Impact of Inter-Generational Transfer through Tax and Social Security Systems on Income Inequality in Japan

Suga, Fumihiko

(出版者 / Publisher)
The Institute of Comparative Economic Studies, Hosei University / 法政大学比較経済研究所

(雑誌名 / Journal or Publication Title)
Journal of International Economic Studies

(巻 / Volume)
35

(開始ページ / Start Page)
31

(終了ページ / End Page)
44

(発行年 / Year)
2021-03

(URL)
https://doi.org/10.15002/00025219

Journal of International Economic Studies (2021), No.35, 31–44

Impact of Inter-Generational Transfer through Tax and Social Security Systems on Income Inequality in Japan

Fumihiko Suga*
Faculty of Economics, Kyushu University

Abstract

In this study, microdata from the *Comprehensive Survey of Living Conditions* (CSLC) is used to investigate how inter-generational transfer through tax and social security systems affects income inequality in Japan. To provide accurate income inequality measures, sampling weights are created using microdata from the *Population Census*. To shed light on income inequality from a broad perspective, four inequality measures are provided based on three definitions of income. It was found that income inequality measures tend to be high when adjusted with our sampling weights. Additionally, it was also found that tax and social security systems not only lower the level of income inequality but also slow the pace of the increase.

Keywords: income distribution, inequality, sampling weight adjustment

JEL Classification Codes: **D31, N35, C83**

1. Introduction

Growing attention is being paid recently to the increase in income and wealth inequality, particularly after Piketty (2014) was published. Along with the global trend pointed out by Piketty (2014), previous studies, such as Tachibanaki (1998), report that inequality has also been rising in Japan.

Tachibanaki (1998) insisted that the level of income inequality in Japan is relatively high among the OECD countries. While it was surprising to note that income inequality in Japan is as high as that of the US, Ohtake (2005) pointed out that the income inequality measures reported by Tachibanaki (1998) were not comparable with those reported by the OECD. Ohtake (2005) found that, if income inequality measures are calculated on the basis of redistributed income, income inequality in Japan is close to the average of the OECD countries. Moreover, Ohtake (2005) pointed that the increase in income inequality in Japan was due to population aging. Tachibanaki (2006) and Oshio (2010), on the other hand, found that an increase in income inequality was observed even within a particular age group, and economists have not yet reached a consensus on the margin by which the increase in income inequality is attributable to population aging.

The main issue in the so-called Tachibanaki-Ohtake dispute is two-fold: First, it is necessary to exercise care over the definition of income used to calculate the income inequality measures. The second point is that income inequality differs across different age groups, and thus changes in the composition of age groups can lead to an increase in income inequality, even if inequality is

* f.suga37@gmail.com.

unchanged within an age group. These two points are closely linked to the redistributive mechanisms of tax and social security systems in the sense that the difference in income inequality measures calculated based on different definitions of income is largely due to tax and social security systems, and tax and social security systems play an important role in mitigating the within- and across-age-group income inequality.

The focus of this study is on the inter-generational redistributive mechanisms of tax and social security systems. Microdata from the *Comprehensive Survey of Living Conditions* (CSLC) is used to investigate how income inequality is affected by inter-generational transfer through the tax and social security systems.

2. Sampling Bias and Weighting Adjustment

2.1. Measurement of Income Inequality

It is well-known that there have been dramatic changes in household composition in Japan, due in part to population aging, the nuclearization of households, and the increase in dual-income households. Previous studies, such as Ohtake (2005) or Kitao and Yamada (2019), pointed out that such changes in the composition of households have substantial effects on the levels and trends of inequality measures. It is therefore important to pay attention to such changes in the composition of households when the effect of tax and social security systems is analyzed.

The problem is that the composition of households in the sample does not necessarily reflect that of the population. If the probability of households to be included in the sample is identical for all the households, the sample distribution would reflect the population distribution. In reality, however, the sample inclusion probability differs considerably across households. For example, a household headed by a young single male is less likely to respond to a survey than a household with a married older household head. Moreover, some surveys intentionally over-sample particular households to shed light on disadvantaged households, for instance. Consequently, the sample distribution deviates from the population distribution, and sampling bias arises.

How then can accurate measures of income inequality be obtained using the sample data? Most public surveys provide sampling weights to correct for the bias. [1] However, even if a weighting adjustment using the officially-provided weights is applied, a sampling bias still remains. To show to what extent sample distribution deviates from the population distribution and by what margin weight adjustment corrects for the bias, the shares of each age category and each household type in the CSLC are shown in Figure 1. As can be seen from the figure, the share of each age category and household category deviate from those calculated from the *Population Census*. In particular, households with a younger head and single households are underrepresented.

Why does the sample distribution not conform to the population distribution even if the weighting adjustment is applied? The reason is that the officially provided weights accompanying the CSLC data correct only for regional disproportionality. Moreover, the sample distribution can deviate from the population distribution even if other household characteristics are taken into account in the process of sampling and creating the provided weights, because researchers sometimes drop non-negligible numbers of observations in the process of data cleaning. Thus, to make sample distribution conform to the population distribution, researchers should ideally create their own original sampling weights.

As mentioned above, this study mainly focuses on the role of inter-generational transfer through tax and social security systems. To do so, it is important to have the sample distribution faithfully

[1] A detailed explanation of sampling weights and how they are used is provided in the following section.

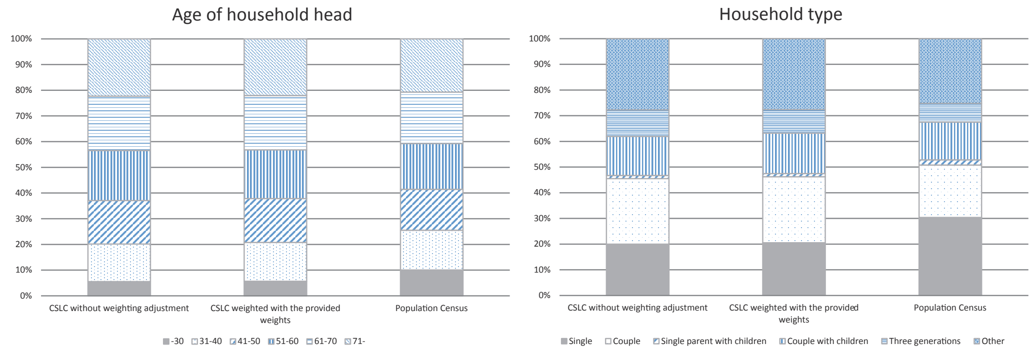

Figure 1: Household share by age category and household type

reflect the population distribution. Otherwise, the proportion of beneficiaries in the sample may be different from that in the population, and the contribution of the tax and social security systems may be under- or over-estimated. This study thus placed emphasis on the weighting adjustment. Original sampling weights were created using microdata from the *Population Census* to provide accurate income inequality measures.

2.2. Weighting adjustment

Weighting adjustment is not a common focus of economic analysis, especially in regression analysis. This is perhaps because economists are interested in relationships between variables and tend to implicitly assume that the parameters relating variables are identical across individuals. Regarding analyses on inequality, however, weighting adjustment is important because inequality measures such as Gini coefficients quantify how many rich or poor people exist and how rich or poor they are. If the sample distribution deviates from the population and rich or poor people are over- or under-represented, income inequality is also likely to be over- or under-estimated.

As mentioned above, it is common for public survey data in Japan to have accompanying sampling weights. The problem is, however, that the weights are created in different ways across different surveys, which can make a huge difference in income inequality measures calculated using different datasets. For example, it is well-known that the value of Gini coefficients calculated using the CSLC data are higher than those calculated using the *National Survey on Family Income and Expenditure* (NSFIE). Moreover, some respondents provide inconsistent responses to survey questions; there are respondents who report that they are working, but earnings are not reported or vice versa, for instance. Such observations have been dropped when analyzing the data. As a result, the sample distribution differs from the raw data and will not conform to the population even if an adjustment using the provided sampling weights is applied. For these reasons, original sampling weights were created as a part of this research.

The most intuitive way of performing weighting adjustment is so-called "cell weighting." The basic idea of cell weighting is that households are sorted into weighting cells defined on the basis of household characteristics, households in each cell being assigned an identical weight. For example, the weighting cell can be defined by household type and households sorted into two cells: single and two-or-more-person households. Suppose we are interested in the (population) mean of income, but the sample is biased in a manner that single households are underrepresented in the sample. The problem, in this case, is that the simple average of income $\overline{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i$ is a biased estimator of the mean if the levels of income differ between single and two-or-more-person households. We then define the sampling weight w_i to observation i and calculate $\tilde{y} = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i$ to obtain an unbiased estimator of the mean of income. In this case, the value of w_i is identical for the households belonging to the same weighting cell (single or two-or-more-person), and the weight assigned to single households is heavier than that assigned to two-or-more-person households.

Let s_j be the share of the households belonging to cell *j* in the population and $\overline{s}_j$ be that in the CSLC. The sampling weight w_j assigned to households belonging to cell *j* is proportional to the ratio of the share of household group *j* in the *Population Census* to that in the CSLC, that is, $w_j \propto s_j / \overline{s}_j$.

An implicit assumption underlying this cell weighting method is that households within each cell are sufficiently homogeneous that households in the sample can be regarded as representative of all households belonging to the weighting cell. This assumption is perhaps not valid in the previous example (of defining the cell by household type). Suppose that the under-sampling of single households in the previous example is due to the low response probability of single households with a younger, working household head. The weighted average $\widetilde{y} = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i$ is no longer an unbiased estimator of the mean, and the sample distribution within the cell to which single households belong differs from the population distribution. In general, the level of income of households with an older, retired head is lower than that of households with a younger, working head. Thus, the mean income is likely to be under-estimated. How can this problem be dealt with? A straightforward way to solve this problem is to employ an additional variable, the age of the household head, to define the weighting cells. That is, households are sorted into four household groups: single households with a head younger than age 60, single households with a single head age 60 or older, two-or-more-person households with a head younger than age 60, and two-or-more-person households with a head aged 60 or older. Let us call such a group a "cell." The variables used to define cells are called auxiliary variables.

To make the households within a cell more homogeneous, more auxiliary variables should be employed. In the previous example, an auxiliary variable, housing status (renters or homeowners), for instance, can be added and eight cells defined. By doing so, the households within each cell become more similar, and the weighted average is expected to be close to the population mean. The problem with using too many auxiliary variables is, however, that there emerge cells with few or no households. For example, if the sample size is 10,000 and the share of single homeowner households with a young head is 0.5%, we would expect to see 50 households in this cell. Since households with a younger head are less likely to respond, the number of households in this cell may be much smaller than 50 households. If there are cells with no observations, the cell would have to be "collapsed" by merging it with neighboring cells. Even if there is no cell containing no observations, cells with few observations can also cause problems. Suppose that there is only one household in a cell. In this case, the weight assigned to this household is likely to be much larger than that assigned to the households in other cells. If the variability of the weight is too large, the variance in the statistics, including income inequality measures, tends to be large.

Weighting adjustment is necessary to obtain accurate income inequality measures, but often increases the variance in income inequality measures. Thus, there exists a tradeoff between precision and variability, and it is necessary to decide on the extent of variance inflation. Previous studies, such as Kish (1992), proposed a variance inflation factor F:

$$F = 1 + CV(w_i)^2$$

where $CV(w_i)^2$ is the coefficient of variation of sampling weight w_i . Five auxiliary variables are employed to define the weighting cells, and the cells are collapsed to maintain the value of F below than that of the provided weights. The auxiliary variables used are the following: i) age of the household head (younger than 40, 40 to 59, 60 or older), ii) household type (single male, single female, a couple, a parent and children, a couple with children, three generations, other), iii) number

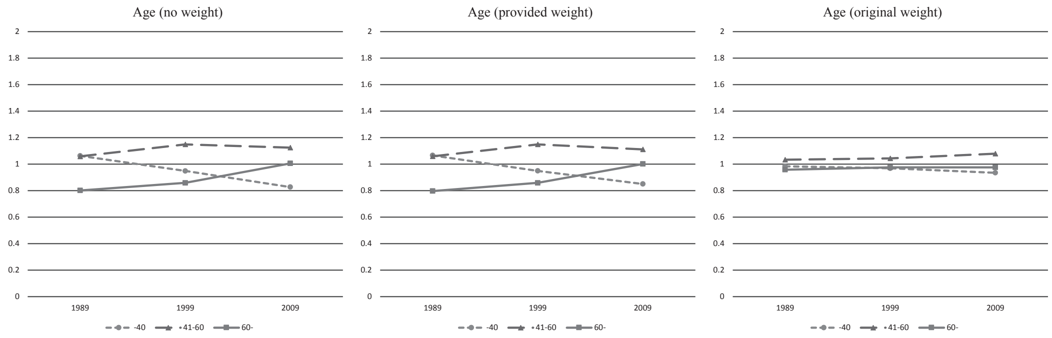

Figure 2-a: Sample disproportionality by weight (age category)

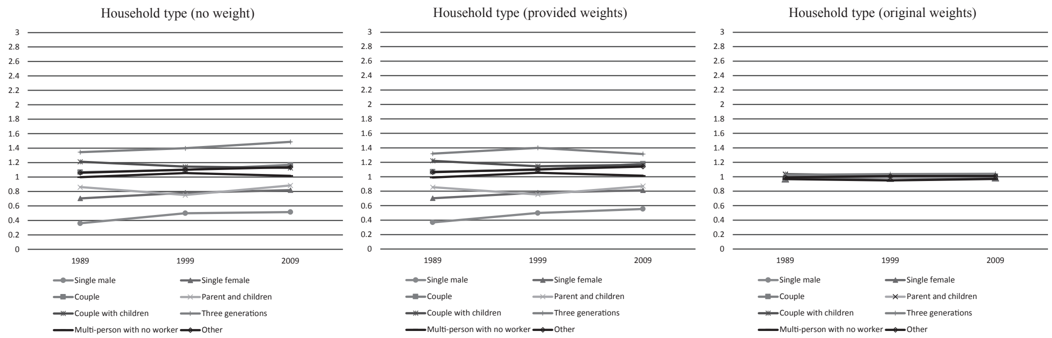

Figure 2-b: Sample disproportionality by weight (household type)

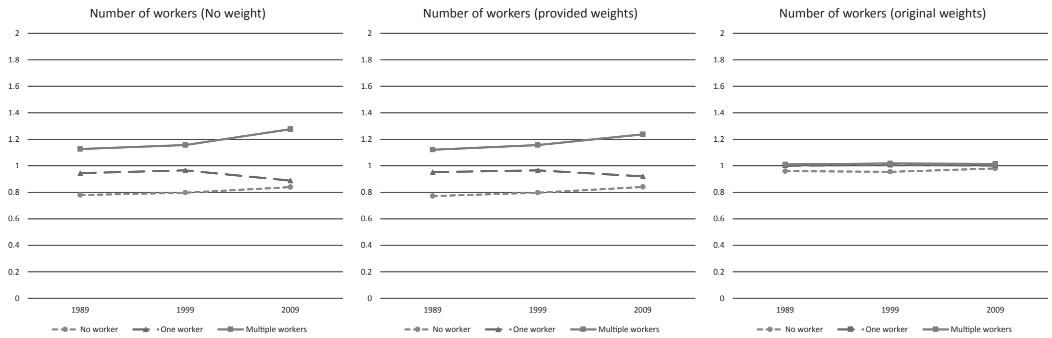

Figure 2-c: Sample disproportionality by weight (number of workers)

of workers (no worker, one worker, two or more workers), iv) area of residence (designated cities, Tokyo special wards, and other cities in five areas (Hokkaido and Tohoku, Kanto, Chubu, Kinki, and Chugoku + Shikoku + Kyushu)), v) homeownership (renter, small homeowner, large homeowner). Thus, there are $3 \times 7 \times 3 \times 7 \times 3 = 1{,}323$ cells in total. However, there exist numerous cells with no observation as well as those that are assigned too heavy a weight. These cells are collapsed in the following manner. First, the no-observation cells are merged with cells in the neighboring area of residence. If there still remain no-observation cells, they are merged with cells in the neighboring age category. The sampling weights are then created, following which F is calculated. If the variance inflation factor F is larger than that of the provided weights, the cell-collapsing procedure is applied

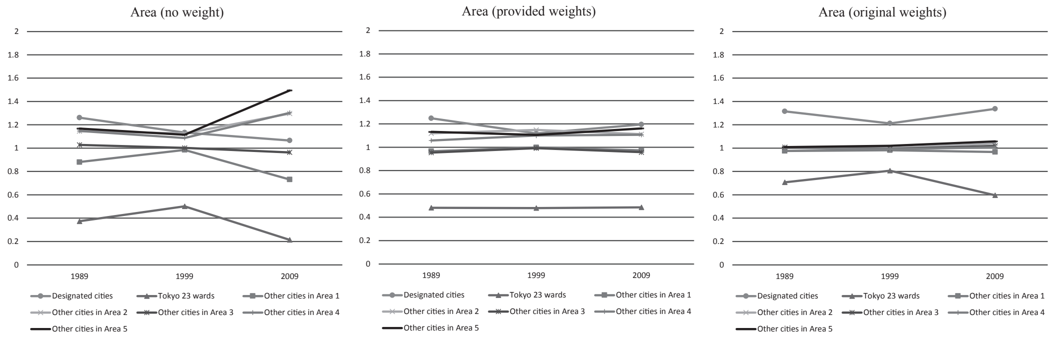

Figure 2-d: Sample disproportionality by weight (area of residence)

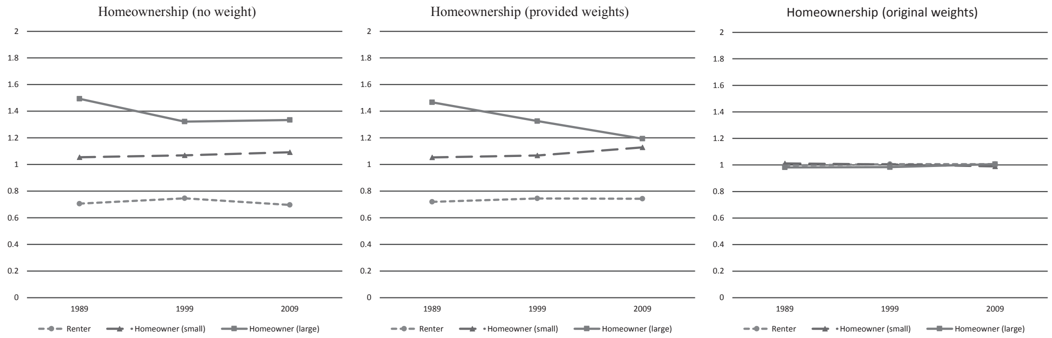

Figure 2-e: Sample disproportionality by weight (homeownership)

to the cell to which the largest weight is assigned. This procedure is repeated until the value of F falls below that of the provided weights.

Figures 2-a to 2-e show the ratio of the share of each category in the CSLC to that in the *Population Census*. If the ratio is larger than 1, households belonging to that category are overrepresented, and vice versa. As can be seen in the figures, adjustment with our original weights generally corrects for the sampling bias. Due to the cells being collapsed with those in the neighboring age or area categories, sampling bias remains for age of the household head and area of residence. Although adjustment with our original weights is not perfect, our original weights perform much better than the provided weights, except for the area of residence.

3. Measurement of income inequality

3.1. Definitions of income

As mentioned above, income inequality measures are sensitive to the definition of income. Since this study focuses on the redistributive mechanisms of the tax and social security systems on income inequality, income inequality measures are calculated based on three definitions: i) initial income defined as pre-tax income including neither pension income nor social security benefits, ii) pre-tax income defined as i) plus pension income and social security benefits, and iii) disposable income defined as ii) minus tax and social security payments.

Moreover, previous studies have shown that income inequality measures depend on whether or not they have been equivalized. As households are rapidly nuclearizing in Japan, it is important to calculate income inequality measures on equivalized income. For example, let us consider a household composed of a household head, the household head's spouse, and the parents of the household head. If the household head and the spouse earn 4 million yen and the household head's parents earn 2 million yen, this household is regarded as one household earning 6 million yen. Suppose that the household head decides to live separately from his or her parents. The household would then be split into two households, one earning 4 million yen and the other earning 2 million yen. This can affect the income inequality measure for household income, while income inequality measures on equivalized income do not change. This is why it is necessary to calculate income inequality measures based on equivalized income when households are nuclearizing.

3.2. Income inequality measures

There are many types of income inequality measures. Since each of them has its own advantages, it is difficult to decide on which measures to use. Here, the most common measures are taken up to shed light on income inequality in Japan from various perspectives. The income inequality measures employed are the following:

i) The Gini coefficient G is defined as

$$G = \frac{1}{2n^2\bar{y}}\sum_{i=1}^{n}\sum_{j=1}^{n}| y_i - y_j |,$$

where y_i denotes the income of individual i, $\bar{y}$ denotes the mean of income, and n is the number of observations.

ii) The relative poverty rate is defined as the share of households whose income falls below half the median income.

iii) The mean log deviation (MLD) is defined as

$$\mathrm{MLD} = \frac{1}{n}\sum_{i=1}^{n}\log\left(\frac{\bar{y}}{y_i}\right).$$

iv) The log variance (LV) is defined as

$$\mathrm{LV} = \frac{1}{n}\sum_{i=1}^{n}(\log y_1 - \overline{\log y}).$$

Since this study placed emphasis on the margin of income by which inequality decreases due to the redistributive mechanisms of the tax and social security systems, details of these income inequality measures are not referred to here.

4. Data

The microdata from the CSLC is used to provide the income inequality measures. To obtain more accurate measures, the sampling weights created from the *Population Census* were used.

4.1. Population Census

The *Population Census* is an exhaustive survey which essentially covers all households in Japan. We can therefore regard the household distribution in the *Population Census* as being the true distribution of households in Japan. Unfortunately, the *Population Census* does not contain information on household income, and therefore income inequality measures cannot be calculated using *Population Census* data. The *Population Census*, however, does contain some important variables related to family characteristics that are also available in the CSLC data. Those variables are therefore used to define weighting cells for the *Population Census* and CSLC data. The sampling weights that make the distribution of CSLC sample conform to the true population distribution were created by using the share of each weighting cell in the *Population Census* as the true population distribution.

Table 1. Summary statistics: *Population Census*

	1990	2000	2010
Number of households	40,272,911	46,191,722	49,723,524
Average age of household head	48.3	51.1	54.2
Share of households with a head aged 60 or older	23.4%	31.3%	40.4%
Share of single households	22.8%	27.3%	30.4%
Share of households with two or more workers	43.1%	37.2%	32.5%
Share of homeowner households	59.5%	60.2%	62.8%

Table 2. Summary statistics: CSLC

	1989	1999	2010
Number of households	35,900	28,307	23,766
Average household pre-tax income	600.4	653.9	550.2
Standard Deviation of household pre-tax income	484.4	562.8	463.7
Average age of household head	47.8	50.7	55.8

Since it is necessary to select the households according to the sample selection criteria applied to the CSLC, microdata from the *Population Census* is used to enable the same sample selection criteria to be applied before calculating the share of population groups. Moreover, as the categories of socio-economic characteristics are defined in the same manner as they are in the CSLC, responding households in the CSLC in each population group can be regarded as representing all of the other households in that cell.

Table 1 presents summary statistics of the *Population Census*. It shows that sampling is disproportional. For example, single households are likely to be under-sampled while homeowners tend to be over-sampled.

4.2. Comprehensive Survey of Living Conditions (CSLC)

The CSLC, conducted by the Ministry of Health, Labour and Welfare, is a representative national survey in Japan. The CSLC is an annual cross-sectional survey, a large-scale survey being conducted only once every three years. The data used here is taken from the large-scale surveys conducted in 1989, 1998, and 2010 to overview the transition of income inequality through the 1990s and the 2000s. The sample size of the large-scale survey is extremely large, but the income questionnaire is distributed only to a subsample: The sample size of the 2010 income data is 26,115 (households). Summary statistics of the CSLC data are provided in Table 2.

A distinguishing feature of the CSLC is that the survey interviewers are people who work at public health and welfare offices or public health centers. The advantage of this feature is that the CSLC sample covers more non-standard households, such as poor single households receiving social welfare, than other government surveys. Hashimoto (2011) pointed, however, that the survey interviewers do not undergo official training for asking survey questions. Thus, there may be some problems such as disproportional response rates across population groups. Another problem associated with the sampling of the CSLC is that, unlike other government surveys, such as the NSFIE, the CSLC do not supplement the non-respondents. Therefore, even if the response rate of households with a young single male is lower than for other households, for example, MHLW does not survey additional households with a young single male. Since the sampling weights provided by MHLW correct only for geographical disproportionality, the response rates may depend on household

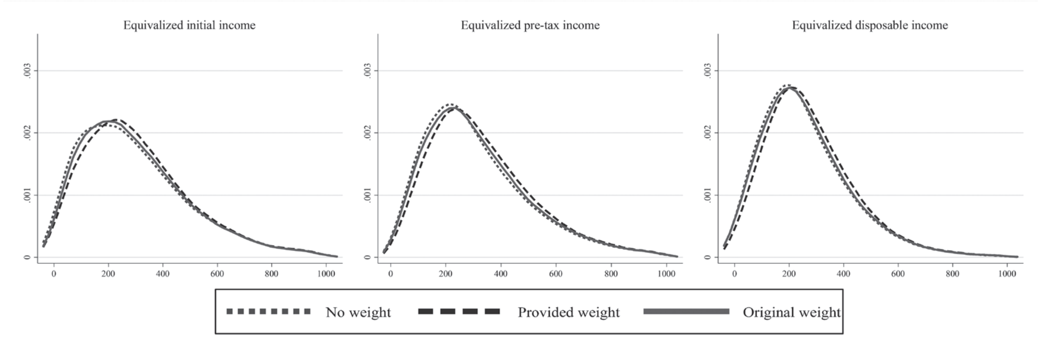

Figure 3: Kernel density by weight (CSLC)

characteristics.

5. Results

5.1. Kernel density of income distribution

To see how income distribution changes with weighting adjustments, let us examine kernel density with and without weighting adjustments. Figure 3 shows the kernel densities produced from the 2010 CSLC data. As can be seen from the figure, the income distribution does not change dramatically. This is perhaps because the CSLC is successful in covering relatively poor households. As Hori, Maeda, and Suga (2020b) pointed out, the income distribution in the NSFIE changes dramatically when weighting adjustment is applied. Thus, the gaps in income inequality measures calculated from the CSLC and the NSFIE are perhaps due to poorer households being underrepresented in the NSFIE.

Weighting adjustment with our original sampling weights bring the share of households belonging to each weighting cell in the CSLC and NSFIE into correspondence with that in the *Population Census*. Thus, adjustment with our sampling weights fills the gap between the income distributions in the CSLC and the NSFIE that can be explained by the auxiliary variables, and it is therefore expected that the income distributions will become similar. To verify this hypothesis, the kernel density was drawn using microdata from the CSLC and the NSFIE. [2] Figure 4-a shows the kernel densities without weighting adjustment obtained from the CSLC and NSFIE data, while Figure 4-b and 4-c present the kernel densities adjusted with the provided weights and our original weights, respectively. What can be seen from Figure 4-a is that the income distribution in the CSLC is more left-skewed than that in the NSFIE. This is perhaps because poorer households are underrepresented in the NSFIE. As can be seen from Figure 4-b, the gap between the income distributions remain even if an adjustment with the provided weights is applied. Finally, Figure 4-c shows that the gap narrows when adjusted with our original sampling weights. It is true that the gap remains for initial income and pre-tax income, but almost disappears for disposable income. Thus, the gap in income distribution between the two surveys can be explained to a considerable extent by the differences in the auxiliary variables (age of household head, household types, number of workers, area of residence, and homeownership rates).

2 The kernel density used here is that presented in Hori, Maeda, and Suga (2020b).

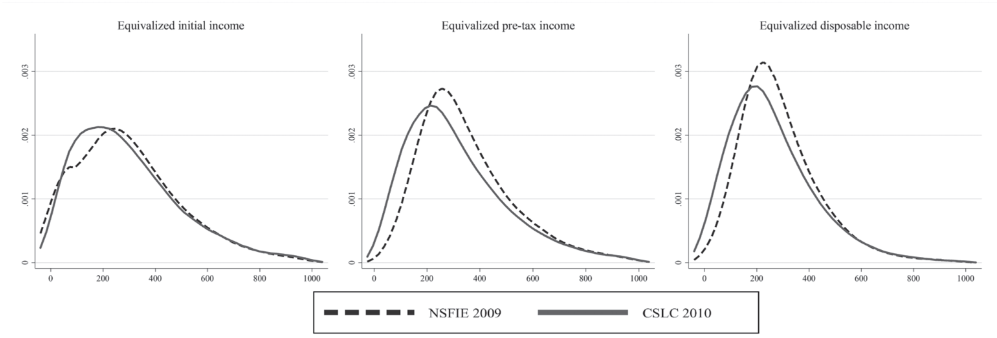

Figure4-a: Kernel density of income distribution by data (no weight)

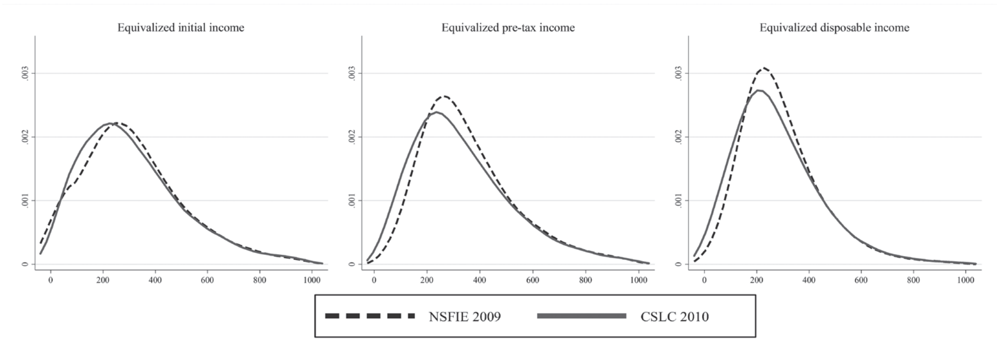

Figure4-b: Kernel density of income distribution by data (adjusted with provided weights)

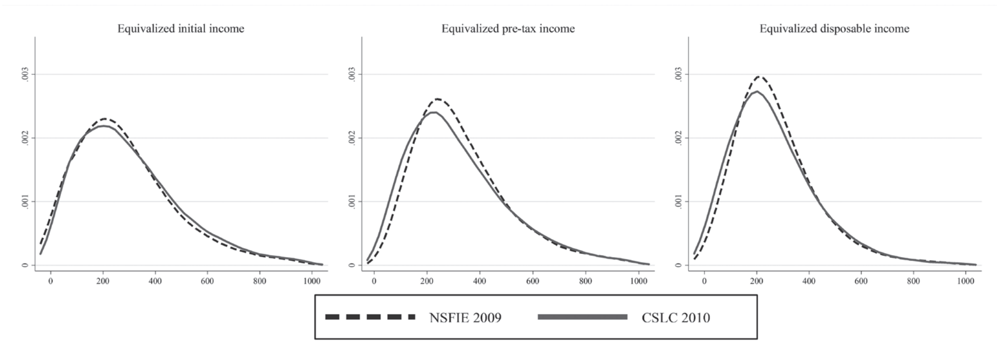

Figure4-c: Kernel density of income distribution by data (adjusted with our original weights)

5.2. Income inequality measures

Figures 5-a to 5-d show income inequality measures with and without weighting adjustment. Only income inequality measures calculated from the CSLC data are presented, since it was confirmed that income distributions are sufficiently close after adjustment with our original weights. It was found that income inequality measures calculated from the CSLC data are not as sensitive as those calculated from the NSFIE data. This is perhaps because the CSLC is more successful in sampling poorer households.

As mentioned above, previous studies revealed that there are huge gaps between income inequality measures calculated from the CSLC and the NSFIE. It was expected that this gap would

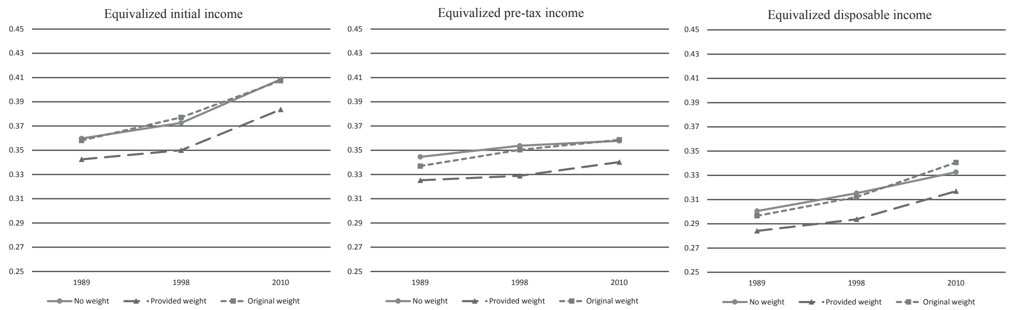

Figure5-a: Gini coefficients by weight (CSLC)

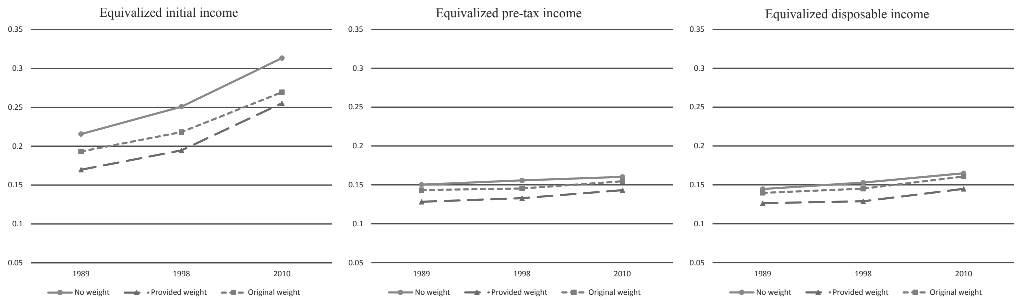

Figure5-b: Relative poverty rate by weight (CSLC)

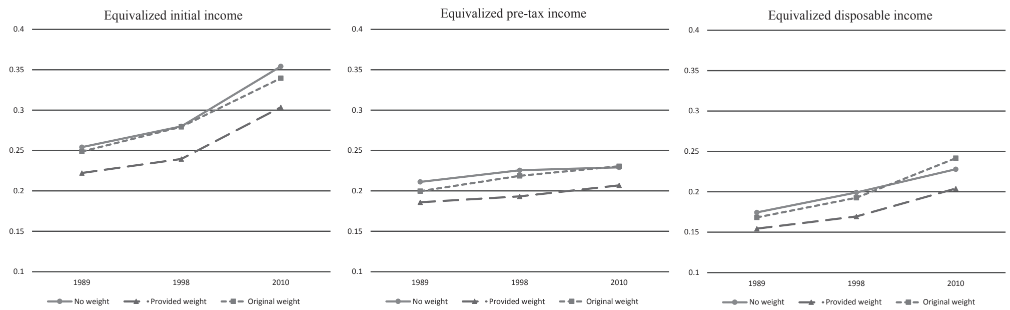

Figure 5-c: MLD by weight (CSLC)

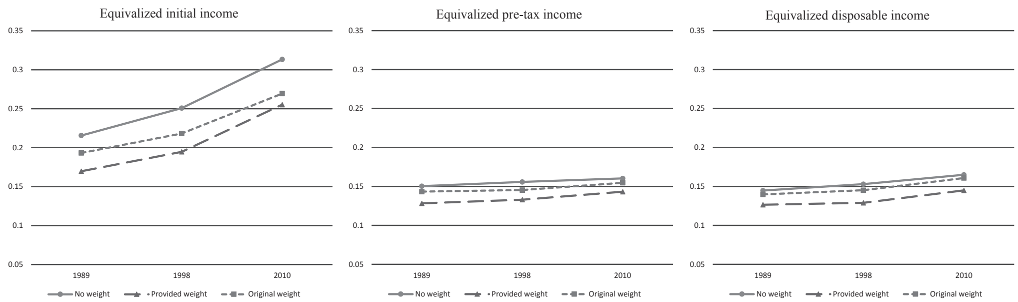

Figure 5-d: LV by weight (CSLC)

be narrowed when adjusted with our original sampling weights. To verify this hypothesis, the Gini coefficient from the 2010 CSLC and the 2009 NSFIE data were calculated with and without weighting adjustment. Without weighting adjustment, the Gini coefficient calculated from the 2010 CSLC data is 0.333 and that calculated from the 2009 NSFIE is 0.302. When adjusted with our original weights, the Gini coefficient calculated from the 2010 CSLC data is 0.341 and that calculated from the 2009 NSFIE is 0.329. Thus, weighting adjustment approximately halved the gap.

As can be seen from Figure 5-a to 5-d, income inequality measures calculated on initial income are considerably higher than those calculated from disposable income. This indicates that redistributive mechanisms of tax and social security systems played an important role in lowering income inequality. Moreover, the pace of increase in income inequality is also lower for the income inequality measures calculated on disposable income. Thus, the inter-generational distributive mechanisms of tax and social security systems not only lower the level but also the speed of increase in the income inequality measures.

5.4. The cause of the increase in income inequality

Let us consider how the contribution of the changes in the composition of households to the increase in income inequality can be measured by using sampling weights. As mentioned, our original

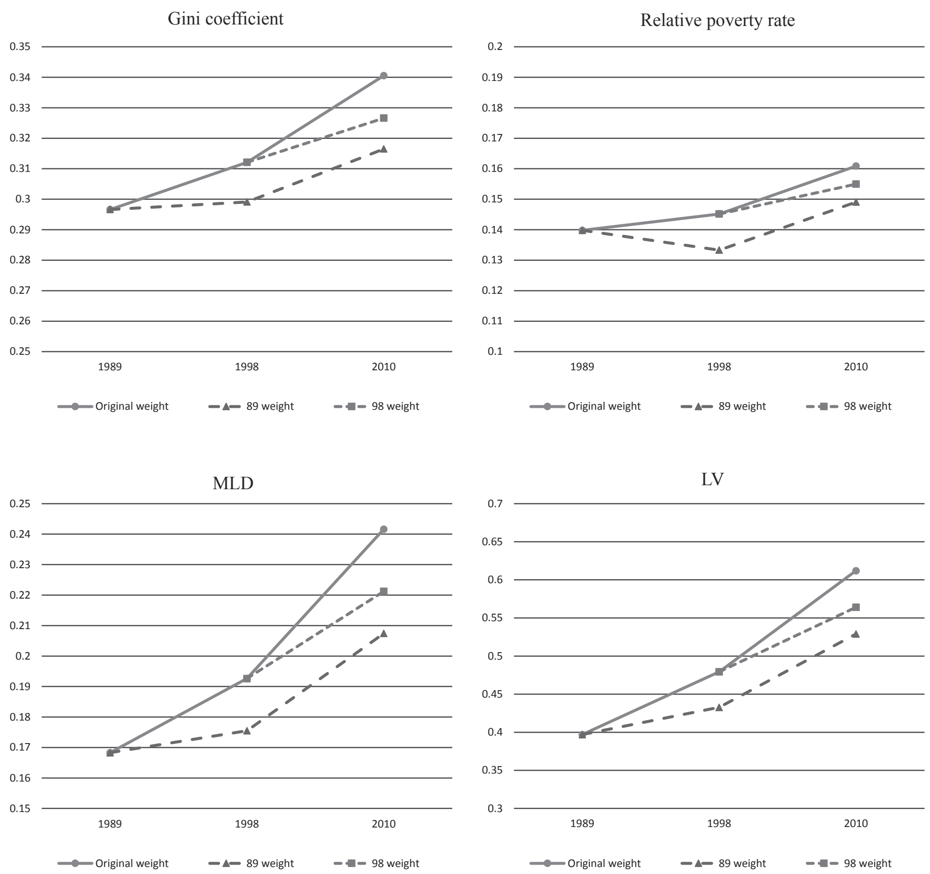

Figure 6: Income inequality measures with fixed share of weighting cells (CSLC)

sampling weights were created by calculating the ratio of the share of households belonging to weighting cell *j* in the *Population Census* to that in the CSLC. Let s_{jt} be the share of households belonging to cell j at year t in the *Population Census* and $\overline{s}_{jt}$ be that in the CSLC. How much income inequality increased can estimated by holding the composition of households at the levels of certain years by using the weights $s_{j1989}/\overline{s}_{j1999}$, $s_{j1989}/\overline{s}_{j2009}$, and $s_{j1999}/\overline{s}_{j2009}$.

Figure 6 shows the income inequality measures adjusted by the weights holding the composition of households at 1989 and 1999 levels. The dashed lines represent income inequality measures holding the household composition at the 1989 level, while the dotted lines represent income inequality measures holding the household composition at the 1999 level. As can be seen from the dashed lines, the greater part of the increase in income inequality during the 1990s can be explained by changes in the composition of households. Thus, as previous studies, such as Ohtake (2005), pointed out, the increase in income inequality during the 1990s can be accounted for by the changes in the composition of households resulting, for example, from population aging, nuclearization of households, or increase in dual-income households. For the 2000s, on the other hand, the changes in the composition of households accounts for at most half of the changes in income inequality measures. Thus, it is necessary to consider factors other than the changes in household composition, such as the collapse of traditional employment systems.

6. Conclusion

The role of inter-generational redistributive mechanisms in income inequality is investigated in this study. Using the microdata from the CSLC, several income inequality measures are provided for different definitions of income to show the margin by which income inequality measures change due to the tax and social security systems. To provide accurate measures, sampling weights were created using microdata from the *Population Census*.

It was found that adjustment with our original sampling weights narrows the gap between the income inequality measures calculated from the CSLC and the NSFIE data. Thus, the income inequality measures obtained are thought to be more accurate than those provided by previous studies using the provided sampling weights. It was also found that the income inequality measures based on equivalized initial income were considerably larger than those based on equivalized disposable income. Moreover, the rate of increase in the income inequality measures based on equivalized initial income was substantially higher than that based on equivalized disposable income. Thus, it can be concluded that the inter-generational redistributive mechanisms of the tax and social security systems not only lower the level but also mitigate the upward trend of income inequality.

Acknowledgements

This paper is a translation of a Japanese article (Suga, 2020) and forms a part of our microdata-based research at the Economic and Social Research Institute (ESRI) on household consumption, labor supply, and macroeconomic policies in Japan. The author would also like to thank Junya Hamaaki, Keiko Murata, Koichiro Iwamoto, and other ESRI colleagues for their support. All results shown in this paper are presented in further detail in Hori, Maeda, and Suga (2020a, b). Special thanks are due to the Statistics Bureau of Japan and the Ministry of Health, Labour, and Welfare for providing the author with the microdata. The views expressed in this paper are those of the author and do not represent those of the institution to which the author belongs.

References

Suga, F. (2020). Zei-Shakaihoshou Seido niyoru Sedaikan Shotoku Saibunpai ga Kakusa ni Ataeru Eikyou (The Effect of Inter-Generational Redistribution through Tax and Social Security Systems on Inequality). in Junya Hamaaki (eds.), *Shoushi Koureika Shakai niokeru Sedaikan Iten to Kazoku*, Tokyo, Japan: Nihon Hyouron Sha, pp. 73-96.

Hashimoto, H. (2011). Kongo no Kokumin Seikatsu Kisochousa no Arikata ni tsuite no Ichikousatsu (A Study on the Nature of Future Basic Surveys of Japanese People's Livelihoods). Kousei no Shihyou, 58(4), 33-39 (in Japanese).

Hori, M., Maeda, S., & Suga, F. (2020a). Income Inequality in Japan: An Analysis using the CSLC dataset. mimeo

Hori, M., Maeda, S., & Suga, F. (2020b). In Search of Accurate Measures of Income Inequality across Japanese Households. mimeo

Kalton, G., & Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19(2), 81.

Kish, L. (1992). Weighting for unequal Pi. *Journal of Official Statistics*, 8(2), 183.

Kitao S. and Yamada T. (2019). Dimensions of Inequality in Japan: Distributions of Earnings, Income and Wealth between 1984 and 2014. RIETI Discussion Paper Series (2019-034).

Lise, J., Sudo, N., Suzuki, M., Yamada, K., & Yamada, T. (2014). Wage, income and consumption inequality in Japan, 1981–2008: From boom to lost decades. *Review of Economic Dynamics*, 17 (4), 582-612.

Ohtake, F. (2005). Nihon no Fubyoudou (Inequality in Japan). Nihon Keizai Shinbunsha, Tokyo (in Japanese).

Piketty, T. (2014). *Capital in the Twenty-First Century*, Cambridge Massachusetts: Belknap Press

Sano, S., Tada, S. & Yamamoto, M. (2015). Setai Chousa no Houhou to Chousa Setai no Seishitsu –Setai Kousei, Nenrei, Gakureki ni kansuru Hikaku– (Methods of Household Survey and the Characteristics of Households – Comparison of Family Composition, Age, and Educational Attainment), Financial Review, (122), 4-24 (in Japanese).

Tachibanaki, T. (1998). Nihon no Keizai Kakusa: Shotoku to Shisan kara Kangaeru (Economic Inequality in Japan: Consideration from the Viewpoint of Income and Assets). Tokyo: Iwanami Shinsho (in Japanese).