



PRI Discussion Paper Series (No.17A-02)

**Estimation of Tax and Social Insurance Burden on Households:**

**Verification of the Validity and Assessment of Actual Status**

*Taro Ohno*

*Associate Professor, Faculty of Economics and Law, Shinshu University*

*Senior Visiting Scholar, Policy Research Institute, Ministry of Finance*

*Takahiro Kodama*

*Researcher, Policy Research Institute, Ministry of Finance*

**January 2017**

The views expressed in this paper are those of the authors and not those of the Ministry of Finance or the Policy Research Institute.

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

# Estimation of Tax and Social Insurance Burden on Households: Verification of the Validity and Assessment of Actual Status<sup>\*</sup>

Taro Ohno<sup>\*1</sup>  
Takahiro Kodama<sup>\*2</sup>

## Abstract

In recent years, household micro data has been streamlined in Japan. Based on this environmental development, the number of analyses using household micro data (questionnaire information) has increased in the fields of taxation and social securities. With the regard to taxes and social insurance premiums for each household account in these analyses, there are two cases: (1) using burdens reported in the questionnaire (Reported Value), and (2) using burdens calculated by applying information, such as family unit, income, etc. reported in the questionnaire to the actual system (Imputed Value).

We used household micro data from the National Survey of Family Income and Expenditure (NSFIE) in 2009 by the Ministry of Internal Affairs and Communications, and estimated the imputed values of tax and social insurance burdens (such as income tax, residence tax, consumption tax, pension insurance premium, health insurance premium, long-term care insurance premium and employment insurance premium) on households. On that basis, we verify the validity of imputed value by comparing it with reported value, and assess the actual status of fiscal burden on household.

The NSFIE is pointed out that the reported values of income tax and social insurance premium are underestimated by the effect of seasonality. Therefore, for verifying the validity of imputed value, we need to focus on the case of residence tax that is less affected by seasonality. And, it was found that the mean disparity between the reported and imputed values was zero and the dispersion was approximately 3% to income. For the implication of the examination, the imputed value has almost no bias in relation to tallied macro values, and it results in an estimated value with almost no error. Therefore, its accuracy can be used sufficiently for the assessment of policies.

The underestimation of reported value may result from the possibility that tax and social premium burdens related to bonuses out of survey period are not reflected completely in the statistics. From the result of examination, such a hypothesis was supported.

In addition, the total fiscal burden on household including taxes and social insurance premiums is progressive both in all samples and in each age bracket. The social insurance premiums and consumption tax have effects to reduce the progressivity. In particular, these are shown up strongly in the elderly generation. Therefore, the progressivity of fiscal burden is smaller in the elderly generation, which implies that intra-generational redistributive effect is weaker in the elderly.

**Key words:** tax, social insurance premium, imputed value, reported value

**JEL Classification:** C15, H24

---

<sup>\*</sup> This research is supported by JSPS KAKENHI Grant Number JP26780176, and uses the information from questionnaires of the National Survey of Family Income and Expenditure (NSFIE) by the Ministry of Internal Affairs and Communications. This report is based on the personal opinions of the authors and does not indicate the official opinions of the institutions to which the authors belong.

<sup>\*1</sup> Associate Professor, Faculty of Economics and Law, Shinshu University; Senior Visiting Scholar, Policy Research Institute, Ministry of Finance

<sup>\*2</sup> Researcher, Policy Research Institute, Ministry of Finance

## 1. Introduction

In recent years, household micro data has been streamlined in Japan. The restructure of tax and social insurance burdens on households is desirable for maintaining the social security system, and thus it is important to evaluate the redistribution effect of changing the tax and social security systems. For doing this, it is necessary to estimate the sophisticatedly imputed value with micro data which reflects heterogeneity across households. In particular, in current Japan with the falling birth rate and the aging population, consumption tax has a critical role to fund the social security cost. Therefore, micro data with households' consumption information is increasingly important to assess the whole picture of tax and social insurance burdens on households<sup>1</sup>.

For household micro data including consumption information in Japan, there exists the National Survey of Family Income and Expenditure (NSFIE) by the Ministry of Internal Affairs and Communications. However, although the NSFIE surveys the burdens of workers' and non-workers' households, it does not survey the burdens of self-employed households. Furthermore, Ohno et al. (2015) show that the income tax and social insurance burdens reported in the questionnaire (Reported Value) are underestimated by the effect of seasonality. Thus, it is difficult to assess the whole picture of tax and social insurance burdens only by using micro data (namely, reported value).

Under this situation, previous studies with using the NSFIE estimate the tax and social insurance burdens by applying information, such as family unit, income, etc. reported in the questionnaire to the actual system (Imputed Value)<sup>2</sup>. Such a method is same as micro simulation that makes original variable virtually by applying available variable to the social system by household, and is able to estimate the burdens of household and variable which are not covered in the survey. However, the imputed value means estimating a variable with strong individuality, such as tax and social insurance burdens, by using limited information, and thus is possible to contain crucial measurement error. Therefore, it is necessary to verify the validity of the imputed value.

The studies on the validity of the imputed value have occurred before, and have received a certain amount of interest from analysts who have implemented microsimulation analysis. For example, Tajika and Yashio (2006a) compares the reported and imputed values of income tax by focusing on the mean value by each income bracket. Tanaka and Shikata (2012) compares the imputed value and macro statistic of tax and social insurance burdens. In addition, Ohno et al. (2015) examine the validity of reported value from the Comprehensive Survey of Living Conditions (CSLC) by the Ministry of Health, Labour and Welfare by comparing it with other surveys, and indicate that it is able to evaluate the validity of the imputed value by comparing it with the reported value of CSLC. And then, they verify the validity of imputed value of income tax and residence tax by using micro data from CSLC, and show that the mean disparity between the reported and imputed values is zero and the dispersion was around 3% over time. Tada et al. (2016) also verify the validity of imputed value of social insurance burdens (such as pension insurance premium, health insurance premium, long-term care insurance premium and employment insurance premium) by using micro data from CSLC, and show that the mean disparity between the reported and imputed values is zero. These examinations imply *that imputed values obtained through microsimulation analysis are estimates with little bias regarding the macro values obtained through collection of data and that there is little margin of error. In that sense, these*

---

<sup>1</sup> There are previous studies which estimate consumption tax burden on households with using micro data, including Ohtake and Kohara (2005), Yashio and Hasegawa (2009), Takayama and Shiraishi (2010, 2011), Shiraishi (2011) and Tanaka (2014).

<sup>2</sup> There are previous studies which analyze tax and social security system with using micro data from the NSFIE, including Tanaka and Shikata (2012), Tanaka et al. (2013), Kitamura and Miyazaki (2013) Miyazaki and Kitamura (2014), and Tanaka (2014).

*imputed values have sufficient precision to be used for the evaluation of policies* (Ohno et al. 2015, p.40).

In this way, previous studies use data from CSLC for verifying the validity of imputed value because the accuracy of reported value exists. Now, we can estimate the imputed value of tax and social insurance burdens because there are also variables required to estimate them in NSFIE. Sano et al. (2015) and Tada and Miyoshi (2015) show that CSLC and NSFIE are consistent with regards to information of family unit and income. Therefore, we can apply the method to estimate imputed value to NSFIE, and furthermore make the imputed value with sufficient precision.

We use household micro data from NSFIE in 2009, and estimate the imputed values of tax and social insurance burdens (such as income tax, residence tax, consumption tax, pension insurance premium, health insurance premium, long-term care insurance premium and employment insurance premium) on households. On that basis, we verify the validity of imputed value by comparing it with reported value. As pointed out above, the NSFIE is shown that the reported values of income tax and social insurance premium are underestimated by the effect of seasonality. Therefore, we focus on the result of residence tax that is less affected by seasonality to verify the validity of imputed value. And, we assess the actual status of the fiscal burden on households.

The outline of this paper will be as follows. First, we confirm data in section 2. Next, we show the method to estimate imputed values of tax and social insurance premiums in section 3. In section 4, we verify the validity of imputed value by examining the disparity between reported and imputed values. In section 5, we assess the actual status of tax and social insurance burdens on households. Finally, we conclude in section 6.

## **2. Data**

### **2.1. Data**

In this analysis, we use micro data from the NSFIE in 2009 (including Household Questionnaire, Yearly Income and Savings Questionnaire, and Family Account Book). This survey is conducted every five years and targets approximately 57,000 family units throughout Japan from September (October in cases of 4,400 single-person households) to November. And, it surveys several household types at the first month of the survey period, yearly income of the last year and consumption expenditure and non-consumption expenditure during survey period. Non-consumption expenditure contains income tax, residence tax, public pension insurance premium, health insurance premium, long-term care insurance premium and other insurance premium (namely, employment insurance premium), so detailed data can be used. One of the aims of this paper is to estimate the imputed values of taxes and social insurance premiums by applying the information of household type and income to real tax and social security system by each family unit, and to verify the validity of the imputed value by comparing it to reported value. Here, we eliminate the following family units that we cannot compare reported value and imputed value.

- Family units which have a member without age or sex information
- Family units which have a member living away from home for business
- Family units which have a member who deviated from the family during survey period
- Family units which have no information on some important items,  
such as income, consumption, taxes and social insurance premiums

After this elimination, the sample size is 49,977 household units<sup>3</sup>.

## 2.2. Income data

We use yearly income data from the “Yearly Income and Savings Questionnaire” (not monthly data from the “Family Account Book”). This data includes the following 10 items.

- (1) Income from employment
- (2) Income from agriculture, forestry and fishery
- (3) Income from business other than (2)
- (4) Income through Piecework
- (5) Income from house and land rents
- (6) Annuities or pensions
- (7) Company and private pension benefits
- (8) Interest and dividends
- (9) Remittance from relatives
- (10) Other income

The “Yearly Income and Savings Questionnaire” of the NSFIE asks yearly income of “household head”, “spouse”, “other household members (aged under 65)” and “other household members (aged 65 or older)”. Thus, in case of the family unit which has more than one “other household member” in each category, namely under-65 and 65-or-older, we can only use total amount of income of members belong to each category. In order to divide the total amount of income into members individually, we accept the following rule.

In the case of income items (1)(2)(3)(4)(6)(7), these income levels seem to be different by sex and age. Therefore, first, we calculate the average income by sex (male or female) and age (15-19, 20-29, 30-39, 40-49, 50-59, 60-69, and 70 or older) group from income data of household heads and spouses whose individual income data are available originally. Second, we divide the total income of “other household members (aged under 65)” and “other household members (aged 65 or older)” into each members by applying ratio of the average income that we calculated above. For instance, consider that one family unit which has two members belong to “other household members (aged under 65)”, and they are “36-years-old male” and “28-years-old female”. Then, we calculate the income from employment of the “36-years-old male” by applying the following formula.

$$A = \alpha \times \beta 1 / (\beta 1 + \beta 2)$$

$\alpha$  = Income from employment of “other household members (aged under 65)”

$\beta 1$  = the average of income from employment of “male / age 30-39” group  
which is calculated from the income data of household heads and spouses

$\beta 2$  = the average of income from employment of “female / age 20-29” group  
which is calculated from the income data of household heads and spouses

---

<sup>3</sup> In the NSFIE, households are classified into “workers households”, “non-workers households”, and “other households” (such as, self-employed workers, etc.). The NSFIE surveys consumption expenditure including taxes and social insurance premiums (reported values) only to “workers households” and “non-workers households”. Therefore, when we verify the validity of imputed value of tax and social insurance burdens, the sample is restricted to “workers households” and “non-workers households”, and thus sample size reduces to 43,391 household units.

In the case of income items (5)(8)(10), we divide total income into each members equally by the number of members who belong to “other household members (aged under 65)” or “other household members (aged 65 or older)”. However, members aged under 15 are not included in the subject for dividing total income.

### **2.3. Consumption data**

We use the consumption data from “Family Account Book”. This questionnaire asks the following 10 items. Here, the total amount of consumption means the sum of these items.

- Food
- Housing
- Fuel, light and water charges
- Furniture and household utensils
- Clothes and footwear
- Medical care
- Transportation and communication
- Education
- Reading and recreation
- Other living expenditure

For estimating consumption tax burden on household, we need to consider the expenditure items on which consumption tax is not imposed, namely “tax-free items”. Here, we define the following subdivided items as “tax-free items”.

- Rents for dwelling
- Rents for land
- Fire and earthquake insurance premium
- Medicines
- Medical treatment
- Dental treatment
- Osteopathic, acupunctural and moxibustic services
- Automotive insurance premium
- School fees
- School textbooks and reference books for study
- Overseas package tours
- Nursery fees
- Nursing care services
- Others

## **3. Estimation Method**

### **3.1. Direct tax**

#### **3.1.1 Income tax and residence tax**

We estimate the imputed values of income tax and resident tax paid by each household. Note that despite income being separated into 10 types under Japanese Tax Act, we use limited types of income such as employment income, employer income, miscellaneous income and real estate income whose data are available in the NSFIE. We assume that each household chooses the separate taxation about their interest or dividends, and thus we omit these kinds of incomes

from the formula of the comprehensive taxation. Here, we calculate total income by the formula as follow. (The parenthesis means item name applied in questionnaire.)

$$\text{Employment income} = [(1)\text{Income from employment}] - \text{Employment income exemption}$$

$$\text{Pension income} = [(6)\text{Annuity or pensions}] + [(7)\text{Company and private pension benefits}] - \text{Public pension, etc. exemption}$$

$$\begin{aligned} \text{Employer income} = & [(2)\text{Income from agriculture, forestry and fishery}] \\ & + [(3)\text{Income from business other than (2)}] \\ & + [(4)\text{Income through Piecework}] \end{aligned}$$

$$\text{Real estate income} = [(5)\text{Income from house and land rents}]$$

$$\begin{aligned} \text{Total income}^4 = & \text{Employment income} + \text{Pension income} + \text{Employer income} \\ & + \text{Real estate income} \end{aligned}$$

Next, according to Ohno et al. (2015), we apply the income exemption and estimate taxable income by each household. The income exemption contains basic exemption, exemption for spouse, special exemption for spouse, exemption for dependents and social insurance premiums exemption. And, these exemptions are applied as follow.

$$\begin{aligned} \text{Temporary taxable income 1} = & \text{Total income} - \text{Basic exemption} \\ & - \text{Social insurance premiums exemption} \end{aligned}$$

$$\begin{aligned} \text{Temporary taxable income 2} = & \text{Temporary taxable income 1} - \text{Exemption for spouse} \\ & - \text{Special exemption for spouse} \end{aligned}$$

$$\text{Taxable income} = \text{Temporary taxable income 2} - \text{Exemption for dependents}$$

First, we calculate temporary taxable income 1 by deducting the basic exemption and social insurance premiums exemption from total income. Second, if there is a spouse who meets the income criteria, we calculate temporary taxable income 2 by applying exemption for spouse and special exemption for spouse to one of a married couple who has higher temporary taxable income 1. Third, if there are dependents, we apply the exemption for dependents to the family member who has highest temporary taxable income 2. In this calculation, we will try two cases in terms of applying social insurance premiums exemption; one is using the reported value of social insurance premium, and another is using the imputed value of that. Finally, we calculate the imputed values of income tax and residential tax by applying the real table of tax rate to their taxable income.

### 3.1.2 Earned income tax

Since the reported value of income tax in the NSFIE is defined as “earned income tax”, we also need to estimate the imputed value of income tax only from “Income from employment”. We calculate tax base on simply earned income by the formula as follow.

---

<sup>4</sup> “Total income” means income after deduction for expenses here.

Temporary taxable income 1 = Employment income – Basic exemption  
– Social insurance premiums exemption

Temporary taxable income 2 = Temporary taxable income 1 – Exemption for spouse  
– Special exemption for spouse

Taxable income = Temporary taxable income 2 – Exemption for dependents

Then, we estimate the imputed value of earned income tax by applying the same table of tax rate as above.

### **3.2. Social insurance premiums**

In Japan, there are several different social insurance systems for people according to their types of employment, therefore in order to calculate imputed values of social insurance premiums, it must be considered which social insurance system each member of household participate in. In this section, according to Tanaka and Shikata (2012), we estimate which system each member participates in at first, and then imputed values of public pension insurance premium, health insurance premium, long-term care insurance premium and employment insurance premium by applying the real formula of these premiums. Noted that amounts of several premiums are different by households' residence in Japan, but enough information about their residence is not available in NSFIE, hence we use the nationwide average to calculate them here.

#### **3.2.1 Public pension insurance premium**

We regard the household members as participants in employee pensions (so-called, second insured persons) if “Income from employments” are more than the amount of “average hourly wage of short-time workers × 30(hours) × 52(weeks)”, the spouses of them as the third insured persons whose annual income are less than 1.3 million yen, and the others as the participants in national pension (so-called, first insured persons). Household members under 19 years old or over 60 years old do not pay their premiums in principle. However, we regard the household members who meet the above income criteria for No.2 category of insured as the participants in employee pension even if they are under 19 or over 60.

We estimating premiums as follow. The first insured persons pay fixed amount of premium (14,660 yen per month in 2009). But, people who meet the income criteria for reduction and exemption always apply the deduction and exemption. Since the premium of second insured persons is paid in equal shares by employer and employee, we estimate their premiums by multiplying their incomes by half of nationwide average premium rate of employee pensions.

#### **3.2.2 Health insurance premium**

We regard household members aged 75 or over as participants in the medical care system for the latter-stage elderly. In the case of members aged 74 or younger, we regard the participants in the employee pensions as participants in the employee health insurances, and others as participants in the national health insurance. But, we regard the household members as the dependents of participants in the employee health insurances if their annual incomes are less than 1.3 million yen and there is other member that participates in the employee health insurances.

We calculate premiums as follow. First, since the premium of employee health insurances is paid in equal shares by employer and employee, we estimate participants' premiums by multiplying their incomes by half of nationwide average premium rate of employee health

insurances. Next, in case of the national health insurance, we estimate the premiums by applying nationwide average premium rate levied on income, nationwide average premium amount levied on property, premium amount per capita and per household, and then summing each household member's premium up. Lastly, in case of the medical care system for the latter-stage elderly, we estimate the premiums by applying nationwide average amount levied on income and levied per household. Besides, we consider the levy limit and reduction in each system.

### **3.2.3 Long-term care insurance premium**

In case of the first insured persons in long-term care insurance (65 years old or older), we estimate the premiums by applying the weighted average premium amount between prefectures. In the case of the second insured persons (between 40 and 64 years old), we consider two cases in terms of participants, such as participants in the national health insurance and participants in the employee health insurances. In the case of the national health insurance, we estimate the premiums by applying nationwide average premium rate levied on income and nationwide average premium amount levied on property, premium amount per capita and per household, and then summing each household member's premium up. In the case of the employee health insurances, since the premium is paid in equal shares by employer and employee, we estimate by multiplying participants' incomes by half of nationwide average premium rate. Besides, we consider the levy limit and reduction in national health insurance, and the upper limit of standard monthly remuneration and bonus in employee health insurances.

### **3.2.4 Employment insurance premium**

We regard the household members as employment insurance participants if "Income from employments" are more than the amount of "average hourly wage of short-time workers  $\times$  20(hours)  $\times$  52(weeks)". We estimate the premiums by multiplying their incomes by the employee burden rate of general industries.

## **3.3. Consumption tax**

We estimate imputed value of consumption tax as follow. First, we compute the monthly amount of taxable consumption by subtracting tax-free items from the total consumption. Next, we multiply it by 12(month) for getting annual amount, and estimate the amount of consumption tax burden by multiplying the annual taxable consumption by consumption tax rate. Here, since the reported amount of consumption in the questionnaire is tax-included, we multiply the annual taxable consumption by 5/105 (because the consumption tax rate was 5% in 2009).

## **4. Validity of Imputed Values**

### **4.1. Comparison of reported and imputed values**

We examine the approximation and disparity between reported and imputed values of tax and social insurance burdens. Here, we use the following disparity rate as index.

$$\text{Disparity rate} = (\text{Reported value} - \text{Imputed value}) / \text{Income}$$

If the disparity rate is larger than zero, it means that reported value is larger than imputed value. In contrast, if disparity rate is smaller than zero, it means that imputed value is larger than reported value.

Table 1 shows descriptive statistics on the disparity rate of tax and social insurance burdens. Figure 1 is histogram of Table 1, and then horizontal axis indicates disparity rate and vertical axis indicates frequency<sup>5</sup>. Here, for estimating imputed values of income and residence taxes, we use the reported values of social insurance premiums in part of the social insurance premiums exemption.

As indicated in Table 1, mean disparity rates are negative in cases of all tax and social insurance premium burdens. Similarly, as indicated in Figure 1, each distribution of disparity rate stands aside to negative direction. Such a negative disparity rate means that imputed value tends to be larger than the reported value. One of reasons is that the reported value is underestimated. Ohno et al. (2015) indicated that the reported values of income tax and social premium burdens are underestimated by the effect of seasonality in the NSFIE, and it may result from the possibility that tax and social premium burdens related to regular income and bonuses out of survey period (from September to November) are not reflected completely in the statistics. Meanwhile, they also indicate that the reported value of residence tax is not underestimated. It may result from the fact that residence tax is imposed on the income of the previous year and collected in monthly installments under this system, thereby is not affected by the seasonality. The underestimation of reported value can have two effects on the disparity rate. First, the underestimation of reported value may lead the disparity rate to negative direction directly. Second, for estimating imputed value of tax including both income and residence taxes, the use of reported values of social insurance premiums may cause “social insurance premiums exemption” to underestimate, thereby cause the imputed value of tax to overestimate. This also may lead the disparity rate to negative direction. These are consistent to the measuring result, and thus can be factors that help to explain the characteristic of the measuring result.

< Insert Table 1 and Figure 1 >

Next, we examine the possibility that the underestimation of reported values of social insurance premiums causes the imputed value of tax to overestimate. For this examination, we use the imputed values of social insurance premiums in “social insurance premiums exemption”. Table 2 shows descriptive statistics on the disparity rate of income and residence taxes. Figure 2 is histogram of Table 2. As indicated in Table 2, compared to the case of using the reported values of social insurance premiums (namely, Table 1), the distribution of disparity rate shifts to positive direction in both income tax case and residence tax case. This indicates that the use of imputed values of social insurance premiums in “social insurance premiums exemption” can mitigate the overestimation of imputed value of tax.

The reported value of residence tax is less affected by the seasonality, and has the validity as an object for comparison. Here, we focus on the result of residence tax in Table 2. The mean disparity rate is 0.95% to income; the 5th percentile is negative 3.89%; the 95th percentile is 3.94%. Therefore, the mean disparity rate is almost zero and dispersion is approximately 3% to income. This is nearly equal to the result of Ohno et al. (2015) using the micro data from the CSLC, which indicated that the mean disparity rate is negative 0.38%; the 5th percentile is negative 3.74%; the 95th percentile is 2.81%. The sample size of this model is approximately

---

<sup>5</sup> It may be not so difficult to lead also to zero imputed value for household units for which the reported value of the burden is zero. It is rather important as a simulation model how close the imputed value can be led to the reported value with respect to household units for which the reported value is positive. Here, we eliminate the household units that both reported and imputed values are zero, and show the measuring result.

43,000 household units, and therefore it is valued that the mean of the sample mean of disparity rate is zero and dispersion is almost zero ( $=0.03/\sqrt{43,000}$ ). Imputed value of microsimulation analysis, etc. have almost no bias in relation to tallied macro values, and it results in an estimated value with almost no errors. Therefore, its accuracy can be used sufficiently for the assessment of policies.

< Insert Table 2 and Figure 2 >

#### **4.2. Verification of a hypothesis on underestimation of reported value**

We focus on the social insurance premium, and empirically verify the factor that causes the underestimation of reported value. If the underestimation of reported value results from the possibility that tax and social premium burdens related to bonuses out of survey period (from September to November) are not reflected completely in the statistics, the distribution of disparity rate will be different according to the difference between households with or without bonuses.

Table 3 shows descriptive statistics on the disparity rate of social insurance premium that are classified by householders' income classes, such as less than 5 million yen, 5 to 10 million yen and more than 10 million yen. Figure 3 is histogram of Table 3. As indicated in Table 3, in lower income class of less than 5 million yen, the mean disparity rate is 0.51%; the 5th percentile is negative 9.04%; the 95th percentile is 10.78%. This means that the mean of distribution is almost zero and the dispersion is bilaterally symmetric. Meanwhile, in the income class of 5 to 10 million yen, the mean disparity rate is negative 2.70%; the 5th percentile is negative 9.88%; the 95th percentile is 3.14%. In the income class of more than 10 million yen, the mean disparity rate is negative 3.97%; the 5th percentile is negative 9.91%; the 95th percentile is 0.43%. In this way, the distribution of disparity rate in higher income class stands aside to negative direction. This implies the possibility that the reported value is underestimated.

< Insert Table 3 and Figure 3 >

Next, Table 4 shows descriptive statistics on the disparity rate of social insurance premium that are classified by householders' occupational categories, such as regular staff, part-time worker and albeit, and dispatched worker from temporary labor agency. Figure 4 is histogram of Table 4. As indicated in Table 4, while the mean disparity rate is negative 2.43% in the regular staff, the mean disparity rate is 0.43% in the part-time worker and albeit. In this way, the distribution of disparity rate in regular staff stands aside to negative direction.

< Insert Table 4 and Figure 4 >

In addition, Table 5 shows descriptive statistics on the disparity rate of social insurance premium that are classified by householders' income classes in the restricted sample to regular staff. Figure 5 is histogram of Table 5. As indicated in Table 5, in lower income class of less than 5 million yen, the mean disparity rate is negative 0.98% and thus the mean of distribution is almost zero. Meanwhile, the mean disparity rate is negative 2.85% in the income class of 5 to 10 million yen, and negative 4.02% in the income class of more than 10 million yen. In this way, the distribution of disparity rate in higher income class stands aside to negative direction.

These results support the effect of social insurance burdens related to bonuses on the underestimation of reported value<sup>6</sup>.

< Insert Table 5 and Figure 5 >

## 5. Actual Status of Tax and Social Insurance Burden on Households

We examine the results of each tax and social insurance premiums (imputed value) by income brackets<sup>7</sup>. First, we show the status of household income and consumption. Table 6 shows the mean income and consumption of households that are classified by income brackets. Here, we use equivalence income decile for making income class. The equivalence income is calculated as follows.

$$\text{Equivalence Income} = \text{Household Income}^8 / \sqrt{\text{the Number of Household Members}}$$

As indicated in Table 6, household income of lowest income bracket, namely 1<sup>st</sup> decile, is 1.391 million yen, and that of highest income bracket, namely 10<sup>th</sup> decile is 13.540 million yen. Meanwhile, household consumption of 1<sup>st</sup> decile is 1.843 million yen, and that of 10<sup>th</sup> decile is 5.032 million yen. In this way, the consumption is larger than income in the 1<sup>st</sup> decile, which indicates that households in this decile spend down their wealth. Base on the life-cycle hypothesis, household smooths their consumption with reflecting the lifetime income. Therefore, the 1<sup>st</sup> decile contains not only poor households but also retired households that are not poor, rather can spend down their wealth. And, when we focus on the number of household members by income bracket, the number of members increases for higher income bracket. Therefore, using equivalence income and consumption reduces the gap between income brackets.

< Insert Table 6 >

### 5.1. Fiscal burdens by income brackets

Table 7 shows the tax and social insurance burdens that are classified by income brackets. Figure 6 is graph of Table 7. Here, we apply 5% of consumption tax rate as of 2009. Burden rate is calculated as follows.

$$\text{Burden Rate} = \text{Burden Amount (Imputed Value)} / \text{Household Income}$$

---

<sup>6</sup> We also examine the distributions of disparity rate that are classified by the number of earners in the household, such as none, one person ,two persons, three or more persons. Table A1 shows descriptive statistics on the disparity rate of social insurance premium that are classified by the number of earners in household. Figure A1 is histogram of Table A1. As indicated in Table A1, the mean disparity rate is 1.04% in the case of a no earner household. Meanwhile, the mean disparity rates are negative 1.8-2.0% in the cases of one or more earner households, and thus the distributions of disparity rate stand aside to a negative direction. Therefore, these results also supports the effect of social insurance burdens related to bonuses on the underestimation of reported value.

<sup>7</sup> Here, we contain not only workers' households and non-workers' households but also self-employed households in the sample.

<sup>8</sup> "Household income" means the initial income which is the income before any deduction here.

As indicated in Figure 6, the burden structures of income tax and residence tax are progressive. For example, income tax is 0.10%, residence tax is 0.34% and the total burden of direct income is 0.44% in 1<sup>st</sup> decile. Meanwhile, income tax is 6.57%, residence tax is 5.16% and the total burden of direct income is 11.74% in 10<sup>th</sup> decile.

Public pension insurance premium of 1<sup>st</sup> decile is 1.26% and that of 10<sup>th</sup> decile is 5.16%, and thus implying that the burden structure of pension insurance premium is also progressive. In Japan, households participate in different pension systems by occupations. The insurance premium of employee pension is constant rate to income, which tends to cause proportional burden structure. The insurance premium of national pension that self-employed workers etc. participate in is constant amount, which tends to cause regressive burden structure. However, the reasons for the progressive burden structure contain that there are exemptions for low income households and that age structure of each income bracket is different. In addition, the burden rate of 10<sup>th</sup> decile is lower than that of 9<sup>th</sup> decile. This reason contains that the burden rate of high income household is restricted by the ceiling of pension insurance premium.

Health insurance premium of 1<sup>st</sup> decile is 3.63% and that of 10<sup>th</sup> decile is 3.53%, and thus implying that the burden structure of health insurance premium is proportional. Similar to pension system, in Japan, households participate in different health insurances by occupations. The insurance premium of health insurance that employees participate in is constant rate to income, which tends to cause proportional burden structure. The insurance premium of National Health Insurance that self-employed workers etc. participate in is mix of constant rate to income and constant amount, which tends to cause regressive burden structure.

Long-term care insurance premium of 1<sup>st</sup> decile is 1.73% and that of 10<sup>th</sup> decile is 0.55%, and thus implying that the burden structure of long-term care insurance is regressive. And, the burden structure of employment insurance premium is proportional. This reason contain that employment insurance premium is imposed on income by constant rate.

Consumption tax of 1<sup>st</sup> decile is 5.32% and that of 10<sup>th</sup> decile is 1.61%, and thus implying that the burden structure of consumption tax is regressive. The burden rate of 1<sup>st</sup> decile is considerably large, which means that low income household bears large tax burden. However, as Yashio and Hasegawa (2009) indicated, low income bracket contains not only low income households through lifetime but also elderly households that face to just temporally low income in a lifecycle. In addition, when we use the burden rate to income for seeing the consumption tax burden, the burden rate on household that spends down their wealth tends to be calculated to become larger.

Figure 7 shows the total burden rate including income tax, residence tax, consumption tax and social insurance premiums. As Figure 7 indicated, the burden structure of tax and social insurance premiums is progressive.

< Insert Table 7 and Figure 6, 7 >

Finally, we compare these results using the imputed values with the case of using the reported values. Ohno et al.(2014) estimated the tax and social insurance burdens on households with using the reported values of the NSFIE. They showed that the burden structures of income tax, residence tax and pension insurance premium are progressive, that the burden structures of health insurance premium, long-term care insurance and consumption tax is regressive, and in addition that the structure of total burden is progressive (Ohno et al. 2014, Figure 1 (2)(3))<sup>9</sup>. These results using the reported values are same as the ones of this paper using the imputed values in terms of the burden structures, except health insurance premium. However, the volumes of burden rates are quite different. Basically, the burden rates of using the imputed values are larger than the ones of using the reported values, especially in the higher income

---

<sup>9</sup> They contain only workers' households and non-workers' households, but not self-employed households in the sample.

bracket. Therefore, the structure of total burden in the case of using the imputed values is more progressive than in the case of using the reported values.

## 5.2. Fiscal burdens by income-age brackets

The amount of tax and social insurance burdens on households are different across generations because of the features of the tax and social security systems, and furthermore the tendencies of income and consumption in a lifecycle. Thus, we examine the fiscal burdens by income-age brackets, where income bracket is income 10 decile, and age bracket contains 15-29, 30-39, 40-49, 50-59, 60-69, 70 and over.

As panel (a) of Figure 8 indicated, direct tax including income tax and residence tax is progressive in each age bracket. But, the direct tax burdens of elderly households are smaller than those of young households even when both households are contained in same income bracket. One of reasons is that tax bases of elderly households are eroded by “deduction for public pension” of income tax etc., and thus their tax burdens are reduced.

Meanwhile, as panel (b) of Figure 8 indicated, social insurance premium is proportional in each age bracket. But, in low income bracket, the burdens of households are relatively small by exemption for each insurance premium. In addition, the social insurance burdens of young households are larger than those of elderly households even when both households are contained in same income bracket. The reason is that the burdens of young households are relatively large by pension insurance premium.

Finally, we look at total burdens on households. As panel (c) of Figure 8 indicated, the burden structure of tax and social insurance premiums is progressive in each age bracket. The social insurance premiums and consumption tax tend to reduce the progressivity. In particular, this is shown up strongly in the elderly generation. Therefore, the progressivity of fiscal burdens is smaller in the elderly generation, which implies that the redistributive effect of intra-generation is weaker in the elderly.

< Insert Figure 8 >

## 6. Conclusion

We used household micro data from the National Survey of Family Income and Expenditure (NSFIE) in 2009, and estimated the imputed values of tax and social insurance burdens (such as income tax, residence tax, consumption tax, pension insurance premium, health insurance premium, long-term care insurance premium and employment insurance premium) on households. On that basis, we verified the validity of imputed value by comparing it with reported value, and assessed the actual status of fiscal burden on household.

The method that evaluates the validity of imputed value by comparing it with reported value has been adopted in Ohno et al. (2015) and Tada et al. (2016) that use the micro data from the Comprehensive Survey of Living Conditions (CSLC). This paper also adopts the same method with using the micro data from the NSFIE. However, the NSFIE is pointed out that the reported values of income tax and social insurance premiums are underestimated by the effect of seasonality. Therefore, for verifying the validity of imputed value, we need to focus on the case of residence tax that is less affected by seasonality. And, it was found that the mean disparity between the reported and imputed values is zero and the dispersion was approximately 3% over income. For the implication of the examination, the imputed value has almost no bias in relation to tallied macro values, and it results in an estimated value with almost no errors. Therefore, its accuracy can be used sufficiently for the assessment of policies.

The underestimation of reported value affects not only the validity for comparing with imputed value, but also estimating the imputed value of tax. This means that when estimating

the imputed value of tax, using reported value of social insurance premium causes underestimation of the “deduction for public pension” of tax, and thus results in overestimation of the imputed value of tax. From the awareness of this issue, when using the imputed value of social insurance premium for the “deduction for public pension”, it was found that the overestimation of the imputed value of tax can be mitigated.

The underestimation of reported value may result from the possibility that tax and social insurance premium burdens related to bonuses out of survey period are not reflected completely in the statistics. From the result of examination, such a hypothesis was supported.

In addition, the total fiscal burden on household including tax and social insurance premiums is progressive both in all samples and in each age bracket. The social insurance premiums and consumption tax have effects to reduce the progressivity. In particular, these are shown up strongly in the elderly generation. Therefore, the progressivity of fiscal burden is smaller in the elderly generation, which implies that intra-generational redistributive effect is weaker in the elderly.

## References

- Kitamura, K. and Miyazaki T. (2013), *Micro Empirical Analysis of Tax Reform: Income Tax and Consumption Tax in Household Economics*, Iwanami Shoten (in Japanese)
- Miyazaki, T. and Y. Kitamura (2014), “Redistributive Effects of Income Tax Rates and Tax Base 1984-2009: Evidence from Japanese Tax Reforms”, *Discussion Paper Series, A No.610*, Institute of Economic Research, Hitotsubashi University
- Ohno, T., M. Nakazawa, K. Kikuta, and M. Yamamoto (2015), “Comparison of Taxes and Social Insurance Premium Burdens in Household Accounts”, *Public Policy Review*, No. 11(4), pp. 547-571
- Ohno, T., M. Nakazawa, K. Matsuda, K. Kikuta, and T. Masuda (2014), “Tax and Social Insurance Burden on Households: Estimating Using National Survey of Family Income and Expenditure”, *Public Policy Review*, No. 10(3), pp. 457-479
- Ohtake, F. and M. Kohara (2005), “Is the Consumption Tax Really Regressive?: Consideration of the Equality of Burden”, *Ronza*, No.127, pp.44-51 (in Japanese)
- Sano, S., S. Tada, and M. Yamamoto (2015), “Method of Household Surveys and Characteristics of Surveyed Households: Comparison regarding Household Composition, Annual Income and Educational Attainment”, *Public Policy Review*, No. 11(4), pp. 505-529
- Shiraishi, K. (2011), “Consumption Tax Payment Level and Regressivity”, Report Paper of the 68th National Meeting of Japan Institute of Public Finance (in Japanese)
- Tada, S. and K. Miyoshi (2015), “Verifying household incomes in Japanese statistics”, *Public Policy Review*, No. 11(4), pp. 531-545
- Tada, S., Ohno, T. and T. Unayama (2016), “Estimation of Social Insurance Premiums using micro data and Verification of the Validity”, *PRI Discussion Paper Series*, 16A-02 (in Japanese)
- Takayama N. and K. Shiraishi (2010), “Burden Level of Consumption Tax in Family Units in Japan”, *Discussion Paper, CIS-PIE DP No.491*, Institute of Economic Research, Hitotsubashi University (in Japanese)
- Takayama N. and Shiraishi K. (2011), “The VAT Credit for Mitigating Regressive Tax Burdens in Japan”, *Discussion Paper, CIS-PIE DP No.503*, Institute of Economic Research, Hitotsubashi University (in Japanese)
- Tanaka S. (2014), “Analysis of Effects of Consumption Tax Measures for Low-Income Earners”, Report Paper of the 71th National Meeting of Japan Institute of Public Finance (in Japanese)
- Tanaka S. and M. Shikata (2012), “Microsimulation of Tax and Social Insurance Premium”, *RISS Discussion Paper Series No. 25*, Research Institute for Socionetwork Strategies, Kansai University (in Japanese)
- Tanaka S., M. Shikata and K. Komamura (2013), “Analysis of Tax and Social Security Burdens on Senior Citizens: Using Individual Questionnaire Data of the National Survey of Family Income and Expenditure”, *Financial Review* No. 115, pp.117-133 (in Japanese)
- Yashio, H. and Y. Hasegawa (2009), “Actual Situation of Consumption Tax Burdens in Household Accounts in Japan”, *Economic Analysis* No. 182, pp.25-47 (in Japanese)

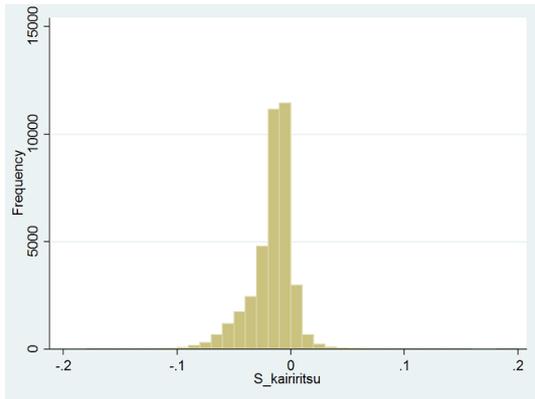
Table 1: Descriptive Statistics on Disparity Rate  
(using the reported values of social insurance premiums)

	Income Tax	Residence Tax	Earned Income Tax
Mean	-1.87%	-0.66%	-1.83%
5 percentile	-6.11%	-3.55%	-6.12%
95 percentile	0.67%	3.12%	0.74%

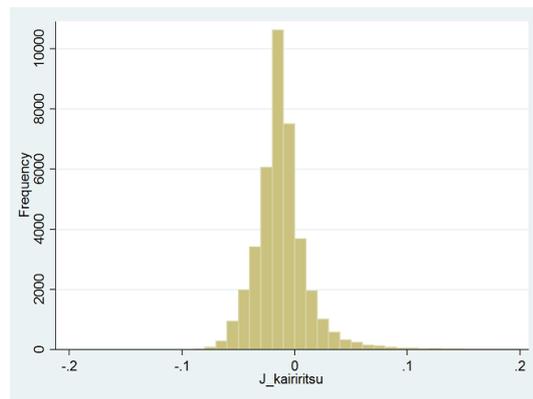
	Pension Insurance Premium	Health Insurance Premium	Long-term Care Insurance Premium	Employment Insurance Premium	Social Insurance Premium
Mean	-0.74%	-0.03%	-0.66%	-0.06%	-1.08%
5 percentile	-7.31%	-4.17%	-2.59%	-0.40%	-9.50%
95 percentile	4.24%	5.73%	0.63%	0.22%	7.15%

Figure 1: Histogram on Disparity Rate  
(using the reported values of social insurance premiums)

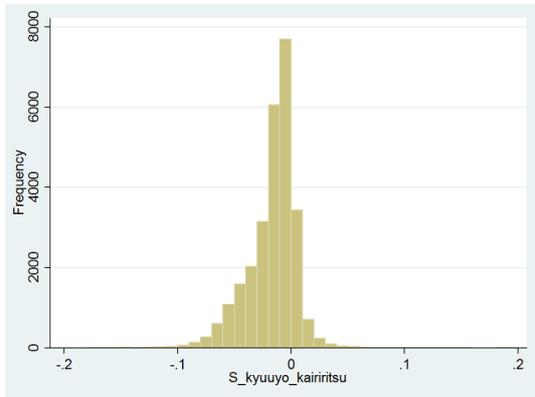
(a) Income tax



(b) Residence tax



(c) Earned income tax



(d) Social insurance premium

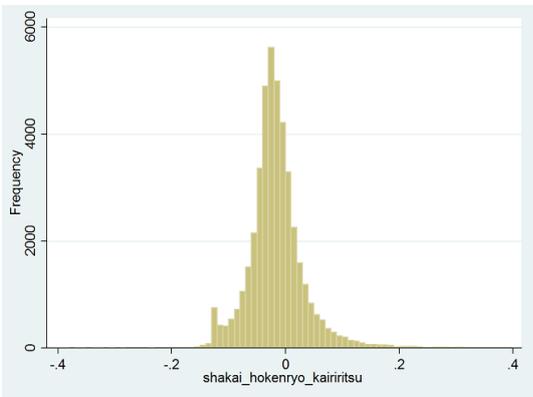
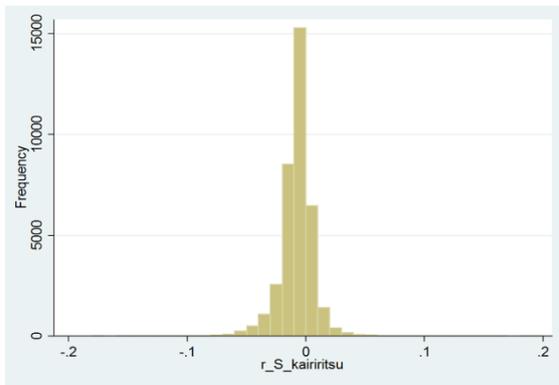


Table 2: Descriptive Statistics on Disparity Rate  
(using the imputed values of social insurance premiums)

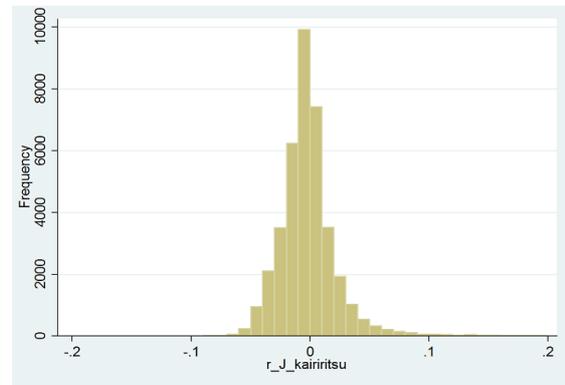
	Income Tax	Residence Tax	Earned Income Tax
Mean	-0.83%	0.95%	-0.65%
5 percentile	-3.43%	-3.89%	-3.72%
95 percentile	1.11%	3.94%	1.46%

Figure 2: Histogram on Disparity Rate  
(using the imputed values of social insurance premiums)

(a) Income tax



(b) Residence tax



(c) Earned income tax

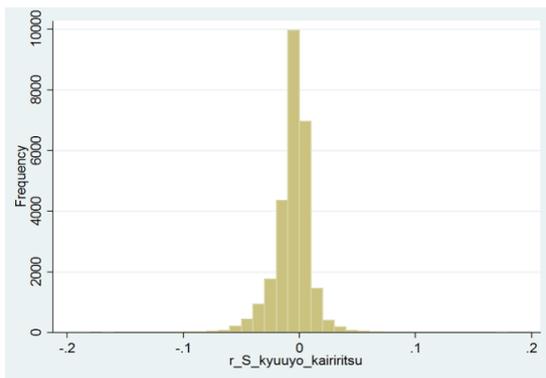
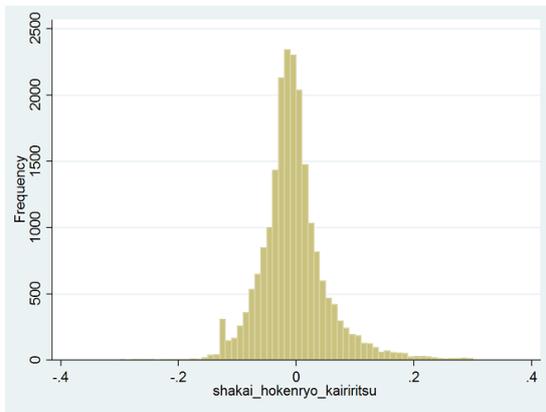


Table 3: Descriptive Statistics on Disparity Rate of Social Insurance Premium  
(by income brackets)

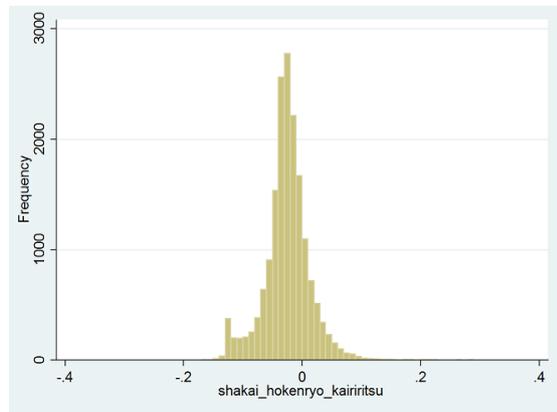
	less than 5 million yen	5 – 10 million yen	more than 10 million yen
Mean	0.51%	-2.70%	-3.97%
5 percentile	-9.04%	-9.88%	-9.91%
95 percentile	10.78%	3.14%	0.43%

Figure 3: Histogram on Disparity Rate of Social Insurance Premium  
(by income brackets)

(a) Less than 5 million yen



(b) 5 – 10 million yen



(c) More than 10 million yen

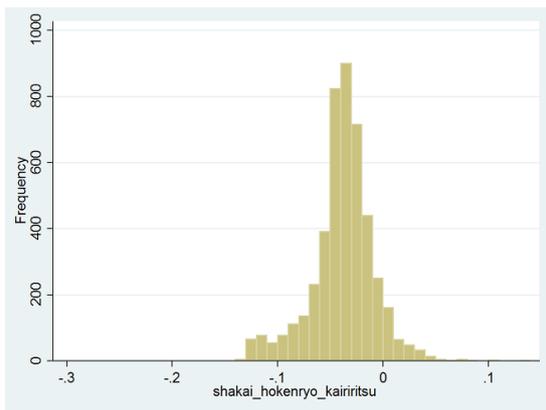
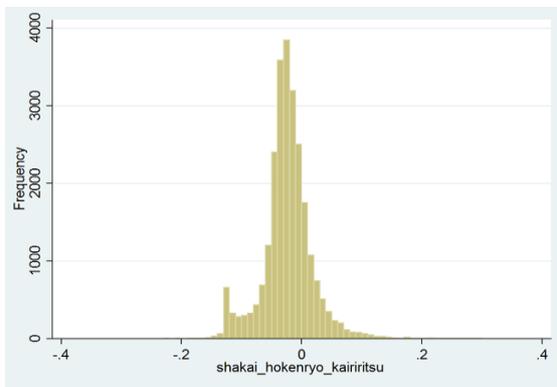


Table 4: Descriptive Statistics on Disparity Rate of Social Insurance Premium  
(by occupational categories)

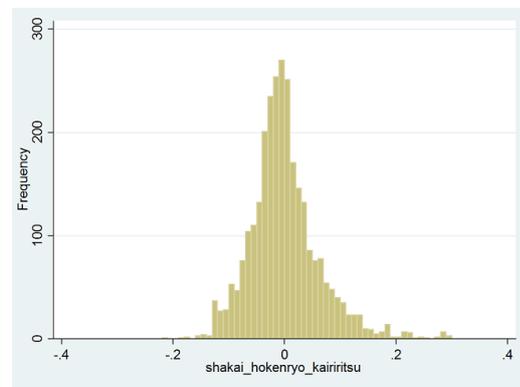
	Regular Staff	Part-time Worker and Albeit	Dispatched Worker
Mean	-2.43%	0.43%	-1.35%
5 percentile	-11.06%	-9.88%	-11.82%
95 percentile	3.99%	13.22%	7.11%

Figure 4: Histogram on Disparity Rate of Social Insurance Premium  
(by occupational categories)

(a) Regular staff



(b) Part-time worker and albeit



(c) Dispatched worker

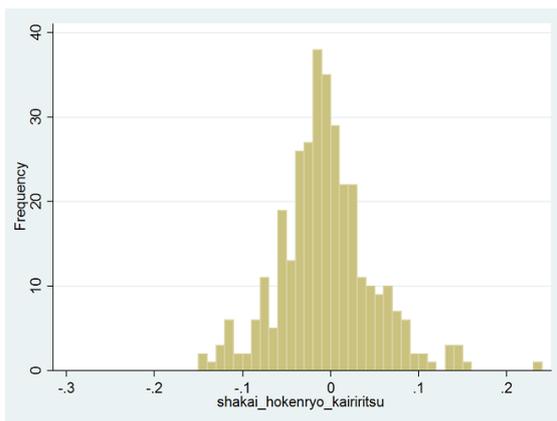


Table 5: Descriptive Statistics on Disparity Rate of Social Insurance Premium

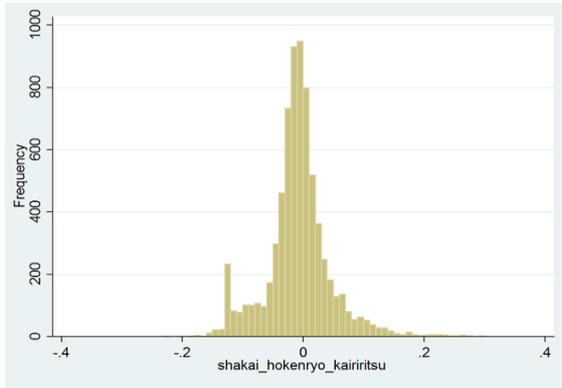
(only regular staff, by income brackets)

	less than 5 million yen	5 – 10 million yen	more than 10 million yen
Mean	-0.98%	-2.85%	-4.02%
5 percentile	-12.03%	-10.64%	-9.84%
95 percentile	7.98%	2.59%	0.03%

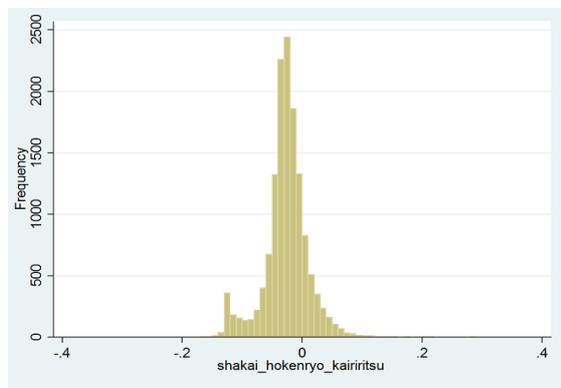
Figure 5: Histogram on Disparity Rate of Social Insurance Premium

(only regular staff, by income brackets)

(a) Less than 5 million yen



(b) 5 – 10 million yen



(c) More than 10 million yen

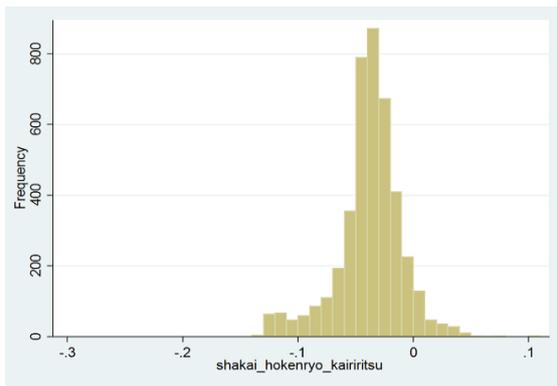


Table 6: Households' Income and Consumption

Income Bracket	Household Income (million yen)	Household Consumption (million yen)	Number of Member (person)	Equivalence Income (million yen)	Equivalence Consumption (million yen)
I	1.391	1.843	2.1	0.993	1.326
II	2.599	2.325	2.5	1.695	1.541
III	3.336	2.619	2.6	2.125	1.693
IV	3.860	2.836	2.6	2.485	1.863
V	4.540	3.091	2.7	2.854	1.975
VI	5.372	3.394	2.9	3.273	2.107
VII	6.184	3.666	2.8	3.787	2.276
VIII	7.287	3.943	2.8	4.459	2.463
IX	8.908	4.350	2.8	5.504	2.715
X	13.540	5.032	2.5	8.870	3.315
Total	5.700	3.309	2.6	3.604	2.127

Table 7: Tax and Social Insurance Burden Rates

Income Bracket											
	Income Tax	Residence Tax	Direct Tax	Pension Ins.	Health Ins.	Long-term Care Ins.	Employment Ins.	Social Ins.	Direct Tax + Social Ins.	Consumption Tax	Direct Tax + Social Ins. + Cons. Tax
I	0.10%	0.34%	0.44%	1.26%	3.63%	1.73%	0.10%	6.73%	7.17%	5.32%	12.49%
II	0.35%	1.10%	1.45%	2.89%	3.94%	1.42%	0.16%	8.41%	9.86%	3.63%	13.49%
III	0.65%	1.77%	2.42%	3.40%	4.13%	1.46%	0.18%	9.17%	11.58%	3.24%	14.82%
IV	0.93%	2.32%	3.25%	3.59%	4.23%	1.38%	0.18%	9.37%	12.62%	3.05%	15.68%
V	1.23%	2.86%	4.09%	4.32%	4.23%	1.11%	0.21%	9.87%	13.96%	2.82%	16.78%
VI	1.49%	3.23%	4.72%	4.79%	4.06%	0.95%	0.22%	10.03%	14.75%	2.63%	17.38%
VII	1.89%	3.67%	5.56%	5.16%	4.01%	0.83%	0.23%	10.23%	15.80%	2.49%	18.28%
VIII	2.43%	4.06%	6.49%	5.58%	3.93%	0.73%	0.24%	10.48%	16.97%	2.26%	19.23%
IX	3.47%	4.56%	8.03%	5.97%	3.90%	0.64%	0.24%	10.75%	18.78%	2.06%	20.84%
X	6.57%	5.16%	11.74%	5.16%	3.53%	0.55%	0.22%	9.45%	21.19%	1.61%	22.79%
Total	2.98%	3.71%	6.68%	4.83%	3.90%	0.88%	0.22%	9.82%	16.50%	2.43%	18.93%

Figure 6: Tax and Social Insurance Burden Rates (Each Burden)

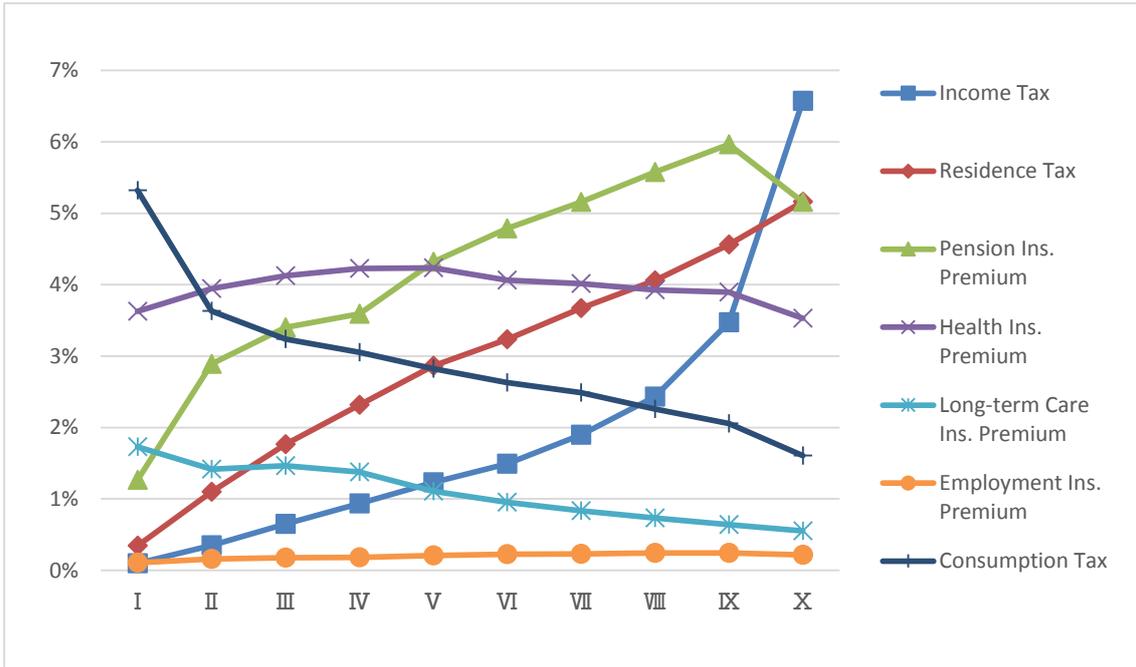


Figure 7: Tax and Social Insurance Burden Rates (Total Burden)

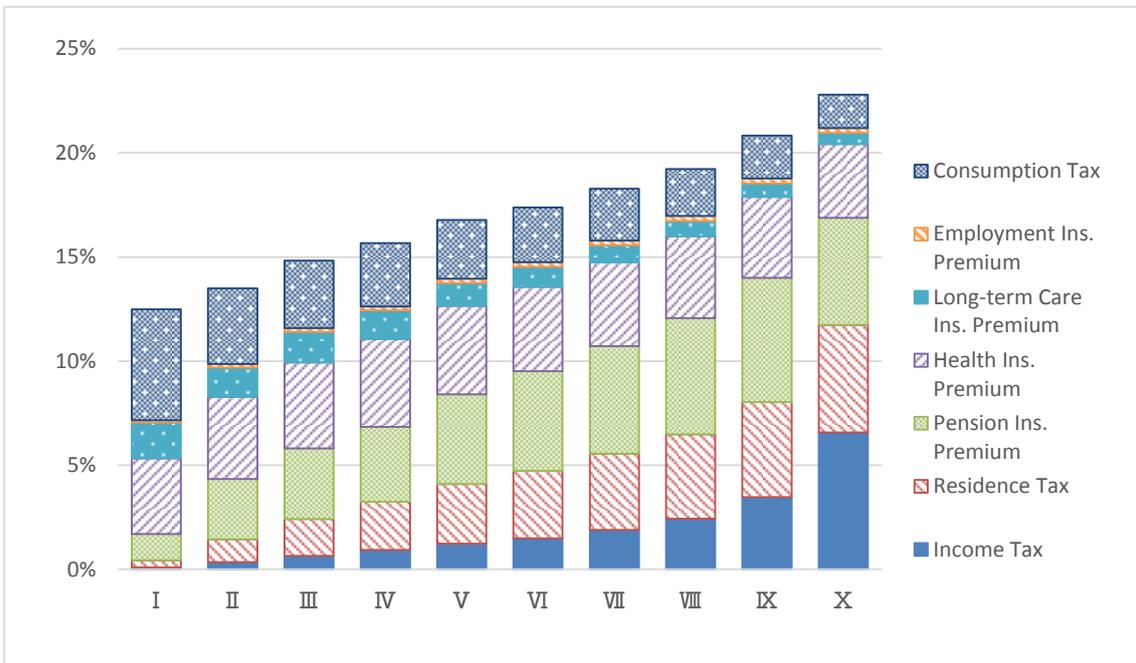
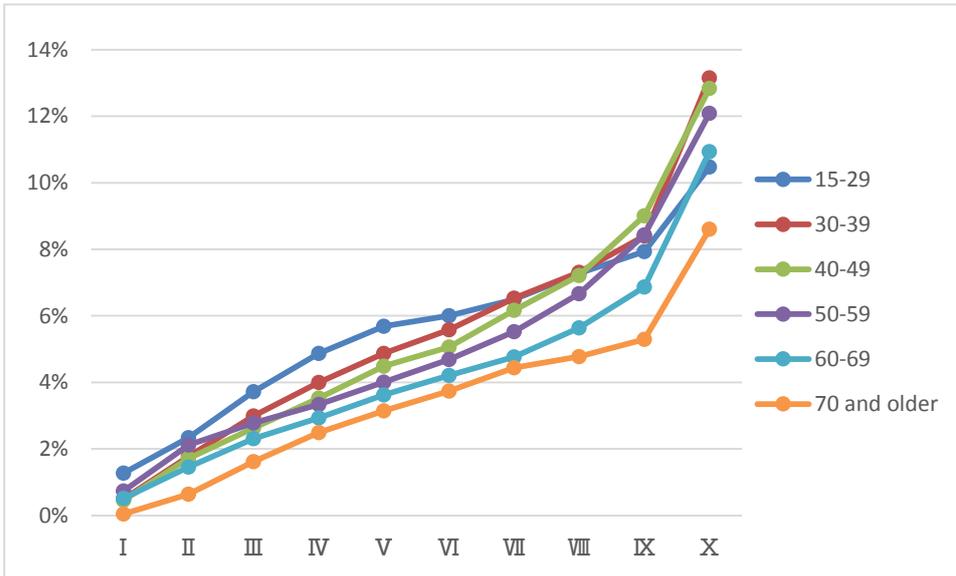


Figure 8: Tax and Social Insurance Burden Rates (By Income-Age Brackets)

(a) Direct tax (including income tax and residence tax)



(b) Social insurance premium

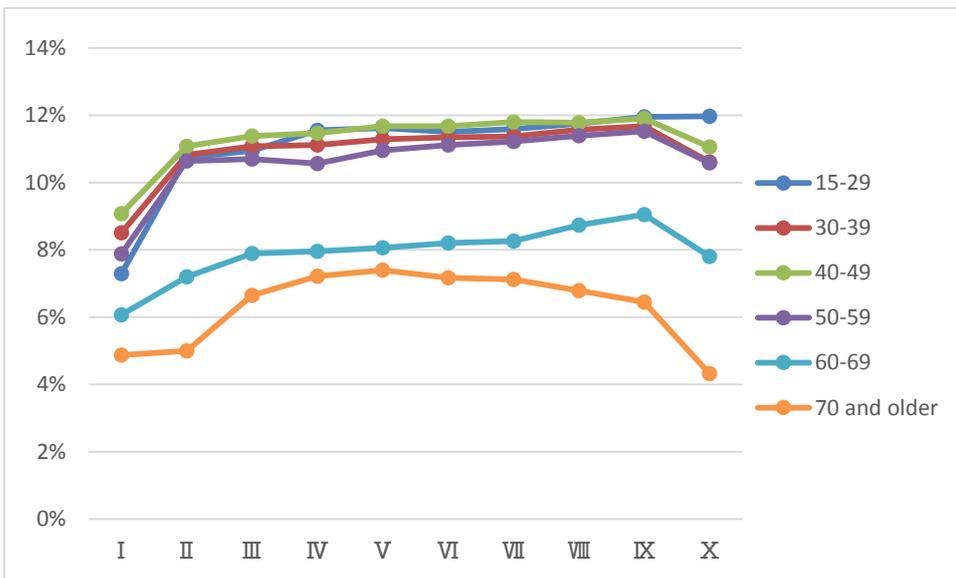


Figure 8: (continued)

(c) Total burden (including direct tax, social insurance premium and consumption tax)

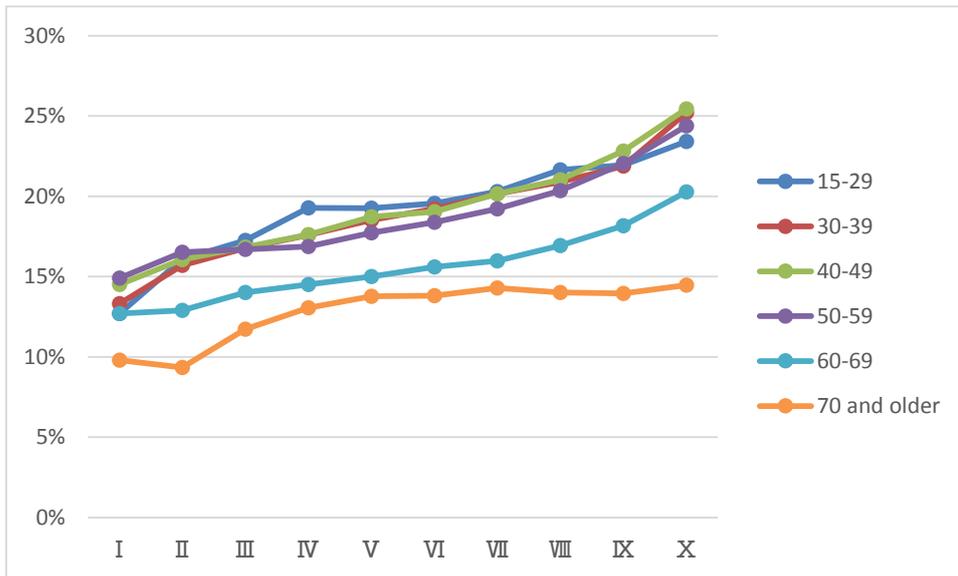


Table A1: Descriptive Statistics on Disparity Rate of Social Insurance Premium  
(by the number of earner)

	the Number of Earner			
	None	1 person	2 persons	3 or more persons
Mean	1.04%	-1.88%	-1.95%	-1.96%
5 percentile	-7.31%	-11.78%	-9.69%	-9.65%
95 percentile	10.13%	6.40%	5.39%	5.91%

Figure A1: Histogram on Disparity Rate of Social Insurance Premium  
(by the number of earner)

