PRI Discussion Paper Series (No.23A-04)

Top Income Shares in Japan from the Survey and Tax Data in 2014 and 2019: Following the Distributional National Accounts Guidelines

Visiting Scholar, Policy Research Institute, Ministry of Finance
Masahiro Mikayama
Visiting Scholar, Policy Research Institute, Ministry of Finance
Tomotsugu Imahori
Chief Economist, Policy Research Institute, Ministry of Finance
Taro Ohno
Senior Economist, Policy Research Institute, Ministry of Finance
Yasutaka Yoneta
Director-General, Policy Research Institute, Ministry of Finance
Junji Ueda

October 2023

Research Department Policy Research Institute, MOF
3-1-1 Kasumigaseki, Chiyoda-ku, Tokyo 100-8940, Japan
TEL 03-3581-4111

Top Income Shares in Japan from the Survey and Tax Data in 2014 and 2019: Following the Distributional National Accounts Guidelines[*1]

Masahiro Mikayama[*2]
Tomotsugu Imahori[*3]
Taro Ohno[*4]
Yasutaka Yoneta[*5]
Junji Ueda[*6]

Abstract

This paper shows the results of estimating top income shares in Japan using the 2014 and 2019 household survey dataset and the newly aggregated tax dataset from income tax microdata following the Distributional National Accounts (DINA) Guidelines. The guidelines are presenting the concepts, data sources, and methods to keep consistency for international comparison and used for making data series in the World Inequality Database (WID). By following the DINA procedure, we combine the survey dataset and the newly aggregated income tax dataset for 2014 and 2019 using the method presented by Blanchet et al. (2022). In addition, we consider various types of incomes which are not included in the survey and tax datasets but are included in the Net National Income (NNI) concept. The main results by using the most recent Japanese survey and tax data show that the top 1% income share in Japan was 8.44%, and the top 10% income share was 33.79% in 2019, which are much lower than the results for the United States and France presented in the World Inequality Report 2022 (Chancel et al., 2021).

Keywords: Top income shares, Inequality, Distributional National Accounts (DINA), Survey data, Tax data, Net National Income (NNI)
JEL Classification: D3, E01, H2, J3

*1 The contents of this paper represent the authors' personal views and not the official views of the organizations to which they belong. The authors express deep gratitude to Chiaki Moriguchi (Hitotsubashi University), Ryo Ishida (Policy Research Institute, Ministry of Finance), Kentaro Asai (World Inequality Lab), Eisuke Sato (Policy Research Institute, Ministry of Finance), and Sehyun Hong (World Inequality Lab) for providing various helpful insights and comments.
*2 Visiting Scholar, Policy Research Institute, Ministry of Finance
*3 Visiting Scholar, Policy Research Institute, Ministry of Finance
*4 Chief Economist, Policy Research Institute, Ministry of Finance
*5 Senior Economist, Policy Research Institute, Ministry of Finance
*6 Director-General, Policy Research Institute, Ministry of Finance

I. Introduction

Since Thomas Piketty published a book in 2013 (Piketty, 2013) focusing on time-series income and wealth inequality data in European countries and the United States, it is attracting more efforts to measure income and wealth inequality based on more reliable data in an internationally comparable way. One of the most systematic efforts to measure inequality in a comparable way across different countries is the Distributional National Accounts (DINA) project implemented by researchers at the World Inequality Lab (WIL). The first version of the guidelines was released in 2016 (Alvaredo et al., 2016), and the most recent version was released in 2021 (Alvaredo et al., 2021). Piketty et al. (2018) published income distribution in the United States, Garbinti et al. (2018) published income distribution in France, and now the WIL publishes income and asset inequality for major countries in the World Inequality Report (WIR), to the extent possible, in accordance with the methodology in the DINA Guidelines. In the latest WIR, World Inequality Report 2022 (Chancel et al., 2021), figures are given for the share of income of the top 1% and the share of the top 10% of income groups in major countries.

The DINA Guidelines set up the concepts and the methods to make internationally harmonized time-series data for income and wealth inequality based on the System of National Accounts (SNA) and provide procedures to use both statistical survey data and tax microdata. This paper shows the results of estimating top income shares in Japan using the 2014 and 2019 household survey data and the newly aggregated tax data from income tax microdata following the DINA Guidelines.

From the perspective of research for Japan's income distribution, there have been extensive works using statistical household survey datasets (Tachibanaki, 1998; Ohtake, 2005; Ohtake and Kohara, 2010)[1)]. However, it has been widely recognized that the statistical survey data lack the sufficient distributional information about the high-income groups, and it is important to utilize both statistical survey datasets and tax datasets. In this context, the most important work was Moriguchi and Saez (2008) which estimated the shares of the high-income groups from 1876 to 2005 by combining semi-aggregated tax datasets for self-assessed income tax and withholding tax for wage income. After their seminal works for estimating top income shares in Japan, there have not been further updates in line with the DINA Guidelines. This paper can shed light on the recent situation of top income shares in Japan using the 2014 and 2019 data, and provide insights about how the top income share results from datasets and methods proposed by the DINA Guidelines could be different from previous seminal works for Japan's income distribution.

[1)] Tachibanaki (1998), Ohtake (2005), and Ohtake and Kohara (2010) are representative papers measuring the Gini coefficient. Ohtake and Kohara (2010) measured the Gini coefficient from 1980 to 2004 using the Household Survey, the Comprehensive Survey of Living Conditions, the National Survey of Family Income and Expenditure, and the Survey on the Redistribution of Income. The results show that the Gini coefficient has increased in all statistics, but the slope varies.

The major contributions of this paper are the following. First, it utilizes the semi-aggregated income tax dataset which was newly created by Kunieda and Yoneta (2023) under the Joint Research Program (JRP) adopted by the National Tax Agency (NTA). The advantage of using the new dataset is that the concept of income for the tax microdata can be appropriately aligned with the concept of the SNA. The most outstanding advantage is that we can accurately remove capital gains and losses from income in the tax microdata. In addition, the tax dataset can provide more information on the distribution of capital income, so we can more accurately reflect the distribution of capital income in our results.

Second, this is the first paper which utilizes the method proposed by Blanchet et al. (2022) to combine the Japanese statistical survey dataset and the tax dataset to estimate the whole income distribution through replacing and reweighting. This approach enables us to fully utilize the information in both statistical survey dataset and tax dataset in a consistent manner to estimate the whole income distribution.

Third, this paper makes extensive efforts to cover different types of incomes which are not covered in the statistical survey dataset and the tax dataset to estimate the whole distribution of the annual Net National Income (NNI) following the concept of the SNA. For this purpose, this paper clarifies the other components of the NNI which are not covered in the household statistical survey dataset and the tax dataset, such as imputed rents, capital incomes, retained earnings by corporations, etc., and then considers the distributions of individual components following the DINA Guidelines.

This paper is structured as follows. Section II describes the definition of income and the characteristics of the data, and Section III explains the method of combining tax and survey data. Section IV explains the method for distributing the income components that are not included in the survey and tax datasets and shows the results of top income shares following the DINA Guidelines. We also compare the results with the estimated results without using the new income tax dataset. Section V provides a summary and the remaining research agenda.

II. Definitions and data

In this section, we introduce the definitions of several concepts that need to be considered as a basis for estimating income distribution, such as population, unit of observation, and types of incomes, which are adopted in this paper referring to the DINA Guidelines. Then we describe the characteristics of the data used in this paper in detail. In particular, the semi-aggregated tax dataset derived from the income tax return data under the JRP adopted by the NTA in 2022 is the key input to the novelty of this paper.

II-1. Definitions

II-1-1. Distributional National Accounts Guidelines (DINA Guidelines)

Since the beginning of the 21st century, there have been active efforts to use tax microdata to look at the historical series of the top income share in a country, and in 2011, a number of researchers created the World Top Income Database (WTID) to share the results. Subsequently, efforts have been made to develop guidelines on how to measure income distribution by combining tax and statistical data based on the SNA concept, in order to enable international comparisons based on a common concept regarding the distribution of income and assets. The first version of the DINA Guidelines was released in 2016 (Alvaredo et al., 2016), and the first edition of the WIR, World Inequality Report 2018, was published in 2017 (Alvaredo et al., 2017) by the WIL. The most recent version of the DINA Guidelines was released in 2021 (Alvaredo et al., 2021), setting standards and clarifying concepts and methods so that the estimation results are comparable over time and across countries.

II-1-2. Unit of observation, population, and income concept

As for the unit of observation for estimating income distribution, the DINA uses “equal-split adults” series showing the results of incomes distributed to adults and distributed equally within couples or households as its benchmark series. It is easy to estimate the series of the equal-split adults in countries where the income tax is levied on joint incomes of couples or overall household incomes, like the United States and France. However, the income tax in Japan is separately levied on individual incomes, and there is no reliable tax data on the overall household incomes, and it is inevitable to use the series of “individualistic adults” distributing income specifically to each individual. Therefore, it should be noted that the results in this paper cannot be perfectly comparable to the benchmark series of other countries following the “equal-split adults” concept.

Regarding the population, the DINA Guidelines recognize “resident and adult population aged 20 years or older” as a benchmark, and this paper also follows the same definition, using “Population Estimates” data provided by the Ministry of Internal Affairs and Communications (MIC).

According to the DINA Guidelines, it is aiming at showing the distribution of the entire NNI, which is equal to Gross Domestic Product (GDP) plus Net Foreign Income (NFI) corresponding to the income generated abroad but accruing to domestic residents minus Consumption of Fixed Capital (CFC). The concept of the aggregate income is fully consistent with the SNA, and therefore, including certain types of incomes, like the undistributed profits of corporations, while excluding capital gains and losses. In this paper, the NNI provided by the Cabinet Office (CAO) is used as the aggregate income in Japan.

In measuring income distribution across individuals, there are four different types of individual incomes in the DINA Guidelines: (1) pretax factor income, (2) pretax post-replacement income (which we will generally abbreviate as “pretax income”), (3) post-tax disposable income, and (4)

post-tax national income. Among them, the pretax income concept is most favored as the benchmark in the DINA Guidelines, and also adopted in this paper. This is because the distribution of pretax income is less sensitive to the age structure of the population in a country than pretax factor income[2),3)].

II-2. Data

The DINA Guidelines provide the method to estimate income distribution by combining statistical survey dataset and tax dataset. The former is expected to convey sufficient information about the income distribution of low- and middle-income groups, while the latter is expected to reflect the distribution of high-income groups accurately.

In this paper, we use the National Survey of Family Income, Consumption and Wealth (FICW, provided by the MIC) as the statistical survey dataset for Japan. It is the largest household survey statistics with 56,000 sample households in 2014 and 78,000 sample households in 2019, carried out once in five years, providing individual income data for the members of households.

Since the FICW is a household-based survey, we need to calculate individual incomes of all the members within households. For this purpose, the sum of major income components was calculated for each member of households as the total income[4)]. As a result, a dataset with approximately 164,000 individuals were obtained from the 2019 FICW survey[5)].

As for the tax dataset, we use the dataset which was newly created by the JRP carried out by the NTA in 2023 (“JRP tax data”) using administrative tax record data from 2014 to 2020. The dataset was provided in a tabulated form as the annex tables of Kunieda and Yoneta (2023) by aggregating individual self-assessed tax return data. The advantage of this dataset is that it provides various tabulations based on different concepts of incomes, including the one in line with the DINA’s income concept, such as excluding capital gains and losses and adding non-taxable incomes and deductions.

---

2) In the DINA Guidelines, it is explained that pretax factor income inequality can be artificially large in economies with large retired people, and pretax national income can satisfy the neutrality condition that, in a hypothetical steady-state economy with 100% replacement rates for pensioners, the cross-sectional inequality of pretax national income should be the same whether it is measured within the entire population (including pensioners) or within the working-age population.

3) Pretax, post-replacement income is classified into (a) pension-based pretax income, which considers replacement related to public pension, and (b) broad-based pretax income, which considers other replacement such as unemployment benefits. This paper uses the pension-based pretax income because the distribution of unemployment benefits is difficult to measure from available data in Japan. The Japanese tax law stipulates that unemployment benefits are non-taxable and therefore, it is difficult to capture the distribution of benefits from tax return data, and the statistical survey data do not have a sufficient number of household samples receiving unemployment benefits.

4) The sum of (1) annual income from employer, (2) income from agriculture, forestry, and fishing, (3) business income other than agriculture, forestry, and fishing, (4) annual income from side job, (5) annual income from rent and land rent, (6) public pension and benefits, (7) corporate pension benefits, and (8) private pension benefits was calculated from the individual data in the FICW.

5) In the FICW, weights are also given for the sample to be assigned back to the population. These are given to each household, and the individual data were given the same weights as households. The total number of weights is 96,551,000 in the 2019 FICW, which is slightly less than the adult population of Japan. Therefore, each weight is multiplied by a certain ratio so that the total of weights is equal to the adult population of Japan.

Therefore, it is not necessary to transform the income information in the tabulation to measure income distribution in line with the DINA. However, although the JRP dataset covers all the self-assessed tax returns (22.7 million taxpayers), it still misses a large part of high-income groups because workers with the earnings of salaries equal to or less than 20 million yen are not required to submit self-assessed tax return and can complete their tax reporting and payment process under the withholding mechanism[6]. So, we need to use an additional tax dataset, the Statistical Survey of Actual Status for Salary (ASS), covering workers subject to withholding taxation mechanism. The ASS is a sample survey conducted by the NTA on payments of salaries to employees at private business establishments.

Without the JRP dataset, previous studies on Japan's top income shares, including Moriguchi and Saez (2008), used the dataset of the Sample Survey for Self-Assessment Income Tax (SAI, provided by the NTA) to see the income distribution of high-income groups. The SAI is an annual sample survey conducted since 1951, and its sample size was 2.3 million in 2019, providing tabulations by income groups with overall taxable incomes. Therefore, it is necessary to make adjustments when trying to estimate the income distribution in line with the DINA Guidelines. We show the results of using the SAI dataset following the methods used in the previous studies, and make a comparison to the results of using the JRP data. Table 1 describes the characteristics of the data used in this paper.

[6] The main conditions for salary earners to be required to file a tax return are (1) earning salary income of more than 20 million yen or (2) earning total income of more than 200 thousand yen except for salary income and retirement income. Based on the above conditions, the term for complete tax filing is that the total income is more than 20.2 million yen, but in this paper, the term "over 20 million yen" is used for simplicity.

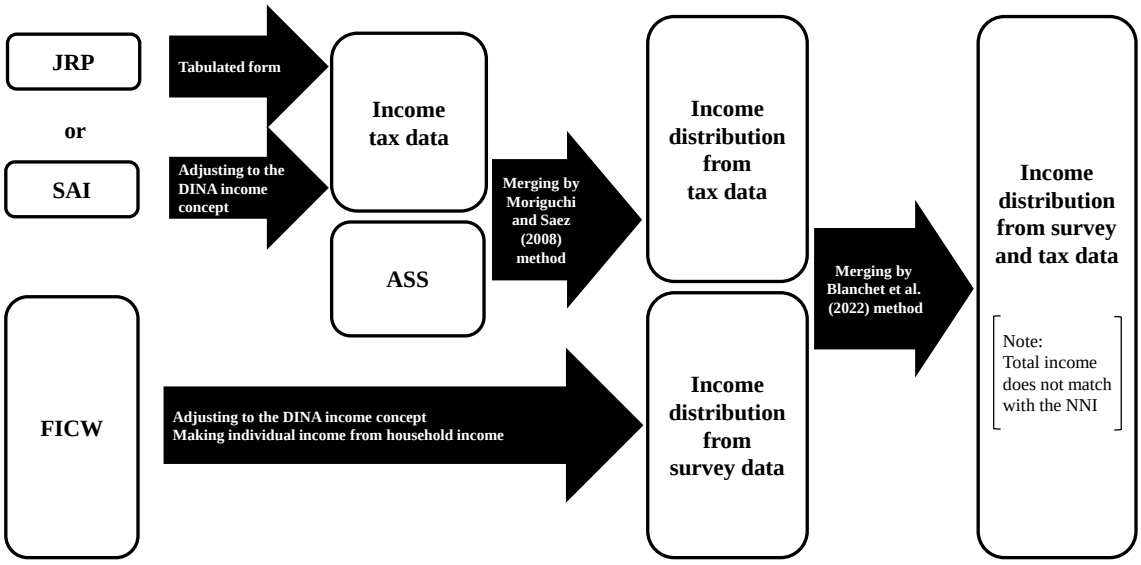
Table 1. Characteristics of survey and tax data

	Statistical survey data	Tax data		
Name	FICW (National Survey of Family Income, Consumption and Wealth)	JRP (Joint Research Program Income Tax Data)	SAI (Sample Survey for Self-Assessment Income Tax Data)	ASS (Statistical Survey of Actual Status for Salary in the Pricate Sector)
Advantage	Including a wide range of households, both working and non-working, as the population.	Covering the entire population of self-assessed income tax return data, including workers earning wage income with 20 million yen and more. Providing various tabulations, including the one in line with the DINA income concept.	Covering all the self-assessed income tax return data as the population, including workers earning wage income with 20 million yen and more.	Including all the employees receiving wage income as the population.
Disadvantage	Limited coverage for top income groups.	Incomes subject to withholding taxation (interest income, dividend income, retirement income, etc.) are not covered. Workers earning below 20 million yen and pensioners receiving below 4 million yen are not covered, unless they submit tax returns by special reasons.	Incomes subject to withholding taxation (interest income, dividend income, retirement income, etc.) are not covered. Workers earning below 20 million yen and pensioners receiving below 4 million yen are not covered, unless they submit tax returns by special reasons. Providing tabulations only for taxable incomes.	Providing tabulations for wage income only. Public employees are not covered.
Population	Approx. 50 million households	Approx. 22 million people	Approx. 6.3 million people	Approx. 60 million people
Sample size	56,352 households (2014) 89,471 households (2019)	Census	1,137,592 people (2014) 2,307,506 people (2019)	298,881 people (2014) 243,018 people (2019)
Available year	1954- (every 5 year)	2014-2020	1951-	1949-

III. Methods of estimating income distribution from survey and tax data

This section describes a method for estimating income distribution using both statistical survey data and tax data following the DINA Guidelines and several additional treatments which are required to use Japan's income data described in Section II. As described in Figure 1, the first step is to derive the income distribution from the tax datasets using the JRP dataset or the SAI dataset merged with the ASS dataset, while preparing the income distribution from survey dataset. This procedure requires to adjust the concepts of incomes in the datasets considering the components of the income datasets respectively following the income concept used in the DINA Guidelines, as explained in Section III-1. Then, merging the ASS dataset by using the method presented in Moriguchi and Saez (2008) is necessary to widen the "trustable span" as explained in Section III-2. Finally, it is necessary to merge the income distribution from the tax dataset and the income distribution from the survey dataset by using the method presented in Blanchet et al. (2022) as explained in Section III-3. The results of the estimated income distribution are presented in Section III-4, while the total income under the procedure in this section does not match with the NNI.

Figure 1. Flowchart of estimating income distribution using Japanese data



III-1. Making income data under the common definition

To measure income distribution by combining various income datasets presented in Section II, it is necessary to make the income data under the common definition with respect to the coverage of the income to be measured so that they can be merged. Since the datasets have their own purposes and data collection methods, the coverage of the overall incomes and the available data for income components in the datasets are profoundly different. In addition, the measurement of the available

income data can vary concerning whether they are evaluated after subtracting costs and other deductible incomes for tax purposes following the “income” concept, or before subtracting these items following the “money received” concept.

Table 2 shows the coverage and the available components data within the overall incomes in the datasets. The “C” mark indicates that the dataset provides the income components data which are sufficiently comprehensive, while the “NC” mark indicates that the data are not comprehensive in the dataset. The “#NA” mark indicates that the data are not provided in the dataset.

For example, the FICW dataset does not include data on capital gains and losses, as well as occasional incomes, because it aims to capture the overall incomes with permanent characteristics. As for the available data components, the FICW dataset and the ASS dataset provide salary data following the money received concept, while the SAI dataset provides salary data following the taxable income concept only. The advantage of the JRP dataset is that it provides salary data following both concepts.

There are several reasons why the income components data cannot be recognized as comprehensive. As for the JRP and SAI tax datasets, the full amount of income cannot be captured if a part of the income is subject to separate withholding taxation. Because the majority of interest and dividends receipts are subject to the withholding taxation mechanism, the data on the interest and dividends income in the tax datasets only reflect a small part of the total financial income. Regarding the FICW data, it is widely recognized that the responses to the survey questions about financial incomes and assets are significantly underreported. This is evident because the total sum of responses to the survey is significantly less than the total amount of financial incomes presented in the SNA data[7)].

[7)] In 2019, the total value of interest and dividends income derived from the FICW data was 1.4 trillion yen, which is 10.6% of the total financial income derived from the SNA data.

Table 2. Available income components data in the datasets

Name of income	Income / Money received	Survey	Tax data		
		FICW	JRP	SAI	ASS
Business income	Income	C	C	C	#NA
Agricultural income	Income	C	C	C	#NA
Forestry income	Income	C	C	C	#NA
Real estate income	Income	C	C	C	#NA
Wages and salaries income	Income	#NA	C	C	#NA
Wages and salaries	Money received	C	C	#NA	C
Miscellaneous income	Income	#NA	C	C	#NA
Pension income (public and private)	Income	#NA	C	C	#NA
Public pension	Money received	C	C	#NA	#NA
Private pension	Money received	C	C	#NA	#NA
Interest	Income*	NC	NC	NC	#NA
Dividends		NC	NC	NC	#NA
Capital gains	Income	#NA	NC	NC	#NA
Occasional incomes	Income	#NA	NC	NC	#NA
Retirement income	Income	#NA	NC	NC	#NA
Social benefits other than public pension	Money received	NC	#NA	#NA	#NA
Other income	Money received	NC	#NA	#NA	#NA

*Interest income is the same as the money received for receipt of interest, while dividends income is the money received for dividends minus interest payments for borrowing to buy the stocks.

Considering the availability of data and the comprehensiveness of each income component, we use the sum of business income, agricultural income, forestry income, real estate income, wages and salaries, and public and private pensions to estimate income distribution by combining the survey dataset and the JRP dataset. Alternatively, we also estimate income distribution by using the SAI dataset to make a comparison with the results using the JRP dataset. In this case, we need to convert the wages and salaries data and public pension data in the SAI dataset from the income concept to the money received concept to follow the common definition with the FICW dataset[8]. Also, we need to subtract other incomes, such as capital gains and occasional incomes by making assumptions about the share of these incomes so that the dataset can be merged to the FICW dataset.

The ASS dataset is used for supplementing income data of wage earners to the JRP and SAI datasets. In this paper, we assume that income other than wages and salaries for the samples included in the ASS dataset is negligible because if wage earners have different types of incomes with equal to or more than 0.2 million yen, they need to submit self-assessed tax returns and they are included in the sample of the JRP and SAI datasets.

III-2. Merging the JRP or SAI dataset with the ASS dataset

In general, survey data are expected to accurately represent the income distribution with respect to

[8] Conversion methods for employment income and pension income are explained in Appendix.

low- and middle-income individuals, while tax data are expected to accurately represent the income distribution with respect to high-income individuals. Therefore, in each country, the DINA Guidelines stipulate that the distribution of survey data should be used for individuals below a certain marginal point income, while the distribution of tax data should be used for individuals above the marginal point income[9]. To apply this method, we need to assure that the tax data can provide sufficiently reliable distribution data above the merging point. The DINA Guidelines provide the concept of the "trustable span" within which the distribution data can be considered reliable because the share of the population covered in the tax data is sufficiently high.

As for Japan's tax dataset, the lower end of the trustable span of the JRP and SAI datasets is recognized to be 20 million yen because only workers with wages and salaries more than 20 million yen are required to submit tax returns of their self-assessed incomes. To avoid the situation in which the merging point is outside the trustable span of the tax dataset, we combine the JRP dataset and the SAI dataset with the ASS dataset, which provides better data on the distribution of workers with the earnings of salaries equal to or less than 20 million yen, in order to extend the trustable span of the tax dataset. For this purpose, we refer to the method used in Moriguchi and Saez (2008).

The first step is to enlarge the population of the ASS dataset to include workers in the public sector, whose wages and salaries are also subject to withholding tax. Because the population of the ASS dataset is approximately 52.5 million with workers in the private sector, while the number of workers in the public sector is approximately 4.2 million, the ASS dataset is inflated with a multiple of 1.1 by assuming that the income distributions of public and private workers are similar.

The second step is to extract and remove duplicated samples included in both the JRP or SAI datasets and the ASS dataset. The duplicated samples are those who earn wages and salaries and submit tax returns for some reason, for example, with wages and salaries of more than 20 million yen. By estimating the number of duplicates and their total income for each income group using the JRP or SAI datasets, they can be removed from the ASS dataset.

The third step is combining the JRP or SAI datasets with the adjusted ASS dataset. Both data are tabulated data, and the income class ranges of these data do not necessarily match. Therefore, both data are interpolated to create a distribution and then combined.

After combining the datasets, the lower end of the trustable span of the combined data is recognized to be 4 million yen.

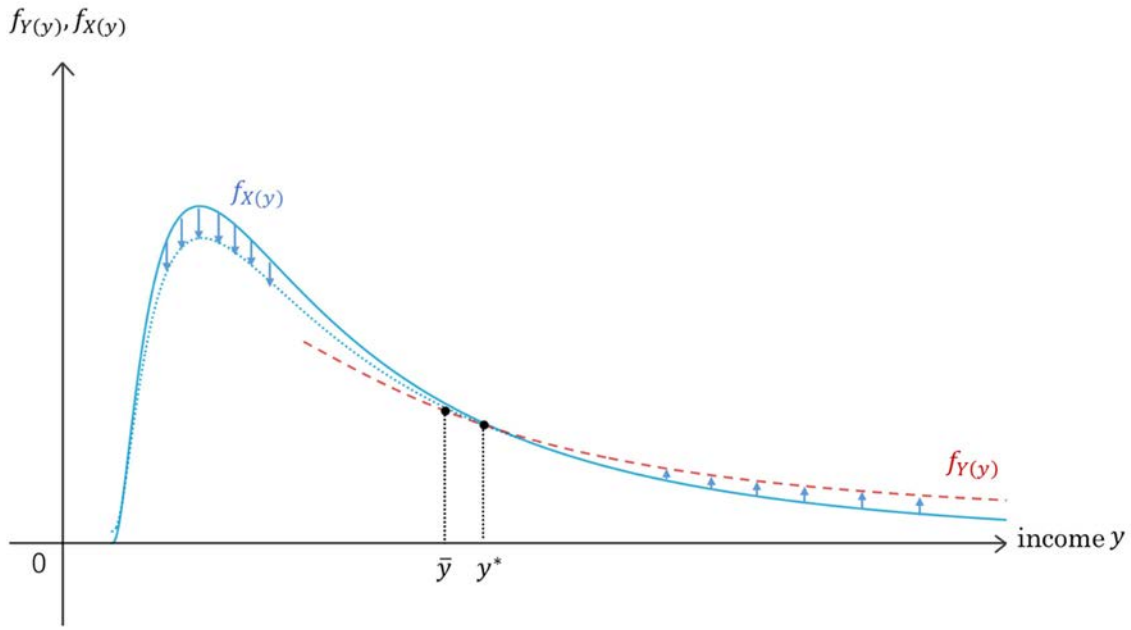
III-3. Merging tax data and survey data

The method of merging the FICW survey dataset and the tax dataset (combination of the JRP and the ASS or the SAI and the ASS) follows the one presented in Blanchet et al. (2022), using the

[9] The method of calculating the merging point income will be explained in Section III-3.

STATA's code "bfmcorr". This procedure tries to exploit the strengths of both the survey dataset and the tax dataset by setting an appropriate merging point and reweighting and replacing the samples. The method is graphically illustrated in Figure 2, where $f_X(y)$ is the probability density function of the survey dataset and $f_Y(y)$ is the probability density function of the tax dataset. The intersection of y^* is called the merging point and different correction procedures are applied below the merging point and above the merging point. The value of the density above the merging point should be increased to the distribution of the survey dataset by replacing it with the tax dataset, because people in high-income groups are underrepresented or unsampled in the survey dataset. Conversely, the value of the density below the merging point should be decreased by reweighting, because people in lower-income groups must be overrepresented in the survey dataset.

Figure 2. Image of the procedure combining survey and tax datasets



Source: Alvaredo et al. (2021, Figure 5.2)

Note: The solid blue line represents the density of the survey dataset $f_X(y)$. The dashed red line represents the density of the tax dataset $f_Y(y)$, which is only available for high-income groups.

III-4. Results of income distribution

Table 3 presents three results of estimating income distribution for 2014 and 2019. The two left-hand columns show the results using a combination of tax and survey data; the first column is the results using the JRP dataset and the second column is the results using the SAI dataset, while the rightmost column shows the results using only the FICW survey data.

Apparently, the results of top income shares using the tax datasets are higher than the ones using only the survey dataset, reflecting the richer information of high-income groups provided in the tax datasets. For example, the top 10% income share in 2019 using the JRP dataset was 33.73%, which is higher than the result using only the FICW (32.96%) by 0.77% point. The top 1% income share in 2019 using the JRP dataset was 8.06%, which is also higher than the result using only the FICW (6.83%) by 1.23% point.

The differences between the estimation results of top income shares using the JRP dataset and the SAI dataset do not seem to be significant. In 2019, the top 10% income share using the SAI dataset was 33.94%, which is a little higher than the result using the JRP dataset (33.73%) by 0.21% point, while the top 1% income share using the SAI dataset was 7.80%, which is slightly lower than the result using the JRP dataset (8.06%) by 0.26% point. Because the SAI dataset needs to be adjusted with several assumptions to make the dataset to be merged to the FICW dataset, it is not ideal to use the SAI dataset to precisely estimate the income distribution following the DINA Guidelines. The advantage of the SAI dataset is that it is available since 1951 and can be used for estimating income distribution for the period in which the JRP dataset is not available.

Table 3. Income distribution of survey and tax datasets before matching to the NNI

<2014>

Income group	Number of adults	JRP+ASS, FICW			SAI+ASS, FICW			FICW only		
		Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)
Full population	104,865,000	100.00%	-	2,757,886	100.00%	-	2,715,072	100.00%	-	2,640,310
Bottom 50%	52,432,500	12.49%	-	728,328	12.28%	-	704,779	12.52%	-	698,843
Bottom 20%(P0-P20)	20,973,000	0.34%	-	69,870	0.30%	-	62,079	0.29%	-	60,121
Next 30%(P20-P50)	31,459,500	12.15%	490,000	1,167,300	11.98%	470,000	1,133,246	12.22%	460,000	1,124,658
Middle 40%(P50-P90)	41,946,000	50.17%	1,350,000	3,559,255	50.22%	1,300,000	3,514,596	51.09%	1,300,000	3,475,575
Top 10%	10,486,500	35.17%	6,070,888	9,700,814	35.25%	6,300,000	9,569,451	34.11%	6,160,000	9,005,501
Top 1%	1,048,650	8.45%	13,688,577	23,303,869	7.87%	13,253,270	21,377,939	7.01%	12,000,000	18,496,977
Top 0.1%	104,865	2.35%	34,705,832	64,693,871	1.90%	30,476,434	51,661,914	1.59%	27,800,000	41,981,543

<2019>

Income group	Number of adults	JRP+ASS, FICW			SAI+ASS, FICW			FICW only		
		Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)
Full population	105,134,000	100.00%	-	3,096,502	100.00%	-	3,006,271	100.00%	-	2,862,276
Bottom 50%	52,567,000	14.40%	-	937,023	14.26%	-	900,671	14.55%	-	874,564
Bottom 20%(P0-P20)	21,026,800	1.06%	-	199,545	1.00%	-	185,013	0.98%	-	174,268
Next 30%(P20-P50)	31,540,200	13.34%	720,000	1,428,675	13.26%	710,000	1,377,777	13.56%	700,000	1,341,428
Middle 40%(P50-P90)	42,053,600	49.78%	1,700,000	3,959,302	49.68%	1,611,664	3,839,006	50.36%	1,560,000	3,704,069
Top 10%	10,513,400	33.73%	6,586,407	10,444,330	33.94%	6,508,804	10,204,215	32.96%	6,300,000	9,432,825
Top 1%	1,051,340	8.06%	14,846,112	24,949,369	7.80%	14,538,370	23,453,367	6.83%	13,320,000	19,558,000
Top 0.1%	105,134	2.18%	36,879,859	67,380,102	1.88%	33,674,809	56,410,535	1.43%	27,180,000	40,959,563

IV. Distributing the other NNI incomes

As explained in Section II-1-2, the DINA Guidelines aim at showing the distribution of the entire NNI for the purpose of comparing income distribution across different countries. The income data used in Section III, hereafter referred to simply as “merged incomes”, were not the same as the NNI mainly for two reasons.

First, because the merged incomes data were selectively picked up in consideration of the availability of distributional data and the comprehensiveness of each income component in the survey and tax datasets. Therefore, the merged incomes were much smaller than the overall household incomes. This was because some types of household incomes, including retirement income, interest, and dividends, were excluded from the survey and tax datasets. In addition, imputed rents of owner-occupied dwellings are not recorded in the survey and tax datasets because there are no cash transactions, while the NNI should include them.

Second, the NNI includes several types of incomes which are not directly distributed to households and thus, are not recorded in the survey and tax datasets and are not included in the merged incomes by any means. Examples of such types of incomes include the retained earnings of corporations, and the incomes of the General Government (GG) and the Non-Profit Organizations Serving Households (NPISH).

In addition, even for the income components included in the merged incomes, such as wages and salaries, there can be discrepancies between the total values of the incomes recorded in the survey and tax datasets and the values in the National Accounts due to measurement issues, such as differences in concepts and data coverage.

To make the comprehensive income distribution of the NNI following the DINA Guidelines, it is necessary to consider the distribution of these types of incomes. In this section, we classify the incomes which are not included in the merged incomes and then explain how the incomes are distributed to individuals to make the final income distribution.

IV-1. Distributing other incomes of households

First, we clarify the income components of households which are not included in the merged incomes in Section III, following the income concept of the National Accounts. Table 4 displays the income components of households for the allocation of primary incomes and the secondary distribution of incomes for households. The second column shows the incomes which are already included in the merged incomes in Section III, and the third column provides the category of other incomes which need to be considered in this section.

Table 4. Income components of households

	Income components	Merged incomes	Other incomes of households
Allocation of Primary Income	**Resources**		
	Operating surplus, net	—	**Imputed rents of owner-occupied dwellings**
	Mixed income, net	**Business income, Agricultural income, Forestry income**	—
	Compensation of employees		
	(1) Wages and salaries	**Wages and salaries**	—
	(2) Employers' social contributions	—	*currnet costs incurred by employers (A)
	Property income		
	(1) Interest	—	**Interest**
	(2) Dividends	—	**Dividends**
	(3) Other investment income		
	a) Investment income attributable to insurance policyholders	—	**Investment income attributable to insurance policyholders**
	b) Investment income payable on pension entitlements	—	*interest costs incurred for retirement benefit obligations (B)
	c) Investment income attributable to collective investment fund share holders	—	**Investment income attributable to collective investment fund share holders**
	(4) Rents	**Real estate income**	—
	Uses		
	(1) Interest	—	**Interest (-)**
	(2) Rents	—	(omitted)
Secondary Distribution of Income	**Resources**		
	Social benefits other than social transfers in kind		
	(1) Social security benefits by cash	**Public pension**	—
	(2) Other social insurance pension benefits	—	**Retirement incomes** form accrual entities (including private pensions)
	(3) Other social insurance nonpension benefits	—	**Retirement incomes** from non-accrual entities
	(4) Social assistance benefits in cash	—	outside the scope of pretax income
	Other current transfers	—	outside the scope of pretax income
	Uses		
	Current taxes on income, wealth, etc.	—	outside the scope of pretax income
	Net social contributions		
	(1) Employers' social contributions	—	*the same amount as (A)
	(2) Households' actual social contributions	—	**Social contributions** (public pension only for the scope of pretax income)
	(3) Households' social contribution supplements	—	*the same amount as (B)
	Other current transfers	—	outside the scope of pretax income

The DINA Guidelines provide methods to be applied to each type of income for making an imputed distribution, either based on external sources or using the same distribution as other types of income in the datasets. If there are high-quality data, it is recommended to adopt detailed methods, while simplified imputations can be used in more problematic contexts[10)].

Table 5 shows the summary of the methods mentioned in the DINA Guidelines and the actual methods used in this paper for individual income types of households in consideration of data availability. The individual methods applied to each type of income are briefly described below.

[10)] The methods of imputing missing pretax income components recommended by the DINA Guidelines are provided in Table 5.4 of Alvaredo et al. (2021).

Table 5. Distribution methods for other income components of households

	Estimated total amounts in 2019 (trillion yen)	DINA allocation method	Allocation method in this paper
(a) Interest	+5.6 (received) -8.7 (paid)	Pretax income (simplified)	Pretax income
(b) Dividends	6.5	Pretax income (simplified)	Amounts shown on tax returns: JRP Amounts withheld: pretax income
(c) Investment income attributable to collective investment fund share holders	0.5	Capital income	Capital income
(d) Investment income attributable to insurance policyholders	8.3	Wages	Wages
(e) Imputed rents of owner-occupied dwellings	21.1	Consumption/Wages	Estimated from the survey dataset (FICW)
(f) Retirement income	8.9	—	Wages
(g) Social contributions for public pension	-20.2	Wages	Estimated from the survey dataset (FICW)

(a) Interest

Most interest income is subject to separate withholding taxation under the Japanese tax system, so there is no reliable information in the tax datasets. The FICW survey dataset does not provide reliable information either about the distribution of financial income, which is the sum of interest and dividends, due to significant underreporting.

Therefore, in accordance with the simplified recommendation of the DINA Guidelines, the distribution of interest is based on the proportion of pretax income. The total amount is controlled by the net interest (received interest minus paid interest) of households in the National Accounts data. Since the total share of net interest is quite small (0.4% of the total NNI), it is unlikely to have a significant impact on the results of income distribution in this paper.

(b) Dividends

Under the Japanese tax system, dividends income is subject to three types of taxation: (1) declared as a part of aggregate income subject to comprehensive taxation (for large amounts of dividends receipts or dividends receipts by majority shareholders), (2) separately declared and subject to different tax rates, and (3) subject to withholding taxation (optional). Therefore, it is not possible to know the full distribution of the overall dividends income from the JRP and SAI datasets, and the FICW cannot provide reliable distributional data due to significant underreporting.

So, we apply the distributional data from the tax datasets to the declared part (1) and (2), while the allocation of the withheld part (3) is based on the proportion of pretax income with the assumption that the total amount of dividends in the National Accounts minus the amount of declared dividends is subject to withholding tax. The annex tables provided in Kunieda and Yoneta (2023) include Case

2 (income data including dividends) and Case 3a (income data excluding dividends), and the difference between these two can be seen as the declared dividends in each bracket[11].

(c) Investment income attributable to collective investment fund shareholders

Investment income attributable to collective investment fund shareholders is the retained profits in collective investment vehicles. Due to the limitation of data on the distribution of such incomes, we adopt the proportion of capital income, which is the sum of interest and dividends, to allocate the income. The total amount is controlled by the amount in the National Accounts data.

(d) Investment income attributable to insurance policyholders

Investment income attributable to insurance policyholders is the income earned by life insurance companies. Because it is difficult to accurately estimate the distribution of insurance premiums paid by individual households, we adopt to allocate the income in proportion to wages and salaries by assuming that the distribution of life insurance policies depends on wages. The total amount is controlled by the National Accounts data.

(e) Imputed rents of owner-occupied dwellings

When households reside in their own dwellings, they do not actually pay rent, but are treated as producing and consuming services equivalent to the rent they would pay if they valued their own housings at market prices. This is called the imputed rents of owner-occupied dwellings. Since imputed rents are not subject to taxation in Japan, the tax dataset does not provide any information on imputed rents.

In this paper, we first estimate the imputed rents of owner-occupied dwellings by using the FICW survey dataset, and then use the value to allocate the total imputed rents in the National Accounts data[12]. The estimates are based on the estimation formula published by the MIC, which consolidates the FICW[13]. For regional classifications, we used the average value for each prefecture[14]. The imputed rents of owner-occupied dwellings, estimated at the household level, are allocated to individuals in proportion to the income of each household member. For the samples replaced in the procedure of merging the datasets, the imputed rents were estimated by using the relationship

---

[11] Since the JRP data are bracketed data, it is difficult to make a simple comparison between the two brackets because the number of people in each bracket differs between Case 2 and Case 3a due to the difference in whether dividends are included or not. Therefore, the interpolation method of Moriguchi and Saez (2008) is used to convert the data into data with thresholds for each percentile and average income in each bracket for comparison. In making the distribution, the assumption is made that the rank order of total income in the sample is not reversed before and after the distribution of dividends.

[12] Since imputed rents of owner-occupied houses are not subject to taxation in Japan, tax data do not provide sufficient information on imputed rents of owner-occupied houses.

[13] See Ministry of Internal Affairs and Communications (2014, 2019).

[14] We confirm that the estimates from the simplified model (household average) are almost the same as the estimates of imputed rents of owner-occupied houses published by the MIC.

between the original incomes and the estimated rents. Finally, the estimated imputed rents of owner-occupied dwellings for each sample are adjusted to the total amount in the National Accounts data.

(f) Retirement income

Retirement income refers to lump sum payments, such as retirement allowance or lump-sum benefit paid at the time of retirement. In Japan, large amounts of retirement benefits are generally paid at the time of retirement to employees of large corporations and public officials who have been continuously employed for a long period of time. Therefore, how the amount received at the time of retirement is treated is an important factor when considering income distribution.

The total amount of retirement income is estimated to be 8.9 trillion yen in 2019, which amounts to 1.6% of annual GDP, or 2.0% of the NNI[15]. However, there are no reliable data on the distribution of lump-sum cash transfers because the FICW does not survey retirement income. In addition, the tax datasets for self-assessment declaration rarely cover retirement income because it is subject to separate withholding taxation.

So, in this paper, we adopt an approach that retirement income is distributed to individual wage earners on an accrual basis by assuming that retirement income is a deferred payment of wages and salaries. The total amount of retirement income derived from the NTA's Annual Statistics Report (ASR) is distributed by using the proportion of wages and salaries in the FICW dataset. For the samples replaced by the procedure of merging the survey and tax datasets, the amount of wages and salaries is appropriately adjusted[16].

(g) Social contributions for public pension

In this paper, we estimate the income distribution for pretax income before the redistribution of tax systems. In the meantime, the benefits of public pension are to be included in the household's receipts, and the social contributions for public pension need to be subtracted. While the total amount of actual social contributions for public pension by households can be derived from the National Accounts data, which amounts to 20.2 trillion yen in 2019, but there are no clear data about the distribution of social contributions at the individual level.

We use a method proposed in Ohno et al. (2022) to estimate the social contributions for public pension at the individual level from the FICW dataset. The first step is to allocate household members

---

15) The amount of retirement income is reported in the ASR by the NTA; 1.5 trillion yen for public workers, and 7.4 trillion yen for others.

16) In the annex tables of Kunieda and Yoneta (2023), Case 3a represents the number of people and average amount in the total income bracket, and Case 3b represents the number of people and average amount of salary earners in the total income bracket. The total income for each bracket is obtained by multiplying the number of people and average income in these two brackets, and the total income and total wage are used to calculate the percentage of wage in each bracket. For the replaced sample, wage is estimated by multiplying the sample's total income by the percentage of wage in the relevant bracket.

in the FICW samples to the categories of public pension systems by considering information on income and relationship between household members. It is necessary to estimate the social contributions for pension because the contribution rates are different by the categories of public pension systems. Category I members (under the National Pension Insurance system) are required to pay a lump-sum contribution, and Category II members (under the Employees' Pension Insurance system) need to pay contribution with a certain proportion of wage income. Category III members (dependent spouses of Category II members) are not required to pay social contributions[17].

The second step is to calculate the average pension contribution ratio (pension contributions divided by total income) for each income bracket. For the samples replaced by merging procedure, the amount is appropriately adjusted. Then, the total amount of actual social contributions is distributed by referring to the estimated individual social contribution, and then subtracted from the other incomes.

IV-2. Distributing primary incomes of other institutional sectors

As shown in Table 6, the NNI in 2019 was 444.9 trillion yen, of which 343.5 trillion yen was allocated to households at the stage of primary income allocation. The share of allocation to the household sector was 77.2% of the total NNI, which means that the remaining 22.8% was not allocated to households. For example, 58.1 trillion yen was recorded as non-distributed corporate income, and 43.0 trillion yen was recorded as income of the GG. To make the distribution of the NNI as a whole, it is necessary to consider how these types of incomes should be imputed to households.

Table 6. Composition of the NNI of institutional sectors

	2014	2015	2016	2017	2018	2019	2020	2021
Gross Domestic Product (GDP)	518.8	538.0	544.4	553.1	556.6	557.9	539.1	549.4
Receipts of Foreign Income (net)	19.3	21.2	18.9	20.5	21.3	21.9	19.6	26.7
Gross National Income (GNI)	538.1	559.2	563.3	573.5	577.9	579.8	558.7	576.0
Consumption of Fixed Capital	-126.2	-128.1	-128.2	-130.1	-132.4	-134.5	-135.6	-138.7
Statistical Discrepancy	-0.2	-0.1	0.0	0.1	-0.1	-0.4	-0.9	1.5
Net National Income (NNI)	**411.7**	**431.0**	**435.1**	**443.6**	**445.5**	**444.9**	**422.1**	**438.8**
[Factor income composition of NNI]								
Compensation of employees	257.5	260.6	267.4	272.1	281.4	286.9	283.2	288.7
Operating surplus and mixed incomes	97.3	107.8	107.0	108.6	100.0	92.8	74.8	76.6
Net receipt of property income from abroad	19.2	21.1	18.8	20.4	21.2	21.7	19.5	26.6
Taxes on production - subsidies	37.7	41.6	41.9	42.5	43.0	43.4	44.7	46.9
[Sectoral composition of net primary income]								
Households	318.2	323.2	327.9	333.5	339.2	343.5	337.8	343.9
NPISH	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
Corporations	58.8	68.7	67.9	69.5	64.1	58.1	39.9	48.1
General Government	34.4	38.8	39.1	40.4	41.8	43.0	44.1	46.5

Source: National Accounts

[17] Under the pension system in Japan, anyone between the ages of 20 and 60 is covered by one of the plans.

The DINA Guidelines provide procedures for the imputation of these incomes to households. In general, we follow the methods proposed by them for the GG and the NPISH to allocate the incomes in proportion to the distribution of the pretax factor income, which is equal to the sum of all pretax income flows accruing directly or indirectly to the owners of the factors of production (labor and capital), before the operation of the tax and transfer system, including indirect taxes, and before the operation of the social insurance systems. When considering undistributed corporate profits, it is recommended to distribute them in proportion to direct or indirect stock ownership. Therefore, we carefully consider the stock ownership in Japan by using the Flow of Funds Accounts (FFA) data provided by the Bank of Japan (BOJ).

(a) Income of the General Government

The primary incomes of the government consist of taxes on production and imports and property income. Taxes on production and imports are distinguished from ordinary taxes imposed on income, wealth, etc. and capital taxes in that they are considered to constitute part of the cost of production for the producer and are recorded only as the GG's receipts in the primary income distribution account. They are broadly divided into (1) taxes on production and (2) other taxes on production. The former includes value-added tax, customs duties, and excise tax, the latter includes property taxes, stamp revenue taxes, etc. Property income of the government consists of net receipts of interest and dividends by the government. A uniform proportionally scaling-up method is applied to the distribution to pretax income for these incomes.

(b) Income of the Non-Profit Institutions Serving Households

The NPISH are organizations that provide social and public services to households without pursuing profits, and specifically include political parties, religious organizations, labor unions, and private schools. The DINA Guidelines recommend distributing the income of NPISH, which is usually a very small amount compared to other incomes, proportionally to the rest of factor income. In this paper, we follow the method proposed in the DINA Guidelines.

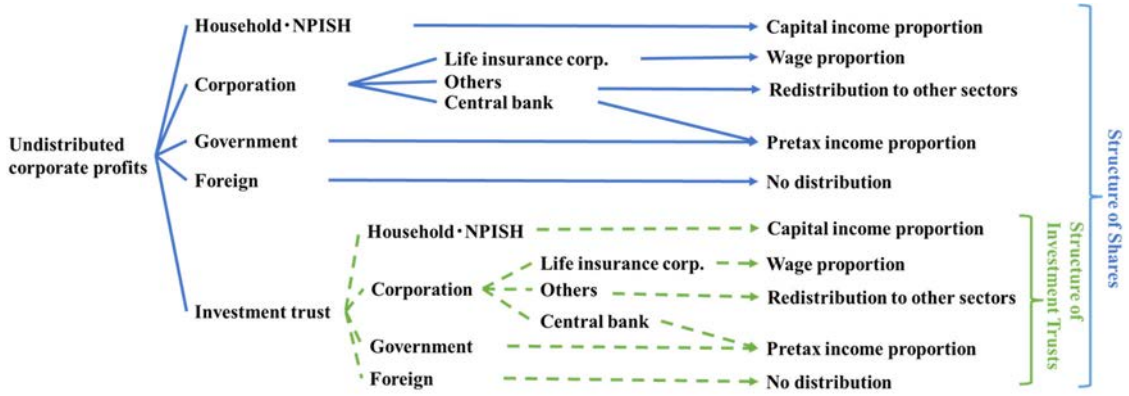
(c) Undistributed corporate profits

Undistributed corporate profits refer to the profits that are not distributed to other sectors through dividends etc., and are the sum of retained earnings of firms before paying corporate income taxes[18]. Since undistributed corporate profits should be imputed to shareholders, the shareholding structure in Japan needs to be carefully considered. The DINA Guidelines propose to use a method of breaking

[18] Since the income concept estimated in this paper is "Pretax, post-replacement income", "Current taxes imposed on income, wealth, etc." is assumed to be pretax for both the household and corporate sectors.

down the shareholders into households, government, and foreign sector. In this paper, we also explicitly consider the shareholding structure through the BOJ, life insurance companies, and investment trusts as shown in Figure 3.

Figure 3. Structure of shareholding and distribution methods



Note: This figure follows the classification of the FFA.

The reason for treating the BOJ separately is that the BOJ has increased its equity holdings in recent years through its ETF holdings. The profits generated by its equity holdings are ultimately paid into the national treasury, and therefore, the BOJ's holdings should be treated in the same manner as the government's holdings. The reason for treating life insurance companies separately is that the main insurance companies in Japan are mutual companies, and the retained earnings need to be imputed to the policyholders, which are mainly households. In addition, the ownership of investment trusts should be carefully considered by using the FFA.

If the imputed shareholders are other than the five sectors (households, government, BOJ, life insurance companies, and foreign sector) through cross-shareholdings, they are also repeatedly imputed to shareholders reflecting the ratio of shareholders to ultimately eliminate the share of the other sectors.

The distribution methods of the undistributed corporate profits imputed to each sector follow the methods presented in the DINA Guidelines. As for households, they are allocated in proportion to capital income. The imputed profits to the government are allocated in proportion to pretax income, as well as the imputed profits to the BOJ. The imputed profits to the foreign sector are not distributed. The imputed profits to life insurance companies are allocated in proportion to wage, on the assumption that the share of insurance policies held by individuals is proportional to wage income.

IV-3. Results of income distribution

The results of the distribution of income other than the merged incomes discussed in the previous

section are shown in Table 7.

For 2019, the results using the JRP dataset show that the shares of the higher income classes, such as the top 10%, 1%, and 0.1%, were higher than the results shown in Table 3. For example, the top 1% share was 8.44%, which is higher by 0.38% point than the result in Table 3. This indicates that the distribution of incomes other than the merged incomes is more skewed toward the top income classes than the distribution of the merged incomes. These conclusions seem to be reasonable because the incomes recognized in this section include dividends, imputed rents, and undistributed corporate profits, which would be expected to be skewed toward the top income classes.

Comparing the 2014 and 2019 results in Table 7, we observe that the income share of the top decile above the top 10% in 2019 is lower than in 2014. This result is also the case for all the cases in Table 3, including the case using only the FICW survey dataset. Although a more detailed study should be needed to explain the recent reduction of the top income shares, it should be noted that in Table 6, the NNI increased from 411.7 trillion yen to 444.9 trillion yen between 2014 and 2019 by approximately 8%, and compensation for employees, including wages and salaries, has risen from 257.5 trillion yen to 286.9 trillion yen by approximately 11%. This means that the increase in the NNI was largely due to the higher contribution of the increase in wages received by employees. Since the distribution of wages is generally considered to be more equally distributed, it seems to be reasonable to understand that the reason for the recent change in the distribution of the NNI can be attributable to the increase in the share of wages and salaries. In the meantime, it should be noted that in this paper we obtain the distribution of income per individualistic adult, not the distribution of income per household. If the labor participation rates of women and the elderly have been increasing in recent years, thereby increasing the incomes of individuals who have not previously earned income, the income share of the top income groups may decline. In addition, the NNI does not include capital gains and losses, and the results in Tables 3 and 7 do not show any distributional impacts of changes in capital gains and losses on individuals. According to Kunieda and Yoneta (2023), the estimated results of the Pareto coefficient of individuals have been gradually increasing in recent years, and this is not contradicting to the results in this paper, because the overall household incomes in their paper include capital gains and losses. To derive further implications for overall inequality and necessary policies in Japan, it is necessary to consider the distributional impacts of capital gains and losses.

Table 7. Distribution of the NNI in Japan

<2014>

Income group	Number of adults	JRP+ASS, FICW			SAI+ASS, FICW		
		Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)
Full population	104,865,000	100.00%	-	3,964,146	100.00%	-	3,964,146
Bottom 50%	52,432,500	12.77%	-	1,068,660	12.80%	-	1,071,555
Bottom 20%(P0-P20)	20,973,000	0.37%	-	107,164	0.33%	-	98,901
Next 30%(P20-P50)	31,459,500	12.40%	728,862	1,709,657	12.47%	728,745	1,719,991
Middle 40%(P50-P90)	41,946,000	49.90%	1,989,607	5,087,668	50.08%	2,002,010	5,114,100
Top 10%	10,486,500	35.19%	8,581,738	13,948,377	34.88%	9,012,682	13,828,803
Top 1%	1,048,650	8.86%	19,466,627	35,140,438	8.07%	19,383,754	31,997,516
Top 0.1%	104,865	2.70%	53,912,742	107,119,133	2.11%	45,379,910	83,594,633

<2019>

Income group	Number of adults	JRP+ASS, FICW			SAI+ASS, FICW		
		Income share	Income threshold (yen)	Average income (yen)	Income share	Income threshold (yen)	Average income (yen)
Full population	105,134,000	100.00%	-	4,229,084	100.00%	-	4,229,084
Bottom 50%	52,567,000	14.69%	-	1,304,371	14.62%	-	1,297,889
Bottom 20%(P0-P20)	21,026,800	1.12%	-	286,687	1.06%	-	275,325
Next 30%(P20-P50)	31,540,200	13.57%	1,032,721	1,982,827	13.55%	1,034,780	1,979,598
Middle 40%(P50-P90)	42,053,600	49.45%	2,331,566	5,369,727	49.31%	2,323,476	5,360,023
Top 10%	10,513,400	33.79%	8,908,574	14,290,558	33.96%	9,106,093	14,361,826
Top 1%	1,051,340	8.44%	20,300,674	35,685,004	8.07%	20,680,559	34,119,270
Top 0.1%	105,134	2.47%	54,266,914	104,441,336	2.07%	48,745,012	87,649,594

Furthermore, Table 8 compares the 2019 results for the top 1% and top 10% income shares with those of other countries. As already mentioned, the results for Japan are not based on the equal-split amount of household income, but show the distribution of income based on the individualistic adults. Therefore, in a strict sense, they are not simply comparable to the results for the United States and France. In general, the distribution based on the individualistic adults is expected to have a higher share of the upper income bracket than the equal-split distribution, since the equal-split treatment results in an equal division of income in the higher income brackets. For Japan, the share of the top 1% is lower than that of the other countries, even though it shows the results based on the individualistic adults. This may indicate that the share of the top 1% income bracket is clearly smaller in Japan than in other countries as for the distribution of the NNI.

Table 8. Income distribution, comparison with other countries

	2014			2019		
	Top1%	Top10%	Bottom50%	Top1%	Top10%	Bottom50%
Japan	**8.9%**	**35.2%**	**12.8%**	**8.4%**	**33.8%**	**14.7%**
United States	19.0%	45.6%	13.1%	19.1%	45.7%	13.6%
France	10.5%	32.6%	21.9%	10.0%	32.4%	22.6%
China	13.7%	41.6%	14.4%	14.7%	42.4%	14.0%
South Korea	9.0%	31.8%	21.3%	11.7%	34.6%	20.8%

Note: The values of the United States, France, China, and South Korea are taken from the WID 2022.

As for the top 10% income share, Japan's figure is lower than that of the United States, but higher than that of France and South Korea. These results may be caused by the fact that the results for Japan are based on the individualistic adults rather than the equal-split amount, and need caution when interpreting the results. If the differences in wages between male and female workers are large, the results for the share of income based on the individualistic adults may be much higher than the equal-split results. We do not have sufficient data for Japan to make the estimation based on the equal-split amount, but this is a major challenge for the future research.

V. Conclusion

This paper shows the results of estimating top income shares in Japan using the 2014 and 2019 household survey dataset and the newly aggregated tax dataset from income tax microdata, following the Distributional National Accounts (DINA) Guidelines. By following the DINA procedure, we combine the survey dataset and the newly aggregated income tax dataset for 2014 and 2019 using

the method presented by Blanchet et al. (2022). In addition, we consider various types of incomes which are not included in the survey and tax datasets, but are included in the Net National Income (NNI) concept. The main results using the most recent Japanese survey and tax data show that the top 1% income share in Japan was 8.44%, and the top 10% income share was 33.79% in 2019, which are much lower than the results for the United States and France presented in the World Inequality Report 2022.

This paper contributes to the literature of the income distribution in Japan in two ways. First, by using the newly available tax dataset, we provide the more accurate and internationally comparable income distribution of pretax incomes in 2014 and 2019 for Japan. Second, we carefully followed the procedures to distribute the overall NNI in Japan in addition to the merged incomes. These were achieved through carefully following the methods provided by the DINA Guidelines.

Finally, it should be noted that several limitations of the dataset do not allow for a complete international comparison. The estimates in this paper are based on the individualistic adults and are not strictly comparable to the results estimated by the equal-split amounts. In addition, for some incomes subject to separate taxation, it is difficult to know their distributions. Further improvements will be possible when additional data related to these issues become available in the future. Also, when interpreting the results in this paper, it is necessary to consider that the analysis is limited to the NNI, not including capital gains and losses, and that further aspects need to be considered when trying to understand the holistic views of inequality in Japan, including the impacts of wealth inequality.

References

Alvaredo, F., Atkinson, A. B., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2016) "Distributional National Accounts (DINA) Guidelines: Concepts and Methods Used in WID.world", *WID.world Working Paper*, 2016/2, World Inequality Lab.

Alvaredo, F., Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2017) "World Inequality Report 2018", World Inequality Lab.

Alvaredo, F., Atkinson, A. B., Blanchet, T., Chancel, L., Bauluz, L. E. et al. (2021) "Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database", *Research Report*, World Inequality Lab.

Blanchet, T., Flores, I., and Morgan, M. (2022) "The Weight of the Rich: Improving Surveys Using Tax Data", *The Journal of Economic Inequality*, 20, pp.119-150.

Chancel, L., Piketty, T., Saez, E., and Zucman, G. (2021) "World Inequality Report 2022", World Inequality Lab.

Garbinti, B., Goupille-Lebret, J., and Piketty, T. (2018) "Income Inequality in France, 1900-2014: Evidence from Distributional National Accounts (DINA)", *The Journal of Public Economics*, 162, pp.63-77.

Kunieda, S. and Yoneta, Y. (2023) "Significance of a Tax Data-Based Analysis of Japanese Income Tax System", *NTC Joint Research Discussion Paper Series*, 230100-01ST, National Tax College. (in Japanese)

Ministry of Internal Affairs and Communications (2014) "Creating Attributable Rent Weights for Owner-Occupied Homes" (in Japanese, last viewed on September 30, 2023) https://www.stat.go.jp/data/cpi/2015/kaisetsu/pdf/fu3.pdf

Ministry of Internal Affairs and Communications (2019) "Methodology for Estimating Attributable Rent of Owner-Occupied Households" (in Japanese, last viewed on September 30, 2023) https://www.stat.go.jp/data/zenkokukakei/2019/pdf/kizoku.pdf

Moriguchi, C. and Saez, E. (2008) "The Evolution of Income Concentration in Japan, 1886-2005: Evidence from Income Tax Statistics", *The Review of Economics and Statistics*, 90(4), pp.713-734.

Ohno, T., Imahori, T., and Kojima, T. (2022) "Reduction and Redistribution Effects of Increasing Income Tax Deductions in Income Taxation and Inhabitant Taxation", *PRI Discussion Paper Series*, 22A-03, Policy Research Institute, Ministry of Finance. (in Japanese)

Ohtake F. (2005), *Inequality in Japan: Illusion and Future of Inequality Society*, Nihon Keizai Shimbun, Inc. (in Japanese)

Ohtake, F. and Kohara, M. (2010) "Income Inequality", in Higuchi, Y. (ed.), *Labor Market and Income Distribution* (Series "The Japanese Economy and Economic Policy during the Bubble /

Deflation Era"), Keio University Press. (in Japanese)

Piketty, T. (2013), *Capital in the Twenty-First Century*, Éditions du Seuil. (in French)

Piketty, T., Saez, E., and Zucman, G., (2018) "Distributional National Accounts: Methods and Estimates for the United States", *The Quarterly Journal of Economics*, 133(2), pp.553-609.

Tachibanaki, T. (1998), *Japan's Economic Inequality: A Consideration from Income and Assets*, Iwanami Shoten, Publishers. (in Japanese)

Appendix: Treatment of data in the Sample Survey for Self-Assessment Income (SAI)

The results of the SAI are only published in a tabulated form, and it is necessary to adjust the data to be merged with the survey dataset under the common income concept. To make the merged income, we need to have the income distribution of the sum of (1) business income, (2) agricultural income, (3) forestry income, (4) real estate income, (5) wages and salaries income, (6) miscellaneous income, and (7) pension income. Therefore, the other incomes need to be deleted from the bracket data. The thresholds of the bracket also need to be revised accordingly. In addition, wages and salaries income and pension income must be corrected to the concept of money received.

The correction method of removing the other incomes is to reduce the amounts of other incomes for each bracket, and the threshold of each bracket is also reduced proportionally. The correction to employment earnings is made by calculating the average employment income in each bracket and adding the estimated deductions. The threshold for each bracket is also increased by the rate of increase in that total income. The conversion from pension income to pension earnings is the same method as in the case of employment income.