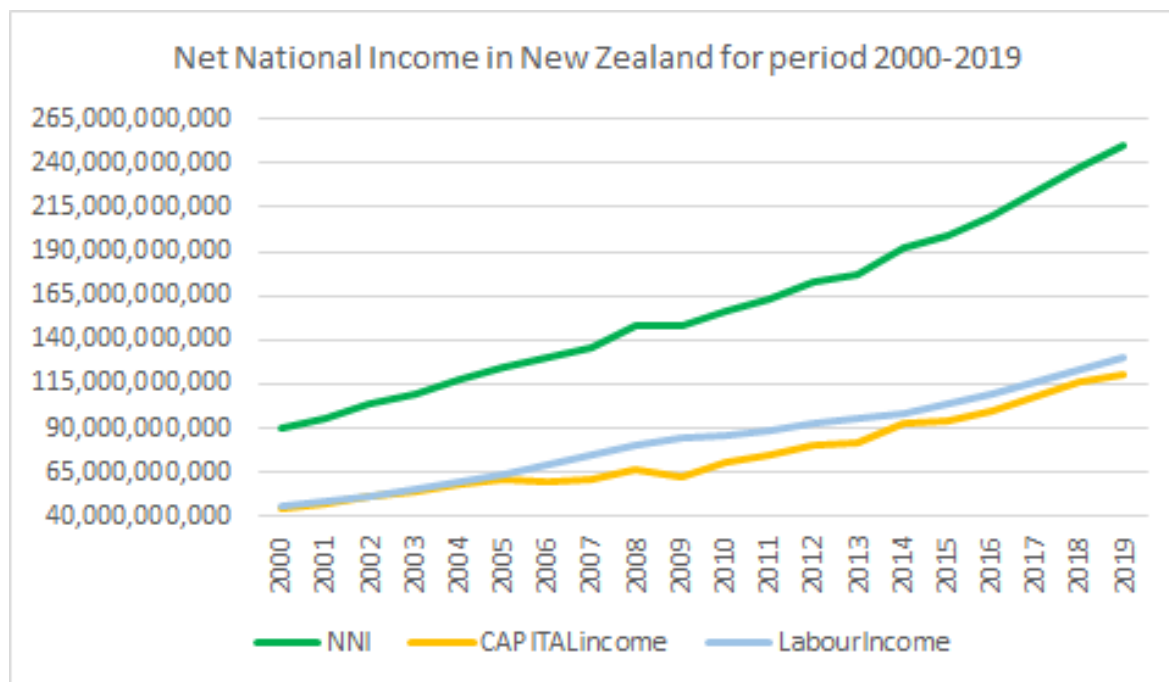
are available. Considering this matter, we also have made assumptions in relevant sections that may seem overly simplified and unrealistic⁵.

Furthermore, although the national accounts aggregate all the available information from survey, balance sheets, tax data and so on, they are still imperfect. *Zucman (2013)*

Those findings that rely on tax data are biased as tax data excludes tax evasion and some forms of income, specifically some components of capital income, are not subject to tax and do not appear on income tax declaration that causes underestimating the inequality. In using survey data, we need to emphasize that there is a strong correlation between survey responses and administrative records that makes HES data imperfect. *Ball and Ormsby (2017)*

⁵E.g, as there are some missed data in micro databases such as rental income in capital income category (it is estimated that 2/3 of capital income are not reported in tax data(*Kergozou, 2017*)) we assume that the ratio of income components for all individuals are equal

Figure 3.4



3.2 Methodology (How to Construct DINA)

3.2.1 Overview of the approach

The methodology is applied to distribute the national accounts values across different income groups largely follows the step-by-step approach recommended by the Distributional National Accounts Guidelines: Methods and Concepts Used in WID.world *Alvaredo et al.* (2016). We present a brief overview of the method below: Step 1: Adjust the National Accounts and count the total aggregation of labour income and capital income in National Accounts. Step 2: Designate each single individual income from micro data sources (HES or fiscal income) to capital, labour, or Pension source categories, then count sum of all income in each category. Step 3: Scale the aggregate amounts from step 2 to the total aggregation of the relevant category of adjusted National Accounts from step 1. Step 4: Calculate a scale factor and use it for scaling individual's income and build the considered series.

3.2.2 The Income Concepts

The income concepts that are used in DINA series are defined in the same manner in all countries and time periods, and aim to be independent from the fiscal legislation of the given country/year. Also, all national account concepts are codified as the official definition in SNA (System of National Accounts) version 2008. One of the central limitations of national account series, specifically GDP, is that they do not provide any information about the extent to which the different social groups benefit from growth. Apart from coping with the gap between national accounts and tax data, DINA series overcome the mentioned lack of national account series as well. The four basic pretax and post-tax income concepts that are useful to measure income inequality are anchored upon the notion of "national income" that is defined as GDP, minus consumption of fixed capital, plus net foreign income. Including capital depreciation would artificially inflate the economic income of capital owners as it does not allow for consumption or the accumulation of wealth. Additionally, including foreign income is important as foreign investment income can be significant for top income earners. At the individual level, income differs whether it is observed before or after the operation of the pension system and government redistribution. We therefore define three income concepts that all add up to national income: pretax factor income, pretax national income, and post-tax national income. The key difference between pretax factor income and pretax national income is the treatment of pensions, which are counted on a contribution basis for pretax factor income and on a distribution basis for pretax national income. Post-tax national income deducts all taxes and adds back all public spending, including public goods consumption. It is worth mentioning that aggregate pretax national income, pretax factor income, and post-tax national income are all equal to aggregate national income, as defined by SNA 2008. As we explain below, we focused on Pretax concepts. Calculating Post-Tax DINA series is out of this study's scope.

3.2.3 Factor Income

Pretax factor income, which for simplicity we sometime refer to as “factor income”, is equal to the sum of all pretax income flows accruing directly or indirectly to the owners of the production factors, labour and capital, before taking into account the operation of the tax/transfer system (including indirect taxes), and before taking into account the operation of the pension system. One problem with this concept of income is that retirees typically have little factor income, so that the inequality of factor income tends to rise mechanically with the fraction of old-age individuals in the population, potentially biasing comparisons over time and across countries. However, we draw useful insights from this concept as well.

3.2.4 National Income

Pretax national income is equal to the sum of all pretax income flows accruing to the individual owners of the production factors, labour and capital, before taking into account the operation of the tax/transfer system, but after taking into account the operation of the pension system. That is, the difference with factor income is that pretax income includes Social Security (old-age, survivor, and disability insurance) benefits, unemployment insurance benefits, and private pension benefits, while it excludes the contributions to Social Security, private pensions, and unemployment insurance.

3.2.5 DINA

The outstanding advantage of our study is using granular income data include a breakdown of income by source. We go through our calculations using two separate frameworks since income source classifications table in HES data are different from those used by IR tax data. This gives rise to separate graphs. In appendix A, income source classification tables for either databases are attached. In both calculations we capture 100 percent of national account by construction.

From Taxable Income to Pretax Factor Income

The starting point of our distributional national accounts is the individual micro data reported by New Zealand Inland Revenue which is the largest sample of taxpayers information containing the income source categories. Tax data contains information about most of the components of Factor and National Income. However, they miss some parts of these components as they are untaxed. The first step is to designate each single income from each individual to one of the three following categories: 'labour income', 'capital income', or 'pension, Benefit, and transfer'. For that, we imputed wage and salary⁶, sole trader PAYE deducted income⁷, sole trader Withholding income⁸, partner income PAYE deducted⁹, partner income Withholding¹⁰, company director/shareholder PAYE deducted¹¹, and company director/shareholder WHT¹² to the labour income category. And partnership income¹³, director/shareholder income¹⁴, sole trader Income¹⁵, and rental income¹⁶ to the capital income category¹⁷. Then, for each category sum up total income of each individual in that category. To distribute Factor Income based on its definition we need to find the total for labour income and capital income. So, sum up all individual incomes in each category to reach the total labour income and total capital income. The next step is comparing these results with their correspondent data in national accounts. As we already mentioned there are always a gap between these results. It is mostly because of the fact that some of the incomes which appear in national accounts are not reported in the tax returns. For instance, rental incomes in national accounts include rent of the houses occupied by their owners, while tax data excludes information about this type of income. To

⁶'WS'

⁷'S01'

⁸'S02'

⁹'P01'

¹⁰'P02'

¹¹'C01'

¹²'C01'

¹³'P00'

¹⁴'C00'

¹⁵'S00'

¹⁶'S03'

¹⁷For items which are known as self employment such as Sole Trader, we consider sole trader PAYE deducted and sole trader Withholding as labour income and sole trader as capital income. That is same for the partner income, the company director/shareholder income, and the partnership income.

cope with the mentioned problem, DINA method scales up/down the income of each individual using a scale factor which is built based on the process below. In this case, we need to find aggregate Labour and Capital income from factor national income (Factor Income). As we mentioned before, according to SNA, aggregate income from two factors of production('labour income' and 'capital income') add up to the Factor Income. Using relevant table in national accounts database¹⁸ containing GNI¹⁹, CFC²⁰ and Total Compensation of employees, as well as following the equations below, we obtained the total amount of capital income and Factor Income.

$$NNI = GNI - CFC \quad (3.1)$$

$$FactorIncome = NNI \quad FactorIncome = LabourIncome + CapitalIncome \quad (3.2)$$

$$TotalCapitalIncome = FactorIncome - TotalCompensationofemployees \quad (3.3)$$

DINA Scale Factor:

Scale factor for labour income:

$$WDINA = \frac{AverageLabourIncomefromnationalaccounts}{AverageLabourIncomefromtaxdata} \quad (3.4)$$

Scale factor for capital income:

$$CDINA = \frac{AverageCapitalIncomefromnationalaccounts}{AverageCapitalIncomefromtaxdata} \quad (3.5)$$

¹⁸SNE table Series, National disposable income account, Nominal, Actual, Total (Annual-March)

¹⁹Gross National Income

²⁰Consumption of Fixed Capital

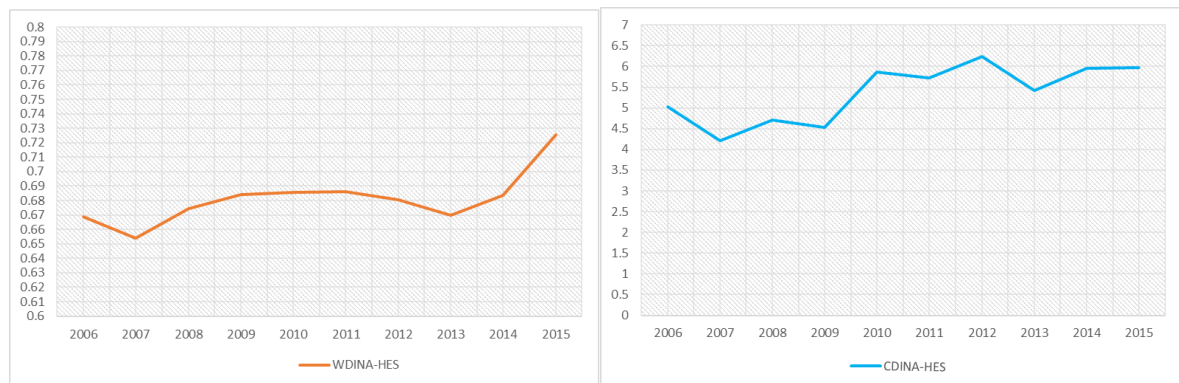
DINA for each individual:

$$IDINA = c \times CDINA + w \times WDINA \quad (3.6)$$

'c' is total income from capital sources and 'w' is total income from labour income sources for each individual.

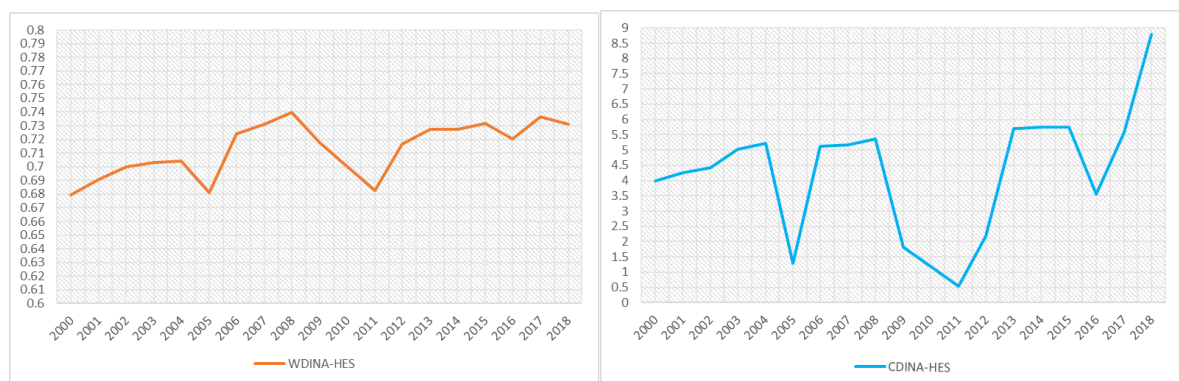
Figure 3.5 compares the scale factors of labour and capital income which are built based on HES data.

Figure 3.5: Scale factors based on HES data



The figures above show the fact that the Household Survey data does not cover the data of capital income properly so its scale factor is significantly bigger than Labour income scale factor. Labour and capital scale factors which are built based on IR data are shown in figure 3.6.

Figure 3.6: Scale factors based on IR data



DINA for each individual:

$$IDINA = c \times CDINA + w \times WDINA \quad (3.7)$$

'c' is total income from capital sources and 'w' is total income from labour income sources for each individual.

We constructed a series from IDINA for all individuals and for each year.

The last step is to find 50th, 90th, 99th percentiles as well as share of bottom 50 percentiles, middle class (50th to 90th percentiles), top 10 percentiles and top 1 percentile using the above series.

From Taxable Income to Pretax National Income

For calculating Pretax National Income, the only difference is to take into account the incomes belong to the 'Pensions, benefits, and transfers' category with a contribution approach. We imputed income from ACC (Accident Compensation corporation)²¹, Pension²², Benefit²³, Paid Parental Leave²⁴, and Student Allowance²⁵ sources to this category. To avoid double counting, we deducted contribution of each individual in this category from their income. In fact, to obtain pension definition of national income we added income receipt from mentioned sources and removed contribution to the private and public schemes for these sources.²⁶ In New Zealand tax system, share of benefits from total income has not been determined. However, to ensure consistency with basic methodology, the total contribution is assumed to be in proportion to total tax income paid. (The computer code is available on appendix B²⁷).

In calculation process above, income from mentioned sources considered as labour income.

²¹'CLM'

²²'PEN'

²³'BEN'

²⁴'PPL'

²⁵'STU'

²⁶New Zealand has a public pension system (New Zealand Superannuation (SNZ)) where all pensioners receive the same amount, NZS is recorded as 'BEN' in tax data.

²⁷Computer code is also in <https://github.com/zarsoleimani/read-me1.git>

From Income Reported in Survey to Factor Income

We used the same instruction method to build DINA using HES data, but it is worth stating the imputation approach here as income sources are labeled differently in this database. Table includes all income sources of HES data is available in appendix.

In group '1', we marked total subgroups '1.1' (income from wage and salaries), '1.3' (income from casual jobs and hobbies), and 70 percent²⁸ of '1.2'(self-employment income) as labour income. Also, total group '2'(investment income including all types of interest, dividend, rent,royalties and other investment incomes),30 percent of self-employment income, trust in group '4'(Other regular and recurring income derived in New Zealand), and trust, interest, dividends and rents in group '5'(overseas income²⁹) are imputed as Capital Income. All group '3' (total pension, superannuation and government transfers), Income from annuities, Income from maintenance, child support or alimony, Income from income protection insurance scheme, Earnings compensation as a dependent of a victim, Job-related superannuation,and Other private superannuation from group '4'(Other regular and recurring income derived in New Zealand) and Pensions, War pensions,Job or private superannuation,Maintenance, child support or alimony payments and Life insurance from overseas (group '6') are in the third category.³⁰

The rest of the process constructing IDINA series and finding share of each percentile have been applied in here exactly as they were for the tax data.

²⁸for this allocation we followed *Kergozou (2017)*

²⁹following the definition of national income which is GDP minus consumption of fixed capital plus net income received from abroad, we add overseas income in calculation process. Foreign income is important because because investment income can be significant for top income earners.

³⁰other income sources which are in non of above categories are left to discuss

4 Results and Discussions

4.1 Introduction

This section presents the results, beginning in section 4.2 which reports distributional national account (Factor Income) evolution using HES database over 2006-2015. Subsection 4.3.1 provides DINA series which are built by fiscal income data and Factor Income concept in period 2000-2018. The DINA series which show distribution of National Income using tax data are stated in subsection 4.3.2

4.2 DINA constructed by HES database

Data from the HES is available for 2006 to 2015, and the achieved sample size is around 300,000 households. Figure 4.1 shows how the pretax income shares of the top 10 percentiles, middle class, and bottom 50 percentiles have evolved since the 2006.

The gap between two income groups and two other categories reaches to the minimum amount in 2009 while top 10 percentiles experience the minimum amount of their income share and middle class experience the maximum amount of that. This phenomenon which coincided with the global financial crises. So we can say that crisis made a more equal society for a short period of time. Share of income for bottom 50 percentiles has not changed significantly during this period and it has been between 12 and 14 percent. Although we can see a slightly more income inequality at the end of the period, in general inequality has not changed over 2006-2015.

Figure 4.2 describes breaking down the share of income of top 10 percentiles into top 1 percentile and the next 9 income percentiles. As lines related to the share of 1 top percentile and share of percentiles 90-99 show, in spite of temporary changes in

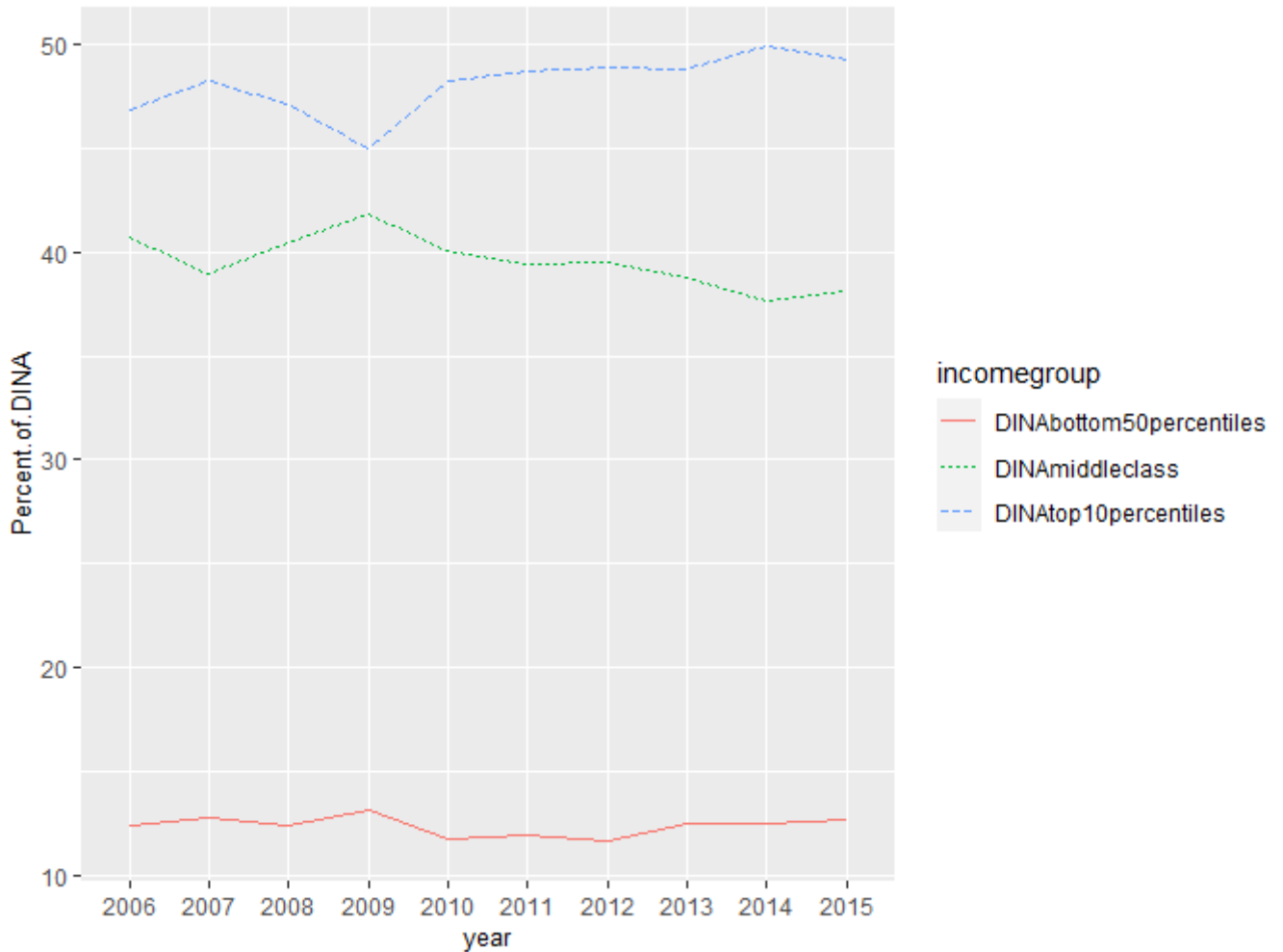


Figure 4.1: National Account (Factor Income) shares by income group-HES data

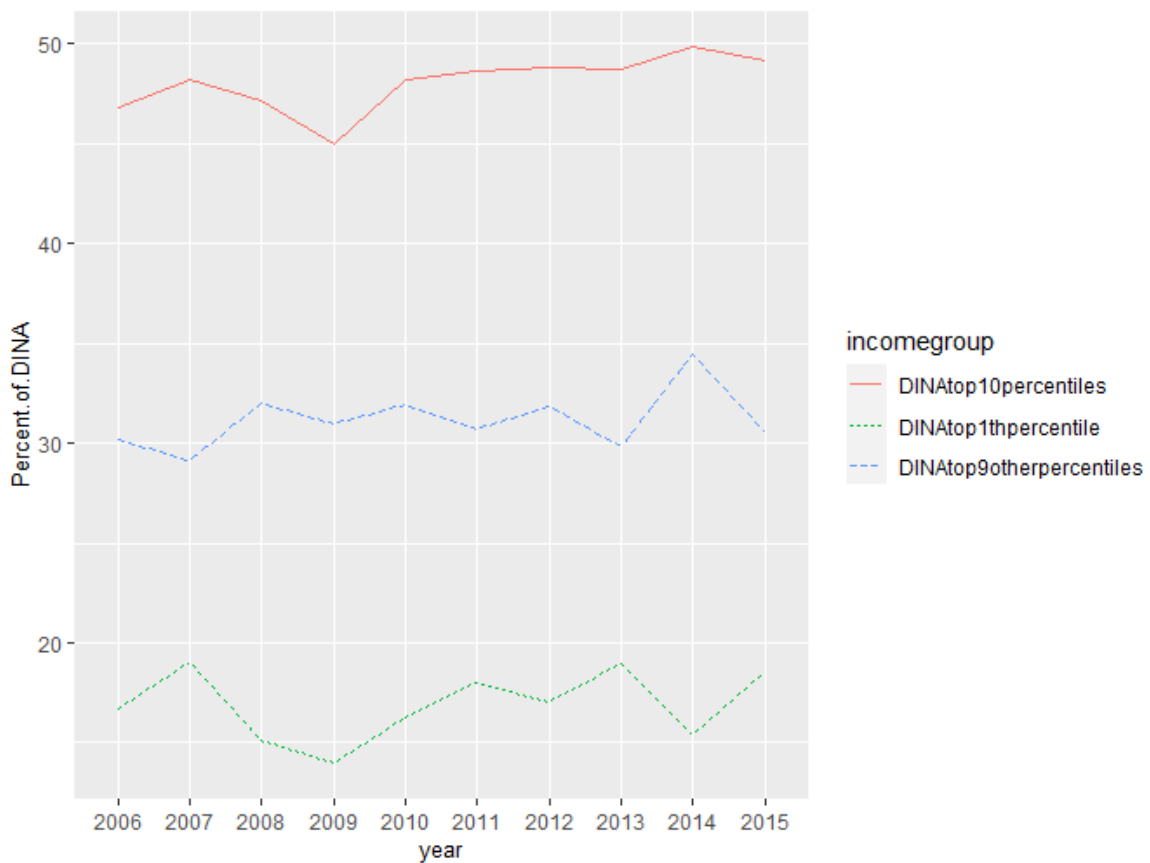
their share of income, they have not changed very much over the mentioned period. In year 2009 we observe an obvious reduction in income share of top 1 percentile that causes sharp drop of income share of top 10 percentiles which coincided with financial crises that started in 2008.

4.3 DINA constructed by IR database

4.3.1 Distributional Factor Income:

Figure 4.3 compares share of bottom 50 percentiles, middle class, and top 10 percentiles from Factor Income. It illustrates that the share of bottom 50 percentiles has increased only 2 percent in that period (8 percent in 2000 to 10 percent in 2018) and we cannot

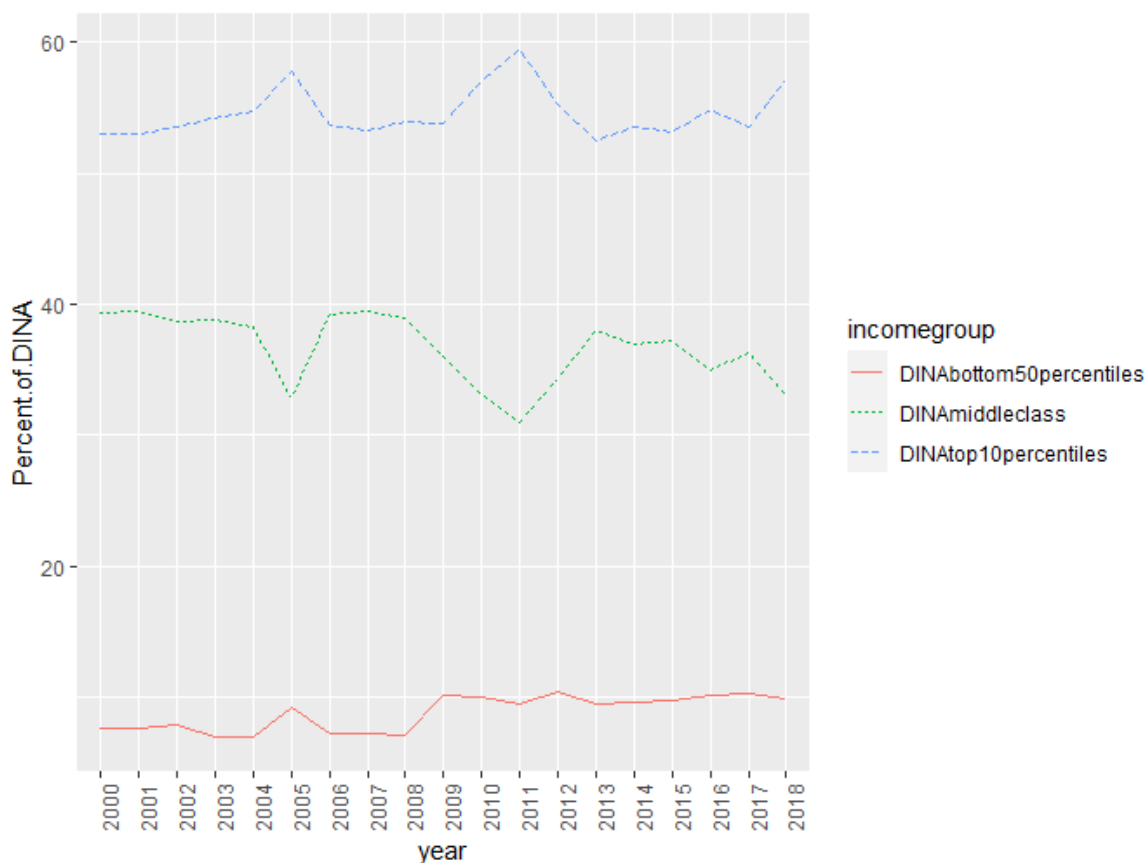
Figure 4.2: National Account (Factor Income) shares by Top income groups-HES data



see any sudden change in that over 20 years. Although the level of the share of other two income groups at the end of the period are almost same as the level of them at year 2000, we can see a slightly bigger gap between middle class and 10 top income group after almost 20 years that shows more income inequality between them. In years 2005 and 2011 Factor income was distributed significantly different than other years and income inequality is at the maximum level for those 2 groups specially in year 2011. In year 2011 share of middle class in New Zealand is at the minimum level and has dropped to 30 percent of total factor income while share of richest 10 percent receive almost 60 percent at the time.

We can investigate distribution of Factor Income of top 10 income percentiles more accurately with a breakdown to top 1 percentile and other 9 income percentiles (90th-99th percentiles). As is clearly shown in figure4.4 the increase of share of top 10 percentiles in years 2005 and 2011 is absolutely because of a significant increase in share of top 1 percentile that even has not covered with a drastic drop in share of

Figure 4.3: National Account (Factor Income) shares by income groups-IR data



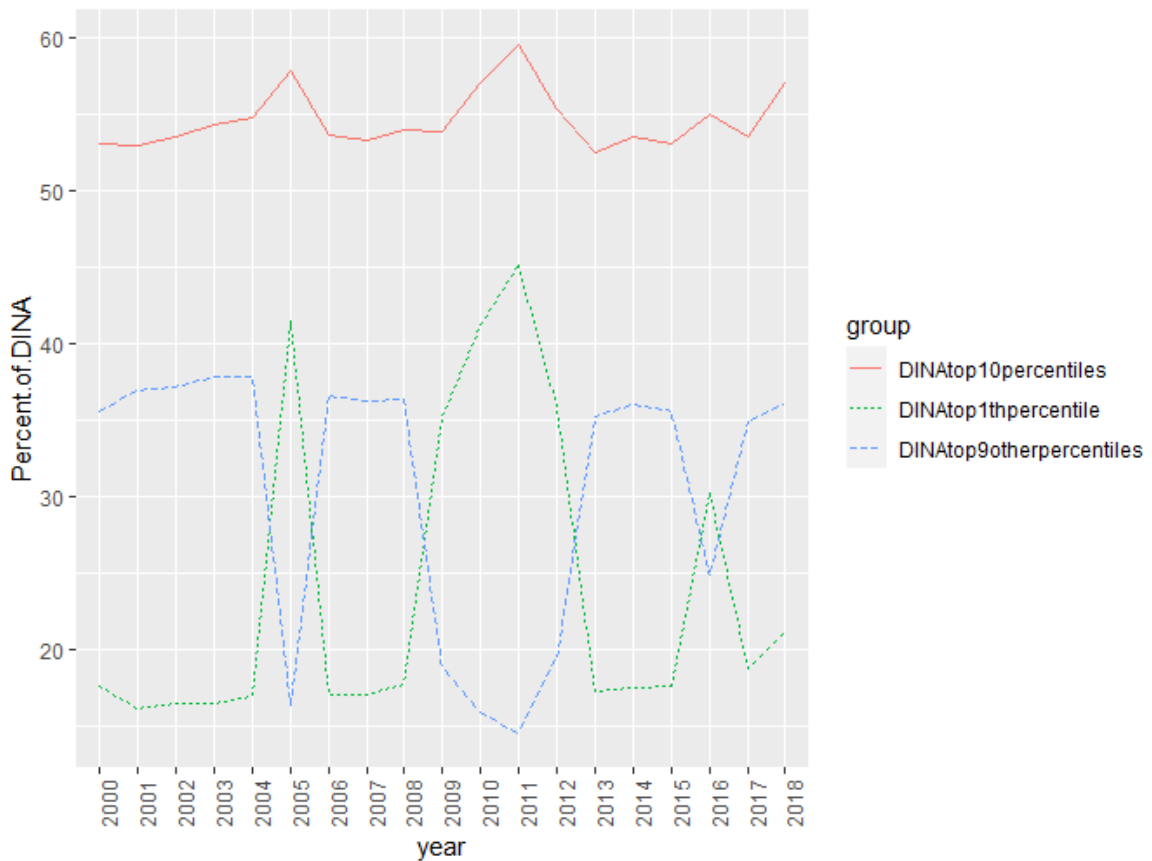
other 9 percentiles from Factor Income. Figure 4: Factor Income shares by top income groups

4.3.2 Distributional National Income

The main benchmark for comparing income inequality between countries is DINA using National Income concept. Figure 4.5 and 4.6 display share of bottom 50 percentiles, middle-class, top 10 percentiles, as well as breaking down the share of top 10 percentiles to share of top 1 percentile and other 9 percentiles from National Income in order. As is to be expected the inequality of National Income tends to be less than inequality of Factor Income which is because of taking into account pension, transfers, and social security in National Income.

For instance, whereas trend of correspondent DINA series for Factor and National Income are very similar, their level are different in favor of less income inequality in

Figure 4.4: National Account (Factor Income) shares by top income groups-IR data



National Income framework. Share of bottom 50 percentiles is between 10 and 12 percent of the National Income while it is 5 to 10 percent of Factor Income. For top 10 percentiles that is 50 to 55 percent of National Income and 55 to 60 percent of Factor Income.

This explanation is applicable for comparing other correspondent series in Factor Income and National Income frameworks.

Figure 4.6 displays share of top income groups from National Income.

Figure 4.5: National Account (National Income) shares by income groups-IR data

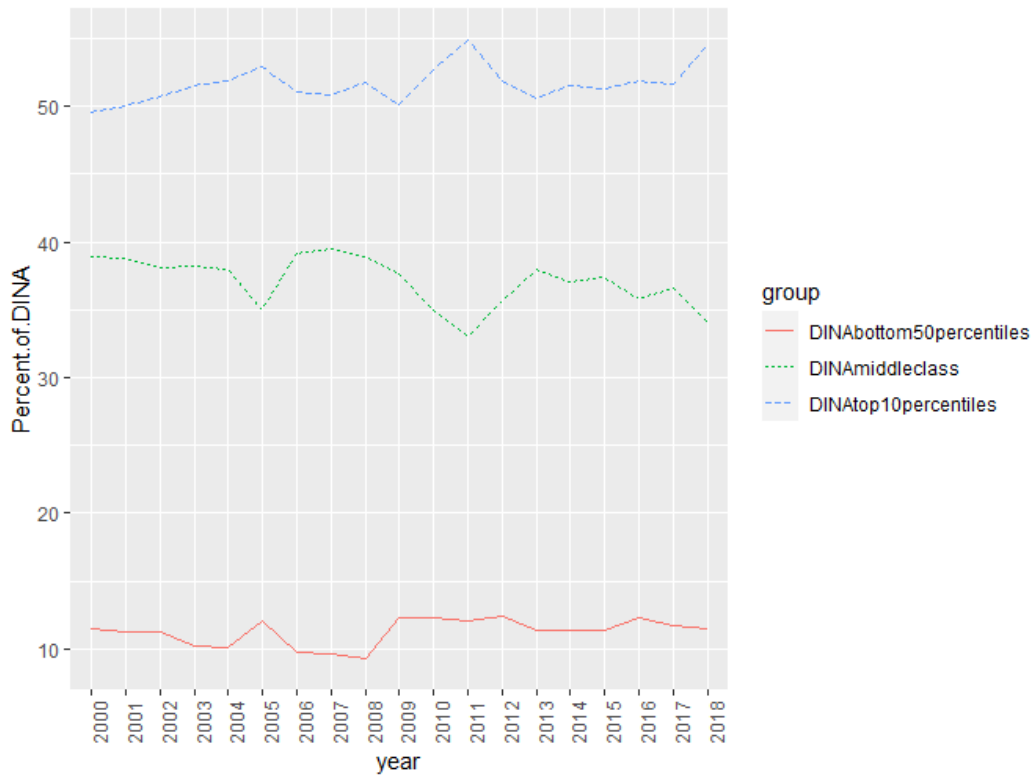
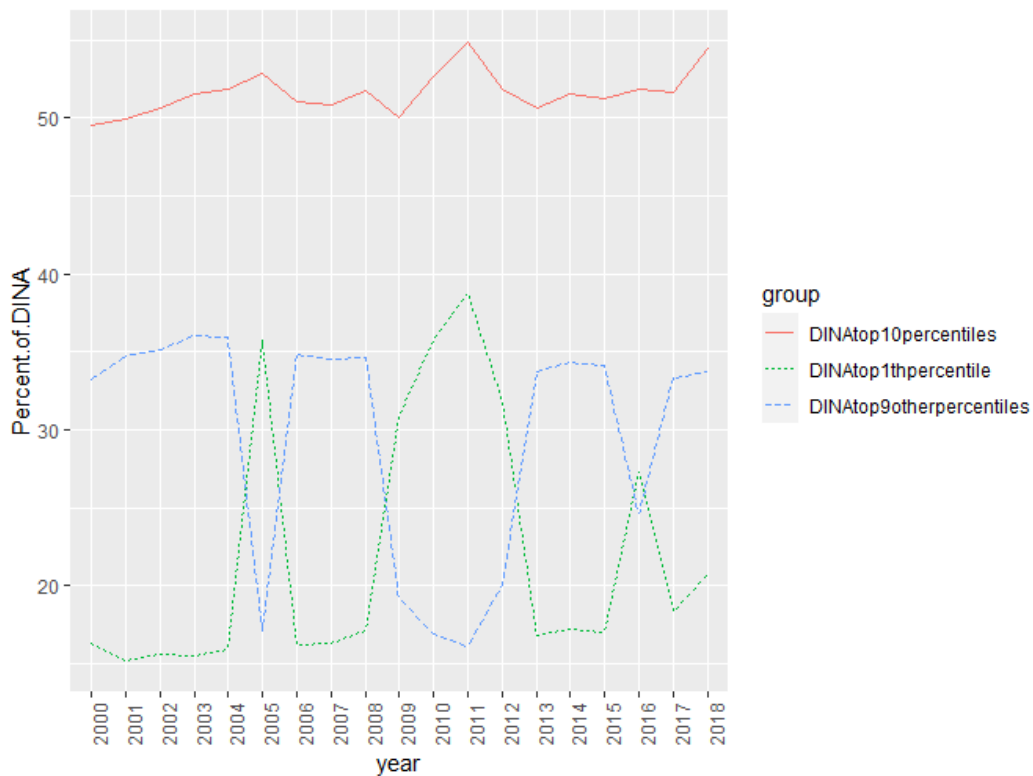


Figure 4.6: National Account (National Income) shares by top income groups-IR data



5 Conclusions and Future work

5.1 Conclusion

In this study we investigated how are national accounts distributed among income percentiles. Our series display the evolution of distribution of National Accounts (either Factor Income or National Income) using two separate data bases (IR data and HES data) so our results have shown how results are different using different database as well. Findings determine that the gap between poor and wealth has changed neither over period 2006-2015, nor in a wider period of time (2000-2018) since share of bottom 50 percentiles from National Income has been between 9 and 13 percent, share of middle class has been between 34 and 39 percent, and share of top 10 percentiles has been between 50 and 55 percent. share top 1 percentile from National Income has increased drastically in years 2005, 2011, and 2016 while share of percentiles 90th-99th has dropped in those years. We should stress again that our methods and results should be viewed not as a final product, but rather as a prototype and part of an on-going attempt to provide more and more complete and transparent inequality statistics. As better sources and methods become available, the results always can be improved accordingly.

5.2 Future directions

There are numerous manners in which these distributional national accounts could be improved upon and developed further. Firstly, this paper uses only two categories (labour and capital income) as source of income and imputes each single income for each individual to one of these two categories producing estimates. However, producing

DINA with more elaborated imputation and unfolded these sources of income to more accurate sources of income would allow for more precise estimates. Secondly, this paper only produces estimates of pre-tax national and factor income. Estimates of post-tax national and factor income would provide a comprehensive view of how government redistribution affects inequality. In addition, breaking down the series to find series which describe share of gender, share of different age groups, and share of capital (or labour) from National Income for each income percentile can reveal valuable facts in terms of distributional national accounts.

Appendices

A Classification of income sources based on HES database

Appendix A includes income classification part of HES data dictionary.

Figure A.1: Income classification in HES

| Level 1, Level 2, Level 3, Level 4 | Source of Income Classification | Source of Income |
|------------------------------------|---|------------------|
| 1 | Income from employment | |
| 1.1 | Income from wages and salaries | |
| 1.1.1 | Income from current wage and salary jobs | |
| 1.1.1.01 | Wages and salaries from 1st current job | Labour Income |
| 1.1.1.02 | Wages and salaries from 2nd current job | Labour Income |
| 1.1.1.03 | Wages and salaries from other current jobs | Labour Income |
| 1.1.1.04 | Bonuses from all current jobs | Labour Income |
| 1.1.1.05 | Commission from all current jobs | Labour Income |
| 1.1.1.06 | Directors fees, honoraria, remuneration for school-board of trustees from all current jobs | Labour Income |
| 1.1.1.07 | Other taxable income from all current jobs | Labour Income |
| 1.1.2 | Income from previous wage and salary jobs | |
| 1.1.2.01 | Wages and salaries from 1st previous job | Labour Income |
| 1.1.2.02 | Wages and salaries from 2nd previous job | Labour Income |
| 1.1.2.03 | Wages and salaries from other previous jobs | Labour Income |
| 1.1.2.04 | Redundancy from 1st previous job | Labour Income |
| 1.1.2.05 | Redundancy from 2nd previous job | Labour Income |
| 1.1.2.06 | Redundancy from other previous jobs | Labour Income |
| 1.1.2.07 | Bonuses from all previous jobs | Labour Income |
| 1.1.2.08 | Commission from all previous jobs | Labour Income |
| 1.1.2.09 | Directors fees, honoraria, remuneration for school-board of trustees from all previous jobs | Labour Income |
| 1.1.2.10 | Other taxable income from all previous jobs | Labour Income |
| 1.2 | Self-employment income | |
| 1.2.1 | Income from current self-employment jobs | |
| 1.2.1.01 | Self-employment income from 1st current job | Self-employment |
| 1.2.1.02 | Self-employment income from 2nd current job | Self-employment |
| 1.2.1.03 | Self-employment income from other current jobs | Self-employment |
| 1.2.2 | Income from previous self-employment jobs | |
| 1.2.2.01 | Self-employment income from 1st previous job | Self-employment |
| 1.2.2.02 | Self-employment income from 2nd previous job | Self-employment |
| 1.2.2.03 | Self-employment income from other previous jobs | Self-employment |

Figure A.2: Income classification in HES

| Source of Income Classification | | Source of Income |
|------------------------------------|--|----------------------|
| Level 1, Level 2, Level 3, Level 4 | | |
| 1.3 | Income from casual jobs and hobbies | |
| 1.3.0 | Income from casual jobs and hobbies | |
| 1.3.0.01 | Income from all casual jobs | Labour Income |
| 1.3.0.02 | Income from hobbies | Labour Income |
| 1.4 | Income from earnings compensation | |
| 1.4.0 | Income from earnings compensation | |
| 1.4.0.00 | Income from earnings compensation | Pension and Transfer |
| 2 | Investment Income | |
| 2.1 | Income from interest | |
| 2.1.0 | Income from interest | |
| 2.1.0.01 | Interest from deposits with banks, building societies, or other financial institutions | Capital Income |
| 2.1.0.02 | Interest from bonds | Capital Income |
| 2.1.0.03 | Interest from stocks or money market funds | Capital Income |
| 2.1.0.04 | Interest from mortgages or loans | Capital Income |
| 2.1.0.05 | Interest from debentures or securities | Capital Income |
| 2.2 | Income from dividends | |
| 2.2.0 | Income from dividends | |
| 2.2.0.00 | Dividends from New Zealand companies | Capital Income |
| 2.3 | Income from rent | |
| 2.3.0 | Income from rent | |
| 2.3.0.01 | Income from rent | Capital Income |
| 2.3.0.02 | Income from Maori or other leased land | Capital Income |
| 2.4 | Income from royalties | |
| 2.4.0 | Income from royalties | |
| 2.4.0.00 | Income from royalties | Capital Income |
| 2.5 | Other investment income | |
| 2.5.0 | Other investment income | |
| 2.5.0.01 | Returns from unit trust | Capital Income |
| 2.5.0.02 | Returns from managed funds | Capital Income |
| 2.5.0.03 | Income from partnership as a non-working shareholder or proprietor | Capital Income |
| 2.5.0.04 | Other sources of investment income | Capital Income |

Figure A.3: Income classification in HES

| Source of Income Classification | | Source of Income |
|------------------------------------|---|----------------------|
| Level 1, Level 2, Level 3, Level 4 | | |
| 3 | Income from government transfers | |
| 3.1 | Pension and superannuation income | |
| 3.1.0 | Pension and superannuation income | |
| 3.1.0.01 | New Zealand Superannuation | Pension and Transfer |
| 3.1.0.02 | Veterans Pension | Pension and Transfer |
| 3.1.0.03 | War Disablement Pension | Pension and Transfer |
| 3.1.0.04 | Surviving Spouse Pension | Pension and Transfer |
| 3.1.0.05 | Other type of New Zealand government pension | Pension and Transfer |
| 3.2 | Other government transfer income | |
| 3.2.0 | Other government transfer income | |
| 3.2.0.01 | In work tax credit / Child tax credit | Pension and Transfer |
| 3.2.0.02 | Minimum family tax credit | Pension and Transfer |
| 3.2.0.03 | Parental tax credit | Pension and Transfer |
| 3.2.0.04 | Family tax credit paid by IRD (excluding WINZ) | Pension and Transfer |
| 3.2.0.05 | Paid Parental Leave paid by Inland Revenue | Pension and Transfer |
| 3.2.0.06 | Other support paid by Inland Revenue | Pension and Transfer |
| 3.2.0.07 | Unemployment Benefit (discont) | Pension and Transfer |
| 3.2.0.08 | Sickness Benefit (discont) | Pension and Transfer |
| 3.2.0.09 | Domestic Purposes Benefit (discont) | Pension and Transfer |
| 3.2.0.10 | Invalids Benefit (discont) | Pension and Transfer |
| 3.2.0.11 | Widows Benefit (discont) | Pension and Transfer |
| 3.2.0.12 | Orphans or Unsupported Childs Benefit | Pension and Transfer |
| 3.2.0.13 | Youth Payment / Young Parent payment (formerly independent youth) | Pension and Transfer |
| 3.2.0.14 | Emergency Benefit | Pension and Transfer |
| 3.2.0.15 | Emergency Maintenance Allowance | Pension and Transfer |
| 3.2.0.16 | Other type of main benefit paid by Work and Income (discont) | Pension and Transfer |
| 3.2.0.17 | Accommodation Supplement | Pension and Transfer |
| 3.2.0.18 | Disability Allowance | Pension and Transfer |
| 3.2.0.19 | Child Disability Allowance | Pension and Transfer |
| 3.2.0.20 | Family tax credit paid by Work and Income | Pension and Transfer |
| 3.2.0.21 | OSCAR subsidy | Pension and Transfer |
| 3.2.0.22 | Special Needs Grant | Pension and Transfer |

Figure A.4: Income classification in HES

| Level 1, Level 2, Level 3, Level 4 | Source of Income Classification | Source of Income |
|------------------------------------|---|----------------------|
| | 3.2.0.23 Temporary Additional Support (formerly special benefit) | Pension and Transfer |
| | 3.2.0.24 Recoverable Assistance Payment (discount) | Pension and Transfer |
| | 3.2.0.25 Special Transfer Allowance (discount) | Pension and Transfer |
| | 3.2.0.26 Other type of supplement or allowance paid by Work and Income | Pension and Transfer |
| | 3.2.0.27 Student Allowance | Pension and Transfer |
| | 3.2.0.28 Jobseeker Support | Pension and Transfer |
| | 3.2.0.29 Sole Parent Support (for those with children under 14) | Pension and Transfer |
| | 3.2.0.30 Supported Living Payment | Pension and Transfer |
| | 3.2.0.31 Other Type of Payment | Pension and Transfer |
| 4 | Other regular and recurring income derived in New Zealand | |
| | 4.1 Private superannuation income | |
| | 4.1.0 Private superannuation income | |
| | 4.1.0.01 Job-related superannuation | Pension and Transfer |
| | 4.1.0.02 Other private superannuation | Pension and Transfer |
| | 4.2 Other regular and recurring income derived in New Zealand | |
| | 4.2.0 Other regular and recurring income derived in New Zealand | |
| | 4.2.0.01 Income from trusts | Capital Income |
| | 4.2.0.02 Income from annuities | Pension and Transfer |
| | 4.2.0.03 Income from maintenance, child support or alimony | Pension and Transfer |
| | 4.2.0.04 Income from educational scholarships | Pension and Transfer |
| | 4.2.0.05 Income from income protection insurance scheme | Pension and Transfer |
| | 4.2.0.06 Earnings compensation as a dependent of a victim | Pension and Transfer |
| | 4.2.0.07 Other sources of regular and recurring income derived in New Zealand | Pension and Transfer |
| 5 | Overseas income | |
| | 5.1 Overseas income | |
| | 5.1.0 Overseas income | |
| | 5.1.0.01 Interest payments | Capital Income |
| | 5.1.0.02 Dividend payments | Capital Income |
| | 5.1.0.03 Rent | Capital Income |
| | 5.1.0.04 Royalties | Capital Income |
| | 5.1.0.05 Social security benefits | Pension and Transfer |
| | 5.1.0.06 Pensions | Pension and Transfer |

Figure A.5: Income classification in HES

| Level 1, Level 2, Level 3, Level 4 | Source of Income Classification | Source of Income |
|------------------------------------|---|----------------------|
| | 5.1.0.07 War pensions | Pension and Transfer |
| | 5.1.0.08 Job or private superannuation | Pension and Transfer |
| | 5.1.0.09 Trusts | Capital Income |
| | 5.1.0.10 Maintenance, child support or alimony payments | Pension and Transfer |
| | 5.1.0.11 Inheritances | Capital Income |
| | 5.1.0.12 Life insurance claims | Pension and Transfer |
| | 5.1.0.13 Gifts of money from overseas households (including family) | Capital Income |
| | 5.1.0.14 Other sources of overseas income | Capital Income |
| 6 | Irregular income | |
| | 6.1 Irregular income | |
| | 6.1.0 Irregular income | |
| | 6.1.0.01 Gifts of money from other New Zealand households | Capital Income |
| | 6.1.0.02 Income from inheritances | Capital Income |
| | 6.1.0.03 Income from matrimonial settlement | Capital Income |
| | 6.1.0.04 Irregular maintenance or child support payments | Pension and Transfer |
| | 6.1.0.05 Lump sum payments from life insurance policies | Pension and Transfer |
| | 6.1.0.06 Educational lump sum bursaries and prizes | Pension and Transfer |
| | 6.1.0.07 Other sources of irregular income | Capital Income |

B R code

Github address of Rcode: <https://github.com/zarisoleimani/read-me1.git>

1. R code for DINA based on HES data:

```
# Code for DINA based on HES data

# for-loop to calculate following variables for
each year (2006 –2015)

## each individual has IDINA=wDINA+cDINA,
## wDINA= individual's wage each year*
(Average National Wage Income/
average wage income regarding to micro-data)
## cDINA=individual's capital income each year*
(Average National Capital Income/
average capital income in micro-data)
## DINA=SUM(IDAINA)

lincome=NA,

wDINA=NA,
```

```
HES<- data.frame(hes_inc_year_nbr=c(2006:2015),
                 numberofmicroindividuals=NA,
                 totalmicrowage=NA,
                 totalmicrocapitalincome=NA,
                 averagemicrowage=NA,
                 averagemicrocapitalincome=NA,
                 averagenationalwage=NA,
                 averagenationalcapita
                 cDINA=NA,
                 IDINA=NA,
                 DINA=NA,
                 bottom50thpercentile=NA,
                 bottom90thpercentile=NA,
                 bottom99thpercentile=NA,
                 DINAbottom50percentiles=NA,
                 DINAbottom90thpercentile=NA,
                 DINAmiddleclass=NA,
                 DINAtop10percentiles=NA,
                 DINAtop1thpercentile=NA,
                 DINAtop9otherpercentiles=NA,
                 wageshareofDINAbottom50thpercentile=NA,
                 wageshareofDINAtop1thpercentile=NA,
                 wageshareofDINAmiddleclass=NA,
                 wageshareofDINAtop10percentiles=NA,
                 wageshareofDINAothertop9percentiles=NA,
                 femalshareofDINAbottom50thpercentile=NA,
                 femalshareofDINAtop1thpercentile=NA,
                 femalshareofDINAmiddlesclass=NA,
                 femalshareofDINAothertop9percentiles=NA,
                 femalshareofDINAtop10percentiles=NA)
```

```
for (i in c(2006:2015)) {

# each individual might have income from different sources
  (either WAGE or CAPITAL)
# finding sum of income from identical sources to calculate
  wDINA, cDINA and lDINA for each individual

yi<- filter(DINADATAf, hes_inc_year_nbr==i )
  sample0<- aggregate(hes_inc_amount_amt~snz_hes_uid , data=yi ,

FUN=sum)

# number of individuals in survey each year
HES$numberofmicroindividuals[HES$hes_inc_year_nbr==i]
<- length(sample0$snz_hes_uid)

# selecting incomes only from wage source using source
income's codes

## wage source

yiwage<- filter(yi, str_detect(hes_inc_income_source_code ,

'^1.1.|^1.3.'))

## 70% OF SELF-Employment consider as Wage & Salary

yi0wage<- filter(yi, str_detect(hes_inc_income_source_code ,
```

```

'^1.2. ')
yi0wage$hes_inc_amount_amt<- 0.7*(yi0wage$hes_inc_amount_amt)

yiTOTALwage<- rbind(yiwage ,yi0wage)

# average wage for each year based on micro-data
sample1<- aggregate(hes_inc_amount_amt~snz_hes_uid ,
data=yiTOTALwage, FUN=sum)

HES$totalmicrowage [ HES$hes_inc_year_nbr==i]<-
sum(sample1$hes_inc_amount_amt)

HES$averagemicrowage [ HES$hes_inc_year_nbr==i]<-
(HES$totalmicrowage [ HES$hes_inc_year_nbr==i ])/
(HES$numberofmicroindividuals [ HES$hes_inc_year_nbr==i ])

# average National Wage
HES$averagenationalwage
[ HES$hes_inc_year_nbr==i ]
<-
mean(yi$ 'Compensation Of Employees -
received '
[yi$hes_inc_year_nbr==i ])/
mean(yi$POPULATION [ yi$
hes_inc_year_nbr==i ])

```

```
# DINA scale for wage: WDINA=
(Average National Wage Income/

average wage income regarding to the micro-data)
WDINA<- (HES$averagenationalwage
[HES$hes_inc_year_nbr==i])/
(HES$averagemicrowage[HES$hes_inc_year_nbr==i])

# individual's weighted wage using WDINA as a scale
sample1$wDINA<- (sample1$hes_inc_amount_amt)*WDINA

## capital source
yicapital<- filter
(yi, str_detect(hes_inc_income_source_code ,
'^2.|4.2.0.01|4.2.0.05|
5.1.0.01|5.1.0.02|
5.1.0.03|5.1.0.04|5.1.0.09'))

## 30% OF SELF-Employment consider as
Capital income

yi0capital<- filter(yi, str_detect
(hes_inc_income_source_code , '^1.2.'))

yi0capital$hes_inc_amount_amt<-
```

```
0.3*(yi0capital$hes_inc_amount_amt)
```

```
yiTOTALcapital<- rbind(yicapital , yi0capital)
```

```
# average capital income for each year
```

```
based on the micro-data
```

```
sample2<- aggregate(hes_inc_amount_amt ~ snz_hes_uid ,
```

```
data=yiTOTALcapital ,
```

```
FUN=sum)
```

```
HES$totalmicrocapitalincome
```

```
[ HES$hes_inc_year_nbr==i ]<-
```

```
sum(sample2$hes_inc_amount_amt)
```

```
HES$averagemicrocapitalincome
```

```
[ HES$hes_inc_year_nbr==i ]<-
```

```
( HES$totalmicrocapitalincome
```

```
[ HES$hes_inc_year_nbr==i ] ) /
```

```
( HES$numberofmicroindividuals
```

```
[ HES$hes_inc_year_nbr==i ] )
```

```
# average National Capital income
```

```
HES$averagenationalcapitalincome
```

```
[ HES$hes_inc_year_nbr==i ]<-
```

```
mean(yi$CapitalIncome
[yi$hes_inc_year_nbr==i])/
mean(yi$POPULATION[yi$hes_inc_year_nbr==i])

# DINA scale for capitalincome: CDINA=
(Average National Capital Income/average capital
income regarding to the micro-data)
CDINA<- (HES$averagenationalcapitalincome
[HES$hes_inc_year_nbr==i])
/(HES$averagemicrocapitalincome[HES$hes_inc_year_nbr==i])

# individual's weighted capital income using
CDINA as a scale
sample2$cDINA<- (sample2$hes_inc_amount_amt)*CDINA

# lDINA=wDINA+cDINA
sample3<- merge(sample1, sample2, by="snz_hes_uid",
all=TRUE)
sample3[is.na(sample3)]<- 0

sample3$lDINA<- sample3$wDINA+sample3$cDINA

yi$lDINAindividuals<- sample3$lDINA
```

```
[match(yi$snz_hes_uid , sample3$snz_hes_uid )]
```

```
# IDINA percentiles
```

```
sample4<- aggregate  
(IDINAindividuals~snz_hes_uid ,  
data=yi , FUN=mean)
```

```
HES$bottom50thpercentile  
[HES$hes_inc_year_nbr==i]<-  
quantile
```

```
(sample4$IDINAindividuals ,  
probs = 0.50)
```

```
HES$bottom90thpercentile  
[HES$hes_inc_year_nbr==i]<-  
quantile (sample4$IDINAindividuals ,  
probs = 0.90)
```

```
HES$bottom99thpercentile  
[HES$hes_inc_year_nbr==i]<-  
quantile (sample4$IDINAindividuals ,  
probs = 0.99)
```

```
yi1<- filter (sample3 ,IDINA<=  
HES$bottom50thpercentile  
[HES$hes_inc_year_nbr==i ] )  
HES$DINAbottom50percentiles  
[HES$hes_inc_year_nbr==i]<-  
(sum(yi1$IDINA )/
```

```
sum(sample3$IDINA))*100
```

```
yi2<- filter (sample3 ,IDINA<=
HES$bottom90thpercentile
[ HES$hes_inc_year_nbr==i ])
HES$DINAbottom90thpercentile
[ HES$hes_inc_year_nbr==i]<-
(sum(yi2$IDINA)/sum(sample3$IDINA))*100
```

```
yi3<- filter (sample3 ,
IDINA<= HES$bottom99thpercentile
[ HES$hes_inc_year_nbr==i ])
HES$DINAbottom99thpercentile
[ HES$hes_inc_year_nbr==i]<-
(sum(yi3$IDINA)/sum(sample3$IDINA))*100
```

```
HES$DINAmiddleclass
[ HES$hes_inc_year_nbr==i]<-
(HES$DINAbottom90thpercentile
[ HES$hes_inc_year_nbr==i ]) -
(HES$DINAbottom50percentiles
[ HES$hes_inc_year_nbr==i ])
```

```
HES$DINAtop1thpercentile
[ HES$hes_inc_year_nbr==i]<- 100-
(HES$DINAbottom99thpercentile
[ HES$hes_inc_year_nbr==i ])
```

```

HES$DINAtop10percentiles
[ HES$hes_inc_year_nbr==i]<- 100-
(HES$DINAbottom90thpercentile
[ HES$hes_inc_year_nbr==i ])

# othertop9percentiles= percentiles90th to 99th
HES$DINAtop9otherpercentiles
[ HES$hes_inc_year_nbr==i]<-
(HES$DINAbottom99thpercentile
[ HES$hes_inc_year_nbr==i ]) -
(HES$DINAbottom90thpercentile
[ HES$hes_inc_year_nbr==i ])

}

HES

# visualizing the results:

DINAbottom50percentiles<- c
(HES$DINAbottom50percentiles
[ HES$hes_inc_year_nbr ==2006:2015])
year<- c(2006:2015)

DINAtop1thpercentile<- c(HES$DINAtop1thpercentile
[ HES$hes_inc_year_nbr ==2006:2015])
year<- c(2006:2015)

```

```
DINAmiddleclass<- c(HES$DINAmiddleclass
[HES$hes_inc_year_nbr==2006:2015])
year<- c(2006:2015)

DINAtop10percentiles<- c(HES$DINAtop10percentiles
[HES$hes_inc_year_nbr==2006:2015])
year<- c(2006:2015)

DINAtop9otherpercentiles<-
c(HES$DINAtop9otherpercentiles
[HES$hes_inc_year_nbr==2006:2015])
year<- c(2006:2015)

graph<- data.frame(supp=rep(c
("DINAbottom50percentiles",
"DINAtop1thpercentile", "DINAmiddleclass",
", "DINAtop10percentiles",
"DINAtop9otherpercentiles"),
each=10),
year=rep(c
("2006", "2007", "2008",
"2009", "2010", "2011", "2012", "2013", "2014", "2015")
,5),
DINA=c(HES$DINAbottom50percentiles
[HES$hes_inc_year_nbr==2006:2015],
```

```
HES$DINAtop1thpercentile
[HES$hes_inc_year_nbr==2006:2015],
HES$DINAmiddleclass
[HES$hes_inc_year_nbr==2006:2015],
HES$DINAtop10percentiles
[HES$hes_inc_year_nbr==2006:2015],
HES$DINAtop9otherpercentiles
[HES$hes_inc_year_nbr==2006:2015])
)
```

```
ggplotDINA<- ggplot(data=graph, aes
(x=year, y= DINA, group=supp, colour=supp))+
  geom_line()+
  geom_point()
```

```
ggplotDINA
```

2. R code for Distributional Factor Income based on IR data:

```
# Code for Distribution of Factor Income
```

```
based on IR data
```

```
# for-loop to calculate needed variables
for each year (2000–2020)
```

```
## each individual has lDINA=wDINA+cDINA,
```

```
## wDINA= individual's wage each year*
(Average National Wage Income/
average wage income regarding to micro-data)
## cDINA=individual's capital income each year
*(Average National Capital Income/
average capital income in micro-data)
## DINA=SUM(IDAINA)

for (i in c(2000:2019)) {

di<- databind1[which
(databind1$inc_tax_yr_year_nbr==i),]

IRD$numberofmicroindividuals
[IRD$inc_tax_yr_year_nbr==i]<-
length(unique(di$snz_ird_uid))

## wage source (for Factor Income version)
diwage<- di[which(di$inc_tax_yr_income_source_code==
"W&S"|di$inc_tax_yr_income_source_code==
"C02"|di$inc_tax_yr_income_source_code==
"P02"|di$inc_tax_yr_income_source_code==
"TPR"|di$inc_tax_yr_income_source_code==
"IRD"|di$inc_tax_yr_income_source_code==
"SNZ"|di$inc_tax_yr_income_source_code==
"C01"|di$inc_tax_yr_income_source_code==
"P01"|di$inc_tax_yr_income_source_code==
"S01"|di$inc_tax_yr_income_source_code=="S02"),]
```

```

IRD$totalmicrowage [ IRD$inc_tax_yr_year_nbr==i ] <-
sum( diwage$inc_tax_yr_tot_yr_amt )

IRD$averagemicrowage [ IRD$inc_tax_yr_year_nbr==i ] <- ( IRD$totalmicrowage [ IRD$
( IRD$numberofmicroindividuals
[ IRD$inc_tax_yr_year_nbr==i ] )

# average National Wage
IRD$averagenationalwage [ IRD$inc_tax_yr_year_nbr==i ]
<- ( IRD$‘ Compensation Of Employees –
received ‘ [ IRD$inc_tax_yr_year_nbr==i ] ) / ( IRD$POPULATION [ IRD$inc_tax_yr_year

WDINA <-
( IRD$averagenationalwage [ IRD$inc_tax_yr_year_nbr==i ] ) /
( IRD$averagemicrowage [ IRD$inc_tax_yr_year_nbr==i ] )

sample1 <- aggregate
( inc_tax_yr_tot_yr_amt ~ snz_ird_uid ,
data=diwage , FUN=sum )

sample1$wDINA <- ( sample1$inc_tax_yr_tot_yr_amt ) * WDINA

## capital source
dicapital <- di [ which
( di$inc_tax_yr_income_source_code ==
”C00” | di$inc_tax_yr_income_source_code == ”S00”
| di$inc_tax_yr_income_source_code == ”S03”
| di$inc_tax_yr_income_source_code == ”P00” ) , ]

```

```
IRD$totalmicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i]<-
sum(dicapital$inc_tax_yr_tot_yr_amt)
IRD$averagemicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i]<-
(IRD$totalmicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$numberofmicroindividuals
[IRD$inc_tax_yr_year_nbr==i])

# average National Capital income
IRD$averagenationalcapitalincome
[IRD$inc_tax_yr_year_nbr==i]<-
(IRD$CAPITALincome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$POPULATION[IRD$inc_tax_yr_year_nbr==i])

# DINA scale for capitalincome:
CDINA=(Average National Capital Income/
average capital income regarding to
the micro-data)
CDINA <-
(IRD$averagenationalcapitalincome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$averagemicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i])
```

```

# individual's weighted capital
income using CDINA as a scale
sample2<- aggregate
(inc_tax_yr_tot_yr_amt~snz_ird_uid ,
data=dicapital , FUN=sum)
sample2$cDINA<-
(sample2$inc_tax_yr_tot_yr_amt)*CDINA

# lDINA=wDINA+cDINA
sample3<- merge(sample1 , sample2 ,
by="snz_ird_uid" , all=TRUE)
sample3[is.na(sample3)]<- 0

sample3$lDINA<- (sample3$wDINA)+(sample3$cDINA)

di$lDINAindividuals<-
sample3$lDINA
[match(di$snz_ird_uid ,
sample3$snz_ird_uid)]

# lDINA percentiles

sample4<- aggregate
(lDINAindividuals~snz_ird_uid ,
data=di , FUN=mean)

IRD$bottom50thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-

```

```
quantile(sample4$IDINAindividuals ,
probs = 0.50)
IRD$bottom90thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
quantile(sample4$IDINAindividuals , probs = 0.90)
IRD$bottom99thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
quantile(sample4$IDINAindividuals , probs = 0.99)
```

```
di1<- sample3 [ which (sample3$IDINA<=
IRD$bottom50thpercentile
[IRD$inc_tax_yr_year_nbr==i ] ) , ]
IRD$DINAbottom50percentiles
[IRD$inc_tax_yr_year_nbr==i]<-
(sum(di1$IDINA)/sum(sample3$IDINA))*100
```

```
di2<- sample3 [ which (sample3$IDINA<=
IRD$bottom90thpercentile
[IRD$inc_tax_yr_year_nbr==i ] ) , ]
IRD$DINAbottom90thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
(sum(di2$IDINA)/sum(sample3$IDINA))*100
```

```
di3<- sample3 [ which (sample3$IDINA <=
IRD$bottom99thpercentile
[IRD$inc_tax_yr_year_nbr==i ] ) , ]
IRD$DINAbottom99thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
(sum(di3$IDINA)/sum(sample3$IDINA))*100
```

IRD\$DINAmiddleclass

```
[ IRD$inc_tax_yr_year_nbr==i ] <-  
( IRD$DINAbottom90thpercentile  
[ IRD$inc_tax_yr_year_nbr==i ] )  
-( IRD$DINAbottom50percentiles  
[ IRD$inc_tax_yr_year_nbr==i ] )
```

IRD\$DINAtop1thpercentile

```
[ IRD$inc_tax_yr_year_nbr==i ] <- 100 -  
( IRD$DINAbottom99thpercentile  
[ IRD$inc_tax_yr_year_nbr==i ] )
```

IRD\$DINAtop10percentiles

```
[ IRD$inc_tax_yr_year_nbr==i ] <- 100 -  
( IRD$DINAbottom90thpercentile  
[ IRD$inc_tax_yr_year_nbr==i ] )
```

othertop9percentiles= percentiles90th to 99th

IRD\$DINAtop9otherpercentiles

```
[ IRD$inc_tax_yr_year_nbr==i ] <-  
  
( IRD$DINAbottom99thpercentile  
[ IRD$inc_tax_yr_year_nbr==i ] )  
-( IRD$DINAbottom90thpercentile  
[ IRD$inc_tax_yr_year_nbr==i ] )
```



```
}
```

```
# visualizing the results:
```

```
DINAbottom50percentiles<- c(IRD$bottom50thpercentile  
[IRD$inc_tax_yr_year_nbr ==2000:2018])  
year<- c(2000:2018)
```

```
DINAtop1thpercentile<-  
c(IRD$DINAtop1thpercentile  
[IRD$inc_tax_yr_year_nbr ==2000:2018])  
year<- c(2000:2018)
```

```
DINAmiddleclass<-  
c(IRD$DINAmiddleclass  
[IRD$inc_tax_yr_year_nbr ==2000:2018])  
year<- c(2000:2018)
```

```
DINAtop10percentiles<- c(IRD$DINAtop10percentiles  
[IRD$inc_tax_yr_year_nbr ==2000:2019])  
year<- c(2000:2018)
```

```
DINAtop9otherpercentiles<- c(IRD$DINAtop9otherpercentiles  
[IRD$inc_tax_yr_year_nbr ==2000:2018])
```

```

year<- c(2000:2018)

graph<- data.frame(supp=rep(c
("DINAbottom50percentiles",
"DINAtop1thpercentile",
"DINAmiddleclass", "DINAtop10percentiles"
,"DINAtop9otherpercentiles"),
each=19),
year=rep(c("2000", "2001", "2002",
"2003", "2004", "2005", "2006",
"2007", "2008", "2009", "2010", "2011",
"2012", "2013", "2014", "2015", "2016", "2017", "2018"),5),
DINApretaxIRfactorIncome=c(IRD$DINAbottom50percentiles
[IRD$inc_tax_yr_year_nbr==2000:2018],
IRD$DINAtop1thpercentile
[IRD$inc_tax_yr_year_nbr==2000:2018],
IRD$DINAmiddleclass[IRD$inc_tax_yr_year_nbr==2000:2018],
IRD$DINAtop10percentiles
[IRD$inc_tax_yr_year_nbr==2000:2018],
IRD$DINAtop9otherpercentiles
[IRD$inc_tax_yr_year_nbr==2000:2018]))

ggplotDINA<- ggplot(data=graph, aes(x=year,
y=DINApretaxIRfactorIncome,
group=supp, colour=supp))+
  geom_line()+
  geom_point()

```

ggplotDINA

3. R code for Distributional National Income based on IR data:

```
# R code for Distributional National Income based on  
IR data
```

```
# for-loop to calculate needed variables for  
each year (2000–2020)
```

```
## each individual has
```

```
IDINA=wDINA+cDINA,
```

```
## wDINA= individual's wage each year*
```

```
(Average National Wage Income/  
average wage income regarding
```

```
to micro-data)
```

```
## cDINA=individual's capital income each year*
```

```
(Average National Capital Income/  
average capital income in micro-data)
```

```
## DINA=SUM(IDAINA)
```

```
for (i in c(2000:2018)) {
```

```
di<- databind1[which
```

```
(databind1$inc_tax_yr_year_nbr==i),]
```

```

IRD$numberofmicroindividuals [ IRD$inc_tax_yr_year_nbr==i] <- length (unique ( di$
diincome <- di
[ which ( di$inc_tax_yr_income_source_code==
"W&S" | di$inc_tax_yr_income_source_code==
"C02" | di$inc_tax_yr_income_source_code==
"P02" | di$inc_tax_yr_income_source_code==
"TPR" | di$inc_tax_yr_income_source_code==
"IRD" | di$inc_tax_yr_income_source_code==
"SNZ" | di$inc_tax_yr_income_source_code==
"C01" | di$inc_tax_yr_income_source_code==
"P01" | di$inc_tax_yr_income_source_code==
"S01" | di$inc_tax_yr_income_source_code=="S02" |
di$inc_tax_yr_income_source_code==
"C00" | di$inc_tax_yr_income_source_code==
"S00" | di$inc_tax_yr_income_source_code==
"S03" | di$inc_tax_yr_income_source_code=="P00" ) , ]

totalIRincome <- sum
(diincome$inc_tax_yr_tot_yr_amt)

## Transfers for National Income version
ditr <- di [ which ( di$inc_tax_yr_income_source_code==
"CLM" | di$inc_tax_yr_income_source_code==
"PEN" | di$inc_tax_yr_income_source_code==
"BEN" | di$inc_tax_yr_income_source_code==
"PPL" | di$inc_tax_yr_income_source_code=="STU" ) , ]

```

```
ditr$tax<- 0
pentotal<- sum(ditr$inc_tax_yr_tot_yr_amt)

sub<- aggregate
(inc_tax_yr_tot_yr_amt~snz_ird_uid ,
 data=diincome , FUN=sum)

sub1<- sub[which
(sub$inc_tax_yr_tot_yr_amt <=14000),]
sub1$tax<- 0.105
sub11<- sum(sub1$inc_tax_yr_tot_yr_amt)
taxsub11<- 0.105*sub11

sub2<- sub[which(sub$inc_tax_yr_tot_yr_amt >
14000 & sub$inc_tax_yr_tot_yr_amt <= 48000),]
sub2$tax<- 0.175

sub21<- sum(sub2$inc_tax_yr_tot_yr_amt)
taxsub21<- 0.175*sub21

sub3<- sub
[which(sub$inc_tax_yr_tot_yr_amt >
48000 & sub$inc_tax_yr_tot_yr_amt <= 70000),]
sub3$tax<- 0.3

sub31<- sum(sub3$inc_tax_yr_tot_yr_amt)
taxsub31<- 0.3*sub31

sub4<- sub[which(sub$
```

```

inc_tax_yr_tot_yr_amt > 70000),]
sub4$tax <- 0.33

sub41 <- sum(sub4$inc_tax_yr_tot_yr_amt)
taxsub41 <- 0.33 * sub41

totaltax <- taxsub11 + taxsub21 +
taxsub31 + taxsub41
taxshare <- (pentotal / totaltax)

sub5 <- rbind(sub1, sub2, sub3, sub4)
diincome$tax <- sub5$tax[match
(diincome$snz_ird_uid,
sub5$snz_ird_uid)]

## wage source (for National Income version)
diwage <- diincome[which
(diincome$inc_tax_yr_income_source_code
== "W&S" | diincome$
inc_tax_yr_income_source_code == "C02" |
diincome$
inc_tax_yr_income_source_code == "P02"
| diincome$
inc_tax_yr_income_source_code == "TPR"
| diincome$
inc_tax_yr_income_source_code == "IRD"
| diincome$
inc_tax_yr_income_source_code == "SNZ" |
diincome$
inc_tax_yr_income_source_code == "C01"

```

```
| diincome$  
  
inc_tax_yr_income_source_code=="P01"  
| diincome$  
inc_tax_yr_income_source_code=="S01"  
| diincome$  
inc_tax_yr_income_source_code=="S02" ), ]  
diwage1<-  
diwage[ which( diwage$tax==0.105) , ]  
diwage1$  
inc_tax_yr_tot_yr_amt<-  
  
( diwage1$inc_tax_yr_tot_yr_amt )  
- ( 0.105* taxshare* diwage1  
$inc_tax_yr_tot_yr_amt )  
  
diwage2<- diwage  
[ which( diwage$tax==0.175) , ]  
diwage2$inc_tax_yr_tot_yr_amt<-  
( diwage2$inc_tax_yr_tot_yr_amt ) -  
( 0.175* taxshare* diwage2  
$inc_tax_yr_tot_yr_amt )  
  
diwage3<- diwage  
[ which( diwage$tax==0.3) , ]  
diwage3$inc_tax_yr_tot_yr_amt<-  
( diwage3$inc_tax_yr_tot_yr_amt ) -  
( 0.3* taxshare*  
diwage3$inc_tax_yr_tot_yr_amt )
```

```

diwage4<- diwage[ which
(diwage$tax==0.33),]
diwage4$inc_tax_yr_tot_yr_amt<-
(diwage4$inc_tax_yr_tot_yr_amt) -
(0.33*taxshare*
diwage4$inc_tax_yr_tot_yr_amt)

diwageTOTAL<- rbind(diwage1 ,
diwage2 , diwage3 , diwage4 , ditr)
IRD$totalmicrowage
[ IRD$inc_tax_yr_year_nbr==i]<-
sum( diwageTOTAL$inc_tax_yr_tot_yr_amt)

IRD$averagemicrowage
[ IRD$inc_tax_yr_year_nbr==i]<-
(IRD$totalmicrowage
[ IRD$inc_tax_yr_year_nbr==i ])/
(IRD$numberofmicroindividuals
[ IRD$inc_tax_yr_year_nbr==i ] )

# average National Wage
IRD$averagenationalwage
[ IRD$inc_tax_yr_year_nbr==i]<-
(IRD$‘ Compensation Of Employees – received ‘
[ IRD$inc_tax_yr_year_nbr==i ])/
(IRD$POPULATION
[ IRD$inc_tax_yr_year_nbr==i ] )

IRD$WDINA[ IRD$inc_tax_yr_year_nbr==i]<-

```

```

(IRD$averagenationalwage
 [IRD$inc_tax_yr_year_nbr==i])
/(IRD$averagemicrowage
 [IRD$inc_tax_yr_year_nbr==i])

sample1<- aggregate(inc_tax_yr_tot_yr_amt ~
snz_ird_uid ,
data=diwageTOTAL, FUN=sum)

sample1$wDINA<- (sample1$inc_tax_yr_tot_yr_amt)*WDINA

## capital source
dicapital<- diincome[which
(diincome$inc_tax_yr_income_source_code
=="C00"
|diincome$
inc_tax_yr_income_source_code==
"S00" |
diincome$inc_tax_yr_income_source_code==
"S03" |
diincome$inc_tax_yr_income_source_code==
"P00"),]

dicapital1<- dicapital
[which(dicapital$tax==0.105),]
dicapital1$inc_tax_yr_tot_yr_amt<-
(dicapital1$inc_tax_yr_tot_yr_amt) -
(0.105*taxshare*
dicapital1$inc_tax_yr_tot_yr_amt)

```

```
dicapital2 <- dicapital
[which(dicapital$tax == 0.175),]
dicapital2$inc_tax_yr_tot_yr_amt <-
(dicapital2$inc_tax_yr_tot_yr_amt) -
(0.175 * taxshare *
dicapital2$inc_tax_yr_tot_yr_amt)
```

```
dicapital3 <- dicapital
[which(dicapital$tax == 0.3),]
dicapital3$inc_tax_yr_tot_yr_amt <-
(dicapital3
```

```
$inc_tax_yr_tot_yr_amt) -
(0.3 * taxshare *
dicapital3$inc_tax_yr_tot_yr_amt)
```

```
dicapital4 <-
dicapital [which
(dicapital$tax == 0.33),]
dicapital4$inc_tax_yr_tot_yr_amt <-
(dicapital4$inc_tax_yr_tot_yr_amt) -
(0.33 * taxshare *
dicapital4$inc_tax_yr_tot_yr_amt)
```

```
dicapitalTOTAL <-
rbind(dicapital1, dicapital2,
dicapital3, dicapital4)
```

```
IRD$totalmicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i]<-
sum(dicapitalTOTAL$
inc_tax_yr_tot_yr_amt)
IRD$averagemicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i]
<-
(IRD$totalmicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$numberofmicroindividuals
[IRD$inc_tax_yr_year_nbr==i])

# average National Capital income
IRD$averagenationalcapitalincome
[IRD$inc_tax_yr_year_nbr==i]<-
(IRD$CAPITALIncome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$POPULATION
[IRD$inc_tax_yr_year_nbr==i])
IRD$averagenationalcapitalincome
[IRD$inc_tax_yr_year_nbr==i]

# DINA scale for capitalincome: CDINA=
(Average National Capital Income/
average capital income regarding
to the micro-data)
IRD$CDINA
[IRD$inc_tax_yr_year_nbr==i]<-
(IRD$averagenationalcapitalincome
```

```

[IRD$inc_tax_yr_year_nbr==i])/
  (IRD$averagemicrocapitalincome
  [IRD$inc_tax_yr_year_nbr==i])

# individual's weighted capital
income using CDINA as a scale
sample2<- aggregate
(inc_tax_yr_tot_yr_amt~snz_ird_uid,
data=dicapitalTOTAL,
FUN=sum)
sample2$cDINA<-
(sample2$inc_tax_yr_tot_yr_amt)*
(IRD$
averagenationalcapitalincome
[IRD$inc_tax_yr_year_nbr==i])/
(IRD$averagemicrocapitalincome
[IRD$inc_tax_yr_year_nbr==i])

# IDINA=wDINA+cDINA
sample3<- merge(
sample1, sample2,
by="snz_ird_uid", all=TRUE)

sample3[is.na(sample3)]<- 0

sample3$IDINA<- (sample3$wDINA)+
(sample3$cDINA)

```

```
di$IDINAindividuals<-
sample3$IDINA
[match(di$snz_ird_uid ,
sample3$snz_ird_uid)]

# IDINA percentiles

sample4<- aggregate
(IDINAindividuals~
snz_ird_uid ,
data=di , FUN=mean)

IRD$bottom50thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
quantile
(sample4$IDINAindividuals ,
probs = 0.50)

IRD$bottom90thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
quantile(sample4$IDINAindividuals ,
probs = 0.90)

IRD$
bottom99thpercentile
[IRD$inc_tax_yr_year_nbr==i]<-
quantile
(sample4$IDINAindividuals ,
probs = 0.99)

di1<- sample3[
which(sample3$IDINA<=
```

IRD\$bottom50thpercentile
[IRD\$inc_tax_yr_year_nbr==i] ,]

IRD\$DINAbottom50percentiles
[IRD\$inc_tax_yr_year_nbr==i]<-
(sum(di1\$IDINA)/sum
(sample3\$IDINA))*100

di2<- sample3 [which (sample3\$IDINA<=
IRD\$bottom90thpercentile
[IRD\$inc_tax_yr_year_nbr==i]) ,]
IRD\$DINAbottom90thpercentile
[IRD\$inc_tax_yr_year_nbr==i]<-
(sum(di2\$IDINA)/sum(sample3\$IDINA))*100

di3<- sample3 [which (sample3\$IDINA <=
IRD\$bottom99thpercentile
[IRD\$inc_tax_yr_year_nbr==i]) ,]
IRD\$DINAbottom99thpercentile
[IRD\$inc_tax_yr_year_nbr==i]<-
(sum(di3\$IDINA)/sum(sample3\$IDINA))*100

IRD\$DINAmiddleclass
[IRD\$inc_tax_yr_year_nbr==i]<-
(IRD\$DINAbottom90thpercentile
[IRD\$inc_tax_yr_year_nbr==i])
-(IRD\$DINAbottom50percentiles
[IRD\$inc_tax_yr_year_nbr==i])

IRD\$DINAtop1thpercentile

```
[IRD$inc_tax_yr_year_nbr==i]<-  
100-(IRD$DINAbottom99thpercentile  
[IRD$inc_tax_yr_year_nbr==i])
```

```
IRD$DINAtop10percentiles  
[IRD$inc_tax_yr_year_nbr==i]<-  
100-(IRD$DINAbottom90thpercentile  
[IRD$inc_tax_yr_year_nbr==i])
```

```
# othertop9percentiles=  
percentiles90th to 99th  
IRD$DINAtop9otherpercentiles  
[IRD$inc_tax_yr_year_nbr==i]<-  
(IRD$DINAbottom99thpercentile  
[IRD$inc_tax_yr_year_nbr==i]) -  
(IRD$DINAbottom90thpercentile  
[IRD$inc_tax_yr_year_nbr==i])
```

```
}
```

```
## visualizing the results:
```

```
DINAbottom50percentiles<-  
c(IRD$bottom50thpercentile  
[IRD$inc_tax_yr_year_nbr  
==2000:2018])  
year<- c(2000:2018)
```

```
DINAtop1thpercentile<-  
c(IRD$DINAtop1thpercentile  
[IRD$inc_tax_yr_year_nbr  
==2000:2018])  
year<- c(2000:2018)
```

```
DINAmiddleclass<-  
c(IRD$DINAmiddleclass  
[IRD$inc_tax_yr_year_nbr  
==2000:2018])  
year<- c(2000:2018)
```

```
DINAtop10percentiles<-  
c(IRD$DINAtop10percentiles  
[IRD$  
inc_tax_yr_year_nbr==2000:2018])  
year<- c(2000:2018)
```

```
DINAtop9otherpercentiles<-  
c(IRD$DINAtop9otherpercentiles  
[IRD$  
inc_tax_yr_year_nbr==2000:2018])  
year<- c(2000:2018)
```

```
graph<- data.frame(supp=rep
(c("DINAbottom50percentiles",
"DINAtop1thpercentile",
"DINAmiddleclass",
"DINAtop10percentiles",
"DINAtop9otherpercentiles"),
each=19),
year=rep(c("2000", "2001", "2002",
"2003", "2004", "2005", "2006",
"2007", "2008", "2009", "2010",
"2011", "2012", "2013", "2014",
"2015", "2016", "2017", "2018"), 5),
DINApretaxIRnationalIncome=
c(IRD$DINAbottom50percentiles
[IRD$inc_tax_yr_year_nbr == 2000:2018],
IRD$DINAtop1thpercentile
[IRD$inc_tax_yr_year_nbr == 2000:2018],
IRD$DINAmiddleclass
[IRD$inc_tax_yr_year_nbr == 2000:2018],
IRD$DINAtop10percentiles
[IRD$inc_tax_yr_year_nbr == 2000:2018],
IRD$DINAtop9otherpercentiles
[IRD$inc_tax_yr_year_nbr == 2000:2018]))

ggplotDINA<- ggplot(data=graph,
aes(x=year,
y=DINApretaxIRnationalIncome,
group=supp, colour=supp))+
  geom_line()+
```

`geom_point()`

`ggplotDINA`

Bibliography

- Alvaredo, F., A. B. Atkinson, T. Piketty, and E. Saez, The top 1 percent in international and historical perspective, *Journal of Economic perspectives*, 27(3), 3–20, 2013.
- Alvaredo, F., T. Atkinson, T. Piketty, and E. Saez, The world top incomes database; 2013, URL <http://topincomes.parisschoolofeconomics.eu>, 2014.
- Alvaredo, F., A. Atkinson, L. Chancel, T. Piketty, E. Saez, G. Zucman, et al., Distributional national accounts (dina) guidelines: Methods and concepts used in wid. world, *WID. world Working Paper*, 2, 2016.
- Alvaredo, F., L. Chancel, T. Piketty, E. Saez, and G. Zucman, Global inequality dynamics: New findings from wid. world, *American Economic Review*, 107(5), 404–09, 2017.
- Atkinson, A. B., and A. Leigh, The distribution of top incomes in Australia, *Economic Record*, 83(262), 247–261, 2007.
- Atkinson, A. B., and A. Leigh, Top incomes in New Zealand 1921–2005: Understanding the effects of marginal tax rates, migration threat, and the macroeconomy, *Review of Income and Wealth*, 54(2), 149–165, 2008.
- Bakker, A., and J. Creedy, Macroeconomic variables and income inequality in New Zealand: An exploration using conditional mixture distributions, *New Zealand Economic Papers*, 33(2), 59–79, doi:10.1080/00779959909544308, 1999.
- Ball, C., and J. Creedy, Inequality in New Zealand 1983/84 to 2012/13, *New Zealand Economic Papers*, 50(3), 323–342, doi:10.1080/00779954.2015.1128963, 2016.
- Ball, C., and J. Ormsby, Comparing the household economic survey to administrative

- records: An analysis of income and benefit receipt, *New Zealand Treasury Analytical Papers (1)*, pp. 1–35, 2017.
- Bozio, A., B. Garbinti, J. Goupille-Lebret, G. Malka, and T. Piketty, Inequality and redistribution in France, 1990-2018: evidence from post-tax distributional national accounts (dina), *World Inequality Lab Working Papers*, 2018.
- Creedy, J., N. Gemmell, and L. Nguyen, Income inequality in New Zealand, 1935–2014., *Australian Economic Review*, 51, 21-40.doi:10.1111/1467-8462.12240.
- Database, W. I., world inequality database (accessed on 28.11.2020), <https://wid.world/>, 2020.
- Dixon, S., The distribution of earnings in New Zealand 1984-1995, *Labour, Employment and Work in New Zealand*.
- Easton, Economic inequality in New Zealand: A user's guide, *New Zealand Sociology*, 28(3), 19, 2013.
- Easton, B., Economic inequality in New Zealand: Update to a user's guide, *New Zealand Sociology*, 29(3), 7, 2014.
- Evans, L., A. Grimes, B. Wilkinson, and D. Teece, Economic reform in New Zealand 1984-95: The pursuit of efficiency, *Journal of Economic Literature*, 34(4), 1856–1902, 1996.
- Fajnzylber, P., D. Lederman, and N. Loayza, Inequality and violent crime, *The Journal of Law and Economics*, 45(1), 1–39, 2002.
- Garbinti, B., J. Goupille-Lebret, and T. Piketty, Income inequality in France, 1900–2014: evidence from distributional national accounts (dina), *Journal of Public Economics*, 162, 63–77, 2018.
- Hsieh, C.-C., and M. D. Pugh, Poverty, income inequality, and violent crime: a meta-analysis of recent aggregate data studies, *Criminal Justice Review*, 18(2), 182–202, 1993.

- Info-Share, Statistics New Zealand Info-share (accessed on 26.11.2020), <http://infoshare.stats.govt.nz/infoshare/>, 2020.
- Kergozou, N., Simplified distribution national accounts: method and estimates for new zealand between 2000 to 2016, 2017.
- NZ, S., Measuring the distribution of household income and outlays within a national accounts framework, *Retrieved from www.stats.govt.nz*, 97 (ISBN 978-1-98-852879-3), 2018.
- NZ, S., Statistics New Zealand (accessed on 26.11.2020), https://www.stats.govt.nz/topics/government-finance?gclid=Cj0KCQiAqo3-BRDoARIsAE5vnaJh22_Bq-0v459qEGziH1gvc7LNA0ZQAqHimWJdEdmzr0mzR6x8VQ8aAufbEALw_wcB//, 2020.
- Piketty, T., E. Saez, and G. Zucman, Distributional national accounts: methods and estimates for the United States, *The Quarterly Journal of Economics*, 133(2), 553–609, 2018.
- Podder, N., and S. Chatterjee, Sharing the national cake in post reform New Zealand: income inequality trends in terms of income sources, *Journal of Public Economics*, 86(1), 1–27, 2002.
- Revenue, I., Tax database (accessed on 28.11.2020), <https://www.ird.govt.nz/>, 2020.
- Winkelmann, L., and R. Winkelmann, Immigrants in New Zealand: A study of their labour market outcomes, New Zealand Department of Labour *Occasional Paper*, 1, 1998.
- Zealand, S. N., A. Evans, R. Crane, L. Cooper, J. Mardula, J. Wilks, C. Surawy, M. Kenny, W. Kuyken, T. Goode, et al., The 2013 census ethnic group profiles, *Wellington: Statistics New Zealand*, 2013.
- Zucman, G., The missing wealth of nations: Are Europe and the US net debtors or net creditors?, *The Quarterly journal of economics*, 128(3), 1321–1364, 2013.