RIETI Discussion Paper Series 13-E-017

Social Security Earnings Test and the Labor Supply of the Elderly: New evidence from unique survey responses in Japan

SHIMIZUTANI Satoshi
RIETI

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 13-E-017

March 2013

Social Security Earnings Test and the Labor Supply of the Elderly: New evidence from unique survey responses in Japan*

SHIMIZUTANI Satoshi†
(RIETI)

Abstract

This study proposes an alternative approach of utilizing direct responses to a survey on the social security earnings test for the elderly to provide new evidence on the sensitivity of the labor supply decision of workers aged between 60 and 64 with respect to the test. Our empirical results show a discouraging effect on working in a large proportion of these workers in Japan, even after correcting for the observed attributes of individuals who reported to be either affected or unaffected.

Keywords: Social security earnings test; *zaishoku rorei nenkin seido*, Labor supply of the elderly; Japan; Wage distribution; DiNardo-Fortin-Lemieux decomposition
JEL classification: H55, J26

* We are grateful to the Ministry of Health, Labour and Welfare for providing us with a micro-level dataset. We also thank Howard Iams, Hidehiko Ichimura, Takashi Oshio, and Izumi Yokoyama for their very constructive comments, which improved the manuscript substantially. The views expressed in this paper are those of the author and do not represent any institution.
† Research Fellow, Research Division, Gender Equality Bureau, Government of Japan; 1-6-1 Nagata-cho, Chiyoda-ku, Tokyo, Japan 100-8914; Tel: +81-3-3581-1854, Fax: +81-3-3592-0408; E-mail: satoshi.shimizutani@cao.go.jp.

1. Introduction

The effect of the social security earnings test on the labor supply of elderly workers has been studied extensively, as evidenced by the tremendous volume of published literature on the subject. While some older literature in the United States found a small effect,[1] the majority of recent studies consistently reveal a large discouraging effect of the earnings test on the labor supply, especially with regard to younger workers, precisely the workers at issue in this study. Among those studies in the United States, Friedberg (2000) employed a "bunch analysis" (analysis of clustering), which examines the concentration of the earnings distribution below the threshold of the earnings tests, and confirmed that the bunching shifted in response to the revision of the earnings test. Moreover, she explored the aggregate impacts of the kinked budget constraint in the earnings test on the conditional hours worked by older workers and showed that the structural estimation yielded a sizable impact from the removal of the earnings test on workers aged 65 and over. While Gruber and Orszag (2003) found a modest labor supply effect, Haider and Loughran (2008) argued that it is hard to detect the effect when the response to the earnings test in the survey data is obfuscated by measurement errors and labor market rigidities. They found a consistent

[1] Burtless and Moffitt (1985) and Gustman and Steinmeier (1985) are earlier works that conclude that the elimination of the earnings tests would have only minor effects on the labor supply.

and substantial response to the earnings test, especially for younger men, if those elements are taken into account. In addition, of the numerous studies conducted in other countries, the majority also found that the earnings test had a sizable labor supply effect. For example, Disney and Smith (2002) examined the effect of eliminating the earnings test in the United Kingdom, and Baker and Benjamin (1999) on eliminating it in Canada; both studies found that eliminating the earnings test increased the employment and earnings of affected male workers.

In Japan as well, many studies have been conducted on the labor supply effect of the social security earnings test, motivated by a popular view in domestic policy debates that the social security earnings test discourages paid work for the elderly through a high effective tax rate on income. When the sum of monthly labor income and public pension income exceeds a certain exemption threshold, the earnings test reduces payments of pension benefits to the beneficiaries of the Employees' Pension Insurance (EPI; *Kosei Nenkin*) program, which covers about half the pensioners in Japan. The earnings test is often viewed as a "punishment" for the elderly labor supply. Indeed, the majority of studies found that the earnings test had a significant labor supply effect. For example, Ogawa (1998) and Iwamoto (2000) built econometric models to simulate the effect of the earnings test and found a sizable effect on labor

supply. On the other hand, there seems to be a negligible effect from the earnings test rule revision. Using a difference-in-difference estimation, Abe (1998) found little effect from a change in the earnings test rule in the 1989 reform on the labor supply of workers aged 60–64, and Ohtake and Yamaga (2003) showed a limited effect from the 1995 reform.

Those previous studies on the labor supply effects of the social security earnings test employ a bunch analysis, a difference-in-difference estimation, or a structural estimation. These approaches are surely effective but not perfect. One caveat of the bunch analysis is that it is hard to detect the labor supply effect when the response to the earnings test in survey data is obfuscated by measurement errors and labor market rigidities (Haider and Loughran [2008]).[2] A difference-in-difference approach needs an appropriate control group, which is often difficult to find since the age range in a treatment group does not overlap that in a control group. Structural estimations require a variety of assumptions of fundamental parameters, which are not necessarily based on solid empirical research and are likely to influence the magnitude of the labor supply effect.

This paper proposes an alternative approach for analyzing the social security

[2] Saez (2010) employs a sophisticated method to analyze whether taxpayers bunch at the kink points of the US income tax schedule generated by jumps in marginal tax rates.

earnings test in Japan, making use of direct responses to a survey question about labor supply. By doing so, we compare the evidence with that in the previous works and explore whether a subjective response tells a similar story. The *Survey on Employment of the Elderly* (SEE; *Konenreisha Shugyo Jittai Chosa*) compiled by the Japanese government is a large-scale, nationwide survey that contains rich information both on employment status and eligibility for social security programs. The SEE explicitly asks the respondents how the earnings test discouraged them from working. This feature enables us to identify a discouraging effect for those who quit working, and the degree to which they circumvent a higher effective tax on labor income for incumbent workers and distill the labor supply effect of the earnings test for workers aged 60–64.[3]

The advantages to examining direct survey responses to the earnings test are summarized in several aspects by which we aim to contribute to the literature in an innovative way. First, this approach is exempt from misidentification when labor market rigidity exists. A bunch analysis implicitly assumes that a labor market is flexible and that workers are able to adjust their earnings and working hours. However, in reality, some workers may not be able to adjust their working hours to earn just

[3] Direct questions regarding social security earnings test responses are available in other countries. To our knowledge, the 1991 interview performed by Social Security Administration in the United States had questions about why people were working and why they were not working. It also asked those not working what wages they would want to go to work. We are grateful to Howard Iams for this information.

below the threshold of the earnings test due to inflexibility in their work schedule. If this is the case, the earning concentration below the threshold is obscured and the discouraging effect is underestimated. The direct response approach enables us to identify those workers who want to work more hours but earn far below the threshold under the earnings test, which is disregarded in a bunch analysis. Second, the direct survey responses enable us to distill the labor supply effect of the earnings test separately from the remaining factors. In most cases, it is difficult to identify whether lower earnings are caused by the earning test or other factors, and we are forced to infer the effect by controlling for factors responsible for earnings using unobserved information, which might be insufficient to derive the effect of interest. The direct survey responses to the earnings test facilitate discerning the effect of the earnings test on labor supply. Finally, if direct responses are informative and can be used to make inferences about the policy effects, it would greatly reduce the costs of policy evaluation.

While we see some notable merits in the use of direct responses to the earnings test, we do not claim that our approach is the best way to evaluate the effect of the earnings test on labor supply. We propose using the direct survey responses to the earnings test as a complement to traditional approaches. We examine whether the gap

in the distribution of wage plus pension benefits, the object of the earnings test, is observed between those reporting the discouraging effect and those not reporting it, even after controlling for the observed characteristics of individuals. We examine the difference in wage distributions using the methodology of DiNardo, Fortin and Lemieux (1996), which enables us to construct a counterfactual distribution as if the observable attributes of both groups were homogenous.[4]

This paper proceeds as follows. The next section briefly describes the earnings rule for workers aged 60–64 in Japan, and Section 3 explains the dataset used in this study. Section 4 presents the results of the direct survey responses in the survey and compares the distribution of wage and pension benefits of both those reporting and those not reporting discouraging effects of the earning test. The last section concludes.

2. Social security earnings test for workers aged 60–64 in Japan

The social security earnings test, which is called the *Zaishoku Rorei Nenkin* program, was first introduced in Japan in 1965 and has been revised every four to six years. The format is different for workers aged between 60 and 64 and those aged 65 and over. We should remember that the earnings test rule is applied only to the

[4] Lemieux (2002) applied this methodology to examine the effect of a change in minimum wage on wage distribution.

beneficiaries of the EPI (*Kosei Nenkin*) program, which applies to employees in the private sector but not to the other public pension programs.[5] The basic rule is that the sum of labor income and the first- (flat-rate component) and second-tier (wage-proportional component) public pension benefit for workers aged 60–64 is reduced when labor income exceeds a certain threshold.

In this study, we confine our interest to EPI pensioners aged 60–64, whose labor supply is presumably the most seriously affected by the earnings test. The earnings test for those workers has multiple thresholds and the reduction rates differ across earnings brackets. In 1989, the object of the earnings test was monthly labor income.[6] In 1995, the object of the earnings test became the sum of monthly labor income and public pension benefits (both first-tier and second-tier). At the same time, the reduction rate was simplified to 20 percent for less than 220,000 yen (about 2,200 US dollars), 60 percent for 220,000 yen–340,000 yen (about 3,400 US dollars), and 100 percent for 340,000 yen and over. In other words, the reduction rate of 20 percent was uniformly applied regardless of the amount of the sum of labor income and public pension

[5] National Pension Insurance (NPI, *Kokumin Nenkin*) for self-employed persons and the Mutual Aid Insurance (MAI, *Kyosai Nenkin*) for employees in the public sector and for private schools.
[6] The reduction rate for workers aged 60–64 was revised in 1989 as follows: 20 percent for less than 95,000 yen, 30 percent for 95,000 yen–114,000 yen, 40 percent for 114,000 yen–138,000 yen, 50 percent for 138,000 yen–165,000 yen, 60 percent for 165,000 yen–185,000 yen, 70 percent for 185,000 yen–210,000 yen, 80 percent for 210,000 yen–230,000 yen, and 100 percent for 230,000 yen and over.

benefits. In 2002 the earnings test rule for those aged 60–64 was revised to raise the threshold to a 60 percent rate for 340,000 yen–370,000 yen, and in 2004 included bonuses as labor income and changed the reduction rate correspondingly (i.e., the threshold for the 20 percent reduction was raised to 280,000 yen). In the 2005 reform, the uniform 20 percent reduction rule was abolished and an eligible person was entitled to receive the full amount of his/her pension benefit if the sum of labor income and public pension benefits was less than 280,000 yen.[7]

3. Data description

This study uses micro-level data from the SEE compiled by the Ministry of Health, Labour and Welfare of the Japanese government. The SEE is performed every four or five years. In this study, we use micro-level data from the SEE surveys complied in 2000, which contain direct responses regarding the earnings test. The individuals in the sample were aged between 55 and 69 and randomly chosen from all regions in Japan. The total sample size was 19,595 in 2000, and the number of persons

[7] The earnings were not tested for workers aged 65 and over after the elimination of the earnings test in 1985. In 2002, the test was revived and the workers who earned more than 370,000 yen from labor income and second-tier pension benefits faced a marginal tax rate of 50 percent (the first-tier pension benefit was not tested for those workers aged 65 and over). In 2007, the earnings test began to be applied to workers aged 70 and over. For those workers aged 65 and over, the first-tier benefit is not tested. See Shimizutani and Oshio (2008) for analysis on the effect of the earnings test on workers aged 65 and over.

aged 60–64 in the sample is 6,692. To our knowledge, the SEE is the only large-scale dataset compiled by the Japanese government and contains a selection of variables that are indispensable to an examination of the labor supply effect of the earnings test, including the individual's demographics and employment status (age, sex, health status, monthly wage, working days per week, hours worked per day, and others) as well as social security eligibility and benefits.[8]

Another feature of the SEE is that it asks the respondents who are beneficiaries of the EPI program explicitly whether the social security earnings test discourages them from working. Since the survey restricts the respondents of the question to those who are eligible for the EPI program, there is no risk of misidentifying other respondents as EPI beneficiaries.[9] The survey asks about the effect of the current earnings test *per se*, which was effective in 2000, on labor supply decisions. We confine our analysis to the respondents of these questions. In addition, we use the respondents who are eligible for Mutual Aid Insurance (MAI), which is exempt from the earnings test, to complement the analysis. Details about the questions will be

[8] There are other large-scale data sets on employment collected by the Japanese government, represented by *Labor Force Survey* [*Rodoryoku Chosa*] or *Basic Survey on Employment Structure* [*Shugyo Kozo Kihon Chosa*], but there is no information on pension eligibility or the direct survey response, which is indispensable to this study.

[9] Some people who are eligible for the EPI program postpone receiving their benefit when they become entitled to it. Because we use information based on actual receipt of the benefit, we are not able to identify persons who are eligible. As a result, we implicitly exclude those who are eligible but do not currently receive any EPI benefits.

provided in later sections.

4. Direct survey response and distribution of wage and pension benefits

In this section, we utilize the respondents' direct responses to the earnings test in the 2000 survey. This survey asked EPI beneficiaries the following questions: "Do you reduce your working hours or days to avoid a reduction in or no receipt of EPI benefits under the social security earnings test?" Each respondent was asked to choose one of three answers: (1) I decided not to work at all; (2) I reduced my working hours or days; or (3) I did not adjust my employment in consideration of the earnings test. The number of respondents eligible for the EPI pension program is 3,555, of whom 2,022 answered these questions in the 2000 survey. We noticed that the number of university graduates who responded to the question was larger than the number of university graduates who did not respond (see Section 4 for a discussion of the bias). Further, we confined our sample to the individuals whose wage and nonwage incomes are available, resulting in a total of 1,790 respondents.

Table 1 reports the labor supply response by males and females in three groups; the "no work group" choosing answer (1), the "restricted group" choosing answer (2), and the "no effect group" choosing answer (3). In the 2000 survey, the SEE asked

respondents to report their monthly labor income (in a 10,000 yen unit integer) earned in September, just one month before the October survey was conducted.[10] It is intuitive to predict that monthly wage income and labor supply would be larger for the "no effect group" and smaller for the "no work group," and indeed this prediction is confirmed by the data. The average monthly wage income is zero yen for the "no work group," and is greatest at 319,600 yen (males) and 148,200 yen (females) for the "no effect group," while the "restricted group" is in between, at 148,500 yen (males) and 93,900 yen (females). We see a reverse pattern in nonwage income: the average nonwage income is 218,000 yen (males) and 74,700 yen (females) for the "no work group," and is 182,600 yen (males) and 73,700 yen (females) for the "restricted group," and is low at 96,500 yen (males) and 50,200 yen (females) for the "no effect group." The sum of wage income and first- and second-tier public pension income, the object of the earnings test, also differs substantially across the groups: 153,300 yen (males) and 58,200 yen (females) for the "no work group," 294,000 yen (males) and 145,900 yen (females) for the "restricted group," and 398,900 yen (males) and 194,500 yen (females) for the "no effect group."

This order is similarly and clearly observed in the labor supply. The average

[10] We believe that the close proximity in time renders the information reliable, but a small number of respondents who did not work to earn a wage reported a non-zero wage. Those respondents were excluded in the analysis.

working hours per day is close to zero for the “no work group,” but it is 6.9 hours (males) and 5.8 hours (females) for the “restricted group,” and 7.9 hours (males) and 6.6 hours (females) for the “no effect group,” implying that the workers in the last category work on a full-time basis. This is also the case for working days per week. The average working days per week is close to zero for the “no work group,” but it is 4.2 days (males) and 4.7 (females) days for the “restricted group,” and 5.2 days (males) and 5.0 days (females) for the “no effect group.” As a result, we see a substantial difference in monthly working hours: 0.1 hours, 29.2 hours, and 41.9 hours for males, and 0.2 hours, 27.4 hours, and 34.5 hours for females for the “no work group,” “restricted group,” and “no effect group,” respectively.

These observations suggest that the earnings test strongly discourages labor supply in a subgroup of the elderly. The sum of wage and pension income of the “no work group” is about 40 percent (males) and 30 percent (females) of that of the “no effect group,” and three quarters of that of the “restricted group.” This difference is also confirmed by comparing the figures of the “no effect group” with those of MAI beneficiaries who are exempt from the earnings test. The table shows that the sum of wage and pension benefits is comparable between the “no effect group” and MAI beneficiaries for both sexes, confirming that the sum of the wage and pension benefits

is substantially smaller for the “no work group” or “restricted group.” Moreover, what we should notice is that the average of the sum of wage and pension income for the “restricted group” is beyond the lower threshold of the earnings test (220,000 yen), and that for the “no effect group” it is beyond the higher threshold (340,000 yen) for males. In contrast, the average of the value for females is below the lower threshold (220,000 yen), even for the “no effect group.”

However, we need to examine not only the average but also the distribution to capture the effect of the earnings test on labor supply. Figure 1 illustrates the distribution of the sum of monthly wage income and first- and second-tier pension income for each group by sex. The range in the X axis is from zero to 800,000 yen, and each cell stands for each 10,000 yen bracket, and the Y axis shows shares in each bin in terms of percentage.[11] If the social security earnings tests restrain labor supply, we would observe a bunch below the thresholds (220,000 yen and 340,000 yen), not for the “no effect group,” but for the remaining groups. Note that the first- and second-tier benefits are computed as full benefits without the earnings test before they were reduced, which is available in the 2000 SEE survey, in order to examine the behavioral responses of the labor supply to the earnings test.

[11] We set the upper limit at 0.8 million yen to make the scale of the figures the same. The proportion of individuals whose sum of wage and pension benefits is greater than 0.8 million and thus who do not appear in the figures is very small.

We observe clearly that the whole distribution is located more to the left-hand side (close to zero) for the "no work group" and more to right-hand side for the "no effect group," implying that the wage plus pension income is lower for that group, which is consistent with the discussion on the average reported in Table 1. In the case of males in the "no work group," a closer look suggests a small bunch below 220,000 yen, while the share between 200,000 yen and 220,000 yen is 21 percent; the corresponding shares in the "restricted group" and "no effect group" are 12 percent and 7 percent, respectively. It seems that there is no bunch below 340,000 yen for any of the groups in the figure. In the case of females, it appears that there is no bunch for any group, and the sum of wage and pension benefits is mostly concentrated below 220,000 yen, even for the "no effect group." The location of the distribution for each group is also confirmed by the reduction rate, which is defined as the share of the amount reduced under the earnings test out of the pension benefits without the test. We also see a difference among the groups. The average reduction rate is zero for the "no work group," 42 percent (males) and 35 percent (females) for the "restricted group," and 60 percent (males) and 45 percent (females) for the "no effect group."[12]

[12] We need to mention two sources of reporting errors in the analysis. First, some respondents reported the values of income, either wage or nonwage, and public pension benefits in round numbers, which is evident in some vertical bars standing out in multiples of five in Figure 1. Second, some respondents may not know the precise full amount of their benefits (the question disappeared in the next SEE survey). Those possible measurement error problems may disturb a

These findings in Table 1 and Figure 1 provide two important observations. First, the figures of the workers who reported being undiscouraged by the earnings test ("no effect group") indeed support their response that they are not concerned about the threshold of the earnings test. They earn above the threshold and the reduction rate is high, and we do not see any bunches on the distribution of the sum of wage and pension income. Second, the wage and pension income and working hours are much lower in both the average and distribution for the "no work group" and "restricted group," a very important observation relating to stimulus of the elderly labor supply. Of course, we should notice a possibility that no bunch below the thresholds for the "no work group" or "restricted group," except for the "no work group" for males below the lower threshold, implies that workers who express concern about the earnings test are not able to adjust their wages close to the threshold for certain reasons such as rigid working hours (see the discussion below).

However, it is too early to conclude that the social security earnings test strongly discourages half of the individuals aged between 60 and 64, since those people are inherently more likely to respond that their labor supply decision is discouraged by the earnings test. If this is the case, the discouraging effect of the earnings test would be

---

bunch analysis.

clearly overestimated. In what follows, we do two things to mitigate this issue. First, we exclude the "no work group" in the following analyses since it is possible that they did not honestly reveal their actual labor supply decision but actually disguised it by reporting the earnings test as a reason for not working to justify their inactivity.[13] Second, even if we assume that the remaining groups—i.e., the "restricted group" and "no effect group"—revealed their actual labor supply decision, we need to explore whether the difference in the distribution of wage plus first- and second-tier pension benefits is observed between those reporting the discouraging effect and those not reporting it, even after controlling for the observed characteristics of the workers. If the difference in the distribution is still observed, the earnings test is more likely to be the reason for the discouraged labor supply.

Indeed, Table 2 shows that the two groups are not homogenous in terms of the direct response to the earnings test, and that they have different characteristics. The proportion of the "restricted group" relative to the "no effect group" is about 40 percent for males and 30 percent for females. The difference in wage income, nonwage

[13] We appreciate the suggestion by the anonymous referees for dealing with this issue. The 2000 SEE contains a question regarding the reason for not working for those respondents who were not working at the time of the survey. Most of them chose either an inability to find an appropriate job or health problems. We do not think that this result simply invalidates the question on the earnings test, because the available choices for the reason for not working did not include the social security earnings test, and a respondent was asked to select only one choice, even if he/she had multiple reasons, which might include the earnings test. However, it is still possible that some respondents chose "I decided not to work at all" due to a reason other than the earnings test.

income, and the sum of wage and full pension benefit is reviewed in Table 1. The respondents' age is younger in the "no effect group" for males. The "no effect group" for males has a higher educational attainment than the same group for females, as is clearly seen in the share of university graduates for males and of high school graduates for females. The difference is also observed in the self-report health status, which is worse in the "restricted group" for males and in the "no effect group" for females. A clearer pattern is observed for the subjective possibility in terms of physical status to work for males: the share of those who are physically able to work is only 47 percent for the "restricted group" but 75 percent for the "no effect group." One half of the individuals in the "restricted group" responded that they are able to work but that work conditions were a factor. Interestingly, the difference between the "restricted group" and the "no effect group" is smaller for females. We do not observe much difference in family size. Lastly, we review the experience of mandatory retirement and job type and firm size prior to mandatory retirement, data which is available in the SEE. The share of those who experienced mandatory retirement is larger for the "restricted group" than the "no effect group" for both sexes, and the share is smaller for females than males. Looking at job type, the dominant share for males is observed in production workers, transportation and communication, and management for the "restricted group," but the

relative share of production workers is smaller for the "no effect group." Job type is more homogenous for females, though the share of administration is larger for the "restricted group." The dominant firm size is 1,000 or more employees for both groups of males and 30–99 employees for both groups of females.[14]

Thus, we should explore the possibility that the difference in the distribution of wage plus pension income is still observed even if the observable characteristics were homogenous between the groups. Concretely, we examine the difference in wage plus pension income distribution using the methodology of DiNardo, Fortin and Lemieux (1996), which enables us to construct a counterfactual distribution as if the observable attributes of both groups were homogenous. This is a semiparametric approach and visually decomposes the change in wage distribution into two parts: the change in the distribution of the attributes and the change in the effect of the attributes on the distribution of wage and pension income.

Concretely, the decomposition approach is described as follows. Denoting *i* as a group, the distributions of wage and pension income for the "restricted group" and the "no effect group" are written as

[14] The information on job type and firm size prior to mandatory retirement includes (1) those at the time of mandatory retirement, and (2) those at the time of leaving a company before the mandatory retirement age. As a result, the sum of the shares of job type or firm size is larger than the share of experiencing mandatory retirement. See Shimizutani and Oshio (2010).

$$f^{restricted}(Y) = \int f^{restricted}(Y \mid X)h(X \mid i = restricted)dX,$$

$$f^{noeffect}(Y) = \int f^{noeffect}(Y \mid X)h(X \mid i = noeffect)dX,$$

respectively, where $f^{restricted}(Y \mid X)$ is the mechanism to determine the wage and pension income of the "restricted group" that maps workers' characteristics *X* to the wage distribution *Y* and $f^{noeffect}(Y \mid X)$ is that of the "no effect group." As *X*, we use the information on the characteristics of workers reported in Table 2, which includes age, educational attainment, subjective health status and physical status, and family size as well as experience of mandatory retirement and job types and firm size prior to mandatory retirement. Moreover, what the distribution of the sum of wage and pension income would be if the distribution of *X* of the "restricted group" is that of the "no effect group" is written as

$$f^{noeffect}_{restricted}(Y) = \int f^{noeffect}(Y \mid X)h(X \mid i = restricted)dX.$$

The DiNardo, Fortin and Lemieux approach employs a "re-weighting" method to estimate the counterfactual distribution. The counterfactual distribution can be

rewritten as:

$$f^{noeffect}_{restricted}(Y) = \int f^{noeffect}(Y \mid X)h(X \mid i = restricted)dX$$

$$= \int \omega f^{noeffect}(Y \mid X)h(X \mid i = noeffect)dX ,$$

where $\omega \equiv \frac{h(X \mid i = restricted)}{h(X \mid i = noeffect)}$. The Bayesian rule produces $\omega = \frac{P(i = restricted \mid X)}{P(i = noeffect \mid X)} \frac{P(i = noeffect)}{P(i = restricted)}$ where the conditional probabilities, $P(i = restricted \mid X)$ and $P(i = noeffect \mid X)$ are propensity scores for the specific observations for those who restricted their labor supply and those who are not affected, respectively, conditioned on *X*, which are calculated by the logit model in this analysis. The terms $P(i = restricted)$ and $P(i = noeffect)$ are calculated based on the proportion of the observations for each group. The counterfactual distribution is computed by the kernel density estimation, using calculated weight ω. The kernel density is also useful when addressing measurement errors, which is the case in this study. To make the results comparable with those from the histogram analysis, we take the level of wage and pension income as the dependent variable, but the results are unchanged when we take the logarithm of the sum as the dependent variable.

In the analysis, we compare the actual distribution for the “restricted group” with

the counterfactual distribution, which is defined as what the density of wage would have been if the attributes of the workers of the “no effect group” were those of the “restricted group.” If the difference in the observable attributes between the two groups produces a difference in the distributions, the counterfactual distribution will overlap with the distribution of the “restricted group.” If this is not the case, however, and the counterfactual distribution still remains to the right of that of the “restricted group,” the discouraging effect of the earnings test between the two groups would not be explained by the difference in the characteristics of the two groups, which is observable.

Figure 2 reports the actual distributions of the sum of wage and pension income of those two groups and the counterfactual distribution as if the observed attributes of the individuals in the “no effect group” were the same as those of the “restricted group.” First, we compare the actual distributions of the groups. For males, when comparing the actual distributions in the two groups, we see that the “peaks” in the distribution of the “restricted group” and the “no effect group” are located at a similar point, but the peak is much smaller for the “no effect group.” In contrast, more concentration is observed for the “no effect group” at the points exceeding 340,000 yen, the higher threshold. For females, the peak of the “restricted group” is located to the left of that of the “no effect group,” which conforms to the histogram analysis in

Figure 1.

What is interesting is the comparison between the actual distribution of the "restricted group" and the counterfactual distribution of the "no effect group," both of which are based on the same observable attributes of the "restricted group." We still see the same observation, since the counterfactual distribution largely overlaps with the actual distribution of the "no effect group." Taking a closer look, in the case of males, we observe that the "peaks" of the actual distribution of "the restricted group" and of the counterfactual distribution of the "no effect group" are located at a similar point, around 220,000 yen, the lower threshold of the earnings test, with a lower peak for the "no effect group." In contrast, a denser concentration is found for the "no effect group" exceeding 340,000 yen, the higher threshold. In the case of females, we see that the counterfactual distribution is located to the right of that of the "restricted group." Interestingly, the peak of the counterfactual distribution of the "no effect group" is located near 220,000 yen, the lower threshold, and the peak of the actual distribution of the "restricted group" is located at a point much lower than the threshold.

These findings show that the difference in the distributions of the two groups does not come from the difference in the observable characteristics of the individuals in the sample. We confirm that the discouraging effect of the earnings test is clearly

observed for a subgroup, and the difference in the distribution of the affected and unaffected groups does not stem from the different characteristics of the individuals in each group. Even if the observable characteristics are homogenous between the two groups, the distribution of the “no effect group” is unchanged and the gap between the two groups remains. These observations reinforce the findings on the discouraging effect of the earnings test on the labor supply for the individuals reporting that they were discouraged by the test. Moreover, even when we do not find a bunch below a threshold of the earnings test, we still find a large, discouraging effect of the earnings test, which has been largely ignored by the previous work on the topic. This finding is consistent with previous works, which found a sizable effect of the earning test on labor supply, including Japanese studies using the structural estimation approach (i.e., Ogawa [1998] and Iwamoto [2000]).

Lastly, we have several reservations regarding the possibility that the labor supply effect is biased. The discouraging effect of the earnings test is underestimated, since some people in the “no work group,” which is excluded in the DFL analysis, are potentially influenced by the earnings test to the greatest extent. As discussed above, this group is excluded because some people are not working due to other reasons than the earnings test, which is clearly evident for the unemployed people in the “no work

group,” who are not discouraged by the earnings test at all in the sense that they are willing to work and thus should be included in the “no effect group.” Unfortunately, the SEE does not allow us to identify the unemployed.

The remaining factors are the sources of overestimation. First, we are forced to limit the sample to those who responded to the question about the effect of the earnings test. As mentioned in the previous section, the proportion of university graduates who responded to that question is larger than the proportion of university graduates who did not respond to the question. Since highly educated people are more likely to understand the structure of the earnings test well, it is possible that their labor supply may be more sensitive to it. Second, we confined our analysis to the “restricted group” and the “no effect group,” since it is possible that some respondents in the “no work group” used the earnings test as a disguising reason. We are not able to completely exclude that possibility in the “restricted group” in the same way. If this is the case, the estimated effect would be overestimated in this study. Third, we aimed to distill the labor supply effect of the earnings test, but the revealed labor supply may be affected by the demand side. The 2009 SEE, which was conducted by the Japan Institute for Labour Policy and Training, shows that, among those workers who reduced their working hours/days due to the earnings test, 28 percent of males and 20 percent of

females responded that they did so according to their employers' instructions. Unfortunately, we are not able to evaluate the demand side effect in the 2000 SEE, but we admit that there is a possibility that it exists.

5. Concluding remarks

The large volume of literature has not reached a consensus on the labor supply effect of the social security earnings test for the elderly. We propose an alternative approach of using direct responses to a survey on the earnings test, and provide new evidence on the sensitivity of the labor supply decision for workers aged 60–64 with respect to the earnings test. We take advantage of micro-level data from the 2000 nationwide survey on the employment of the elderly to examine the change in the labor supply effect for those aged 60–64. Our analysis shows a discouraging effect, even after correcting for the observable attributes among those individuals who reported being affected and being unaffected.

In this study, we argue that a direct response to the question of the labor supply effect would complement the traditional methodology used to examine the labor supply effect of the earnings test. We show that making use of direct responses is informative for evaluating policy effects, and that the methodology produces similar results to

previous studies. A more proper wording of the questionnaire would allow us to identify the discouraging effect, especially in the “no work group.” Further study should examine the effect of the social security earnings test, considering other important factors influencing the labor supply decision, such as health status, family relationships, and labor-leisure choices, as well as the retirement decision of the elderly. Moreover, quantifying the discouraging effect would be indispensable for policymaking. The findings in this study suggest that the abolition of the uniform 20 percent reduction rate in 2005 may have stimulated the labor supply, especially for females, since the density of the distribution is more concentrated under 220,000 yen for females than for males. On the other hand, Shimizutani and Oshio (2008) shows that the repeal of the earnings test in 1985 for those aged 65–69 and the abolishment of the 20 percent reduction rule did not affect the labor supply of the elderly. Further in-depth analysis on labor market conditions for workers in their 60s may clarify future policymaking.

References

Abe, Y. (1998) “1980–1990 Nendai no Dansei Koureisha no Roudou Kyokyu to Zaishoku Rorei Nenkin Seido” [Labor Supply of the Elderly Male and the Earnings-Tested Pension in the 1980s and 1990s]. *Nihon Keizai Kenkyu* [*JCER Economic Journal*] Vol. 36, pp. 50–82.

Baker, M. and D. Benjamin (1999) “How Do Retirement Tests Affect the Labour Supply of Older Men?” *Journal of Public Economics* Vol.71, No.1, pp. 27–52.

Burtless, G. and R. A. Moffitt (1985) “The Joint Choice of Retirement Age and Post-Retirement Hours of Work.” *Journal of Labor Economics* Vol.3, pp. 209–236.

DiNardo, J (2002) “Propensity Score Reweighting and Changes in Wage Distributions” (mimeograph). Ann Arbor, MI: University of Michigan.

DiNardo, J., N. Fortin and T. Lemieux (1996) “Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach.” *Econometrica* Vol.65, No.5, pp.1001–1046.

Disney, R. and S. Smith (2002) “The Labour Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom.” *Economic Journal* Vol. 112 (Issue. 478), pp. C136–C152.

Friedberg, L. (2000) “The Labor Supply Effects of the Social Security Earnings Test.” *Review of Economics and Statistics* Vol. 82, No.1, pp.48–63.

Gruber, J. and P. Orszag (2003) “Does the Social Security Earnings Test Affect Labor Supply and Benefits Receipt?” *National Tax Journal* Vol.56, No.4, pp. 755–773.

Gustman, A. L. and T. L. Steinmeier (1985) “The 1983 Social Security Reforms and Labor Supply Adjustments of Older Individuals in the Long Run.” *Journal of Labor Economics* Vol.3, pp. 237–253.

Haider, S. and D. Loughran (2008) “The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data.” *Journal of Human Resources* Vol.43, No.1, pp. 57–87.

Iwamoto, Y. (2000) “Zaishoku Rorei Nenkin Seido to Koureisha no Shugyo Kodo” [The Social Security Earnings Test and Labor Supply of the Elderly]. *Kikan Shakai Hosho Kenkyu* [*Quarterly of Social Security Research*] Vol.35, No.4, pp. 364–376.

Lemieux, T. (2002) “Decomposing Changes in Wage Distributions: A Unified Approach.” *Canadian Journal of Economics* Vol.35, No.4, pp. 646–688.

Ogawa, H (1998) “Nenkin ga Koureisha no Shugyo Kodo ni Ataeru Eikyo ni tsuite” [On the Impact of the Public Pension on Labor Supply of the Elderly]. *Keizai*

Kenkyu [*Economic Review*] Vol.49, No.3, pp.245–258.

Ohtake, F. and H. Yamaga (2003) “Zaishoku Rourei Nenkin Seido to Dansei Koureisha no Roudou Kyoukyu” [Social Security Earnings Test and the Labor Supply of Elderly Males]. Edited by the National Institute of Population and Social Security Research. *Sentaku no Jidai no Shakai Hosho*, Tokyo: University of Tokyo Press.

Saez, E. (2010) “Do Taxpayers Bunch at Kink Points?” *American Economic Journal: Economic Policy* Vol. 2, No.3, pp. 180–212.

Shimizutani, S. and T. Oshio (2008) “The Labor Supply Effect of the Social Security Earnings Test Revisited: New Evidence from Its Elimination and Revival in Japan.” Discussion paper No. 410, Project on Intergenerational Equity and Center for Intergenerational Studies (CIS), Institute of Economic Research, Hitotsubashi University, Tokyo.

Shimizutani, S. and T. Oshio (2010) “New Evidence on Initial Transition from Career Job to Retirement in Japan.” *Industrial Relations* Vol.49, No.2, pp. 248–274.

Shimizutani, S. and I. Yokoyama (2009) “Has Japan’s Long-Term Employment Practice Survived? Developments Since the 1990s.” *Industrial and Labor Relations Review* Vol.62, pp. 311–324.

Table 1 Direct response to the earnings test in the 2000 survey

(1) Male

	No work group	Restricted group	No effect group	MAI beneficiaries
Wage income	0.00	14.85	31.96	15.03
(ten thousand yen)	(0.00)	(13.96)	(28.74)	(21.58)
Nonwage income	21.84	18.26	9.65	22.17
(ten thousand yen)	(9.57)	(9.09)	(11.68)	(6.68)
Wage + pension income	15.33	29.40	39.89	35.78
(ten thousand yen)	(8.13)	(15.09)	(27.60)	(20.70)
Working hours per day	0.06	6.89	7.89	4.65
	(0.73)	(2.02)	(1.87)	(3.68)
Working days per week	0.01	4.17	5.22	3.09
	(0.18)	(1.38)	(1.22)	(2.50)
Monthly working hours	0.13	29.20	41.93	22.31
	(1.63)	(14.23)	(14.57)	(20.46)
Number of observations	275	247	645	229

(2) Female

	No work group	Restricted group	No effect group	MAI beneficiaries
Wage income	0.00	9.39	14.82	3.85
(ten thousand yen)	(0.00)	(5.68)	(14.92)	(7.55)
Nonwage income	7.47	7.37	5.02	17.05
(ten thousand yen)	(6.37)	(5.91)	(4.98)	(5.89)
Wage + pension income	5.82	14.59	19.45	20.41
(ten thousand yen)	(5.39)	(6.34)	(15.07)	(9.23)
Working hours per day	0.04	5.84	6.59	1.92
	(0.43)	(2.06)	(2.34)	(3.22)
Working days per week	0.03	4.69	5.04	1.31
	(0.40)	(1.38)	(1.30)	(2.14)
Monthly working hours	0.16	27.41	34.47	8.68
	(1.99)	(13.33)	(16.43)	(15.99)
Number of observations	265	75	247	92

Note: The figures in parentheses are the standard deviation.

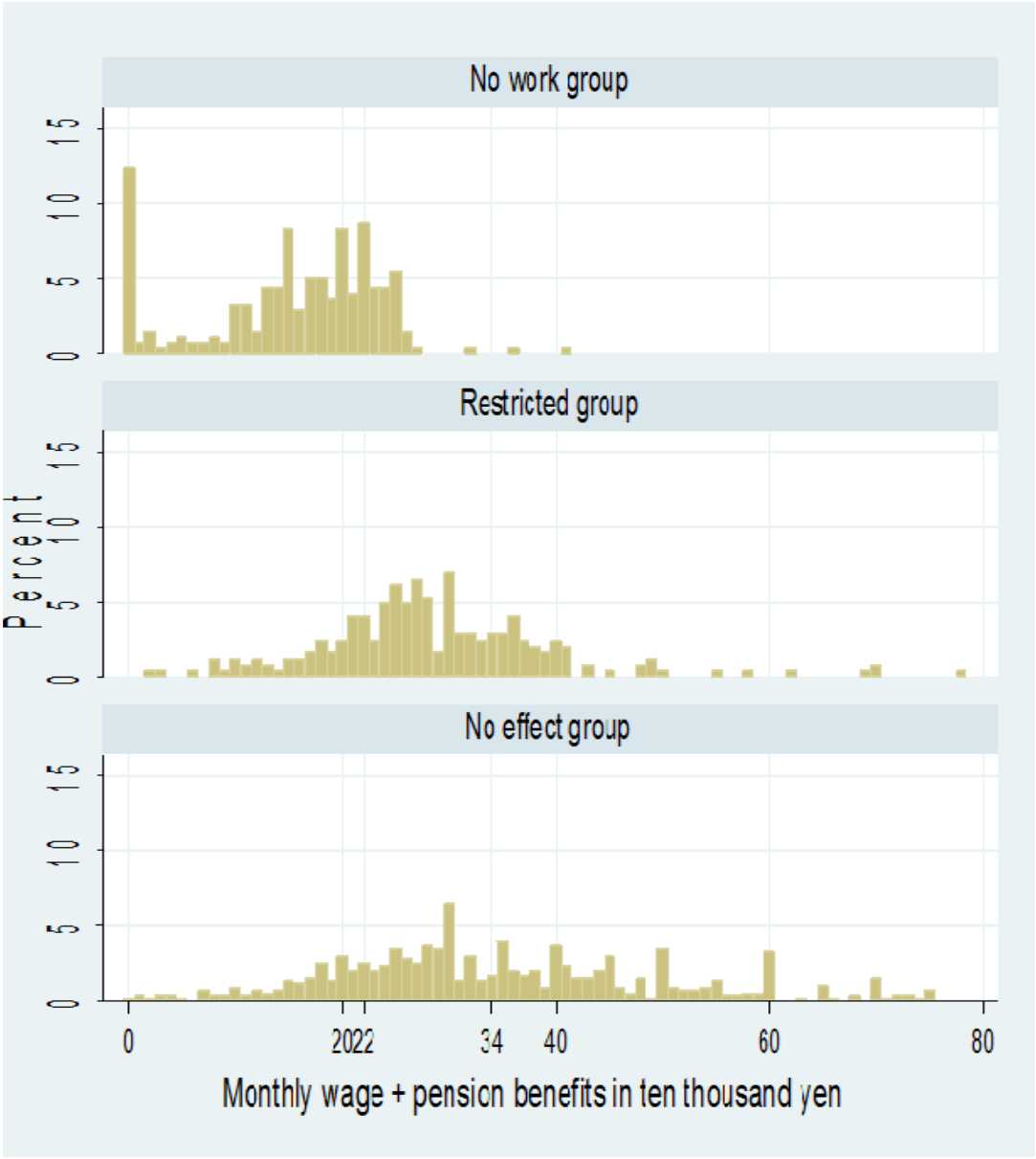
Table 2 Summary statistics in the 2000 survey

	Male				Female			
	Restricted group		No effect group		Restricted group		No effect group	
	mean	S.D.	mean	S.D.	mean	S.D.	mean	S.D.
Wage income	14.854	13.956	31.964	28.744	9.387	5.683	14.822	14.925
Nonwage income	18.259	9.087	9.647	11.680	7.373	5.909	5.020	4.981
Sum of wage and full pension income	29.405	15.087	39.888	27.595	14.587	6.340	19.449	15.070
Age 60 (dummy variable)*	0.166	0.373	0.212	0.409	0.253	0.438	0.243	0.430
Age 61 (dummy variable)*	0.190	0.393	0.208	0.406	0.213	0.412	0.150	0.358
Age 62 (dummy variable)*	0.235	0.425	0.186	0.389	0.200	0.403	0.247	0.432
Age 63 (dummy variable)*	0.215	0.411	0.222	0.416	0.240	0.430	0.198	0.400
Age 64 (dummy variable)*	0.194	0.396	0.172	0.378	0.093	0.293	0.162	0.369
Educational attainment (junior high school)*	0.393	0.489	0.326	0.469	0.533	0.502	0.421	0.495
Educational attainment (senior high school or two-year college)*	0.466	0.500	0.442	0.497	0.440	0.500	0.559	0.498
Educational attainment (university)*	0.138	0.345	0.229	0.421	0.013	0.115	0.016	0.126
Health status (healthy)*	0.838	0.369	0.864	0.344	0.840	0.369	0.834	0.373
Health status (not healthy)*	0.146	0.354	0.116	0.321	0.160	0.369	0.134	0.341
Health status (sick)*	0.016	0.126	0.020	0.141	0.000	0.000	0.032	0.177
Physical status (possible to work on a full-time basis)	0.466	0.500	0.753	0.431	0.413	0.496	0.466	0.500
Physical status (possible to work depending on work conditions)	0.530	0.500	0.245	0.430	0.547	0.501	0.514	0.501
Physical status (impossible to work)	0.004	0.064	0.002	0.039	0.040	0.197	0.020	0.141
Family size	2.854	1.142	2.918	1.385	2.662	1.599	2.813	1.554
Experience of mandatory retirement*	0.713	0.453	0.443	0.497	0.400	0.493	0.239	0.427
Job type prior to mandatory retirement*								
expert or technical	0.049	0.215	0.054	0.227	0.067	0.251	0.020	0.141
management	0.142	0.349	0.130	0.337	0.013	0.115	0.008	0.090
administration	0.089	0.285	0.054	0.227	0.173	0.381	0.069	0.254
sales	0.049	0.215	0.056	0.230	0.027	0.162	0.028	0.166
services	0.004	0.064	0.019	0.135	0.067	0.251	0.061	0.239
security guard	0.020	0.141	0.022	0.146	0.000	0.000	0.000	0.000
transportation and communication	0.146	0.354	0.056	0.230	0.013	0.115	0.000	0.000
production workers	0.300	0.459	0.122	0.328	0.187	0.392	0.150	0.358
agriculture, forestry, and fishery	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.064
Firm size prior to mandatory retirement*								
1-4 employees	0.000	0.000	0.006	0.079	0.013	0.115	0.028	0.166
5-29 employees	0.134	0.341	0.053	0.224	0.093	0.293	0.057	0.232
30-99 employees	0.134	0.341	0.068	0.252	0.147	0.356	0.126	0.332
100-299 employees	0.085	0.279	0.099	0.299	0.040	0.197	0.065	0.247
300-999 employees	0.105	0.308	0.070	0.255	0.120	0.327	0.020	0.141
1,000 or more employees	0.296	0.457	0.163	0.369	0.093	0.293	0.040	0.197
government employees	0.045	0.207	0.053	0.224	0.040	0.197	0.008	0.090
Number of observations	247		645		75		247	

Note: * refers to a dummy variable. The summary statistics for each prefecture dummy (47 dummies) are omitted.

Figure 1 Distribution of wage plus pension benefits

(A) Male



(B) Female

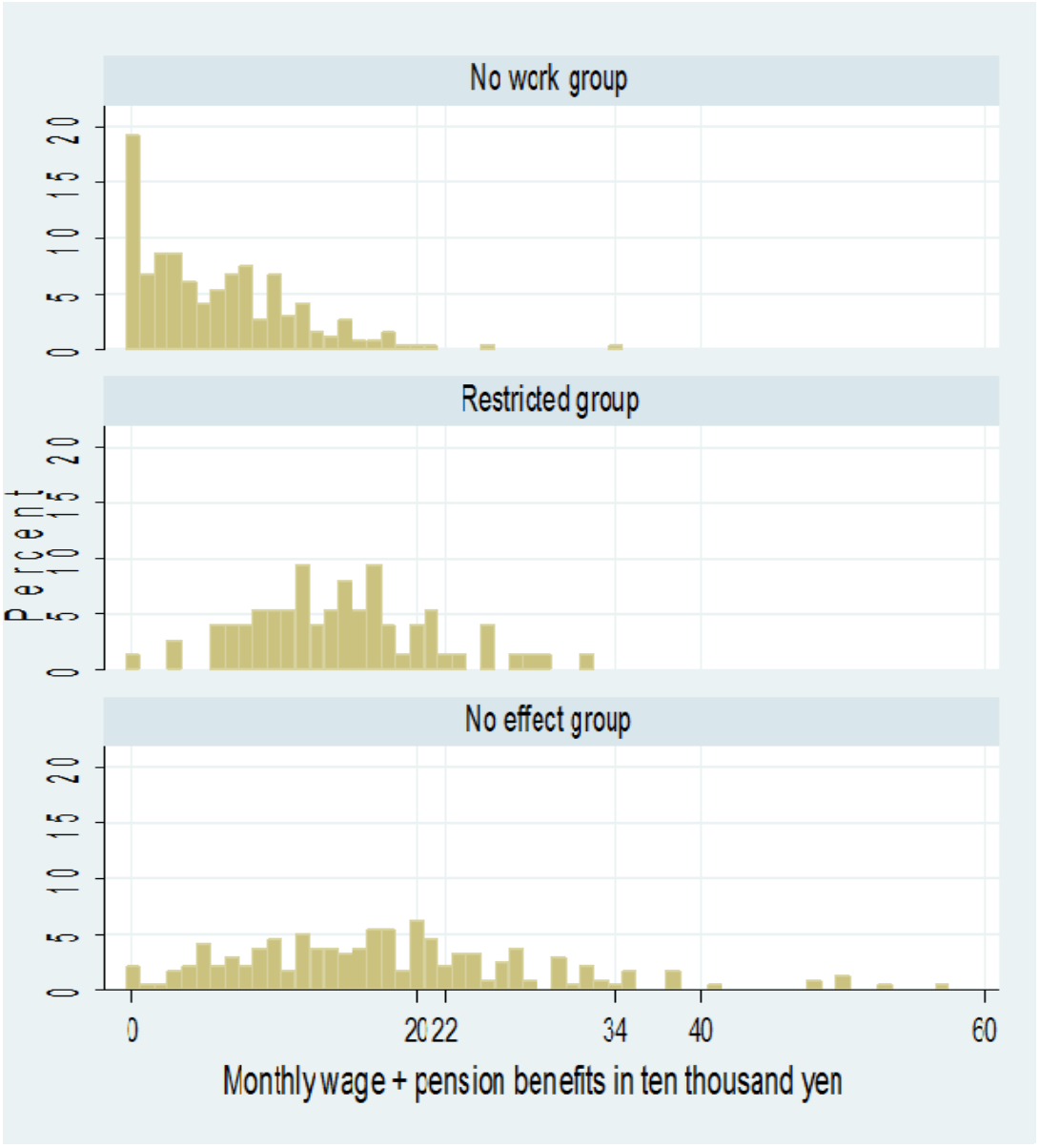
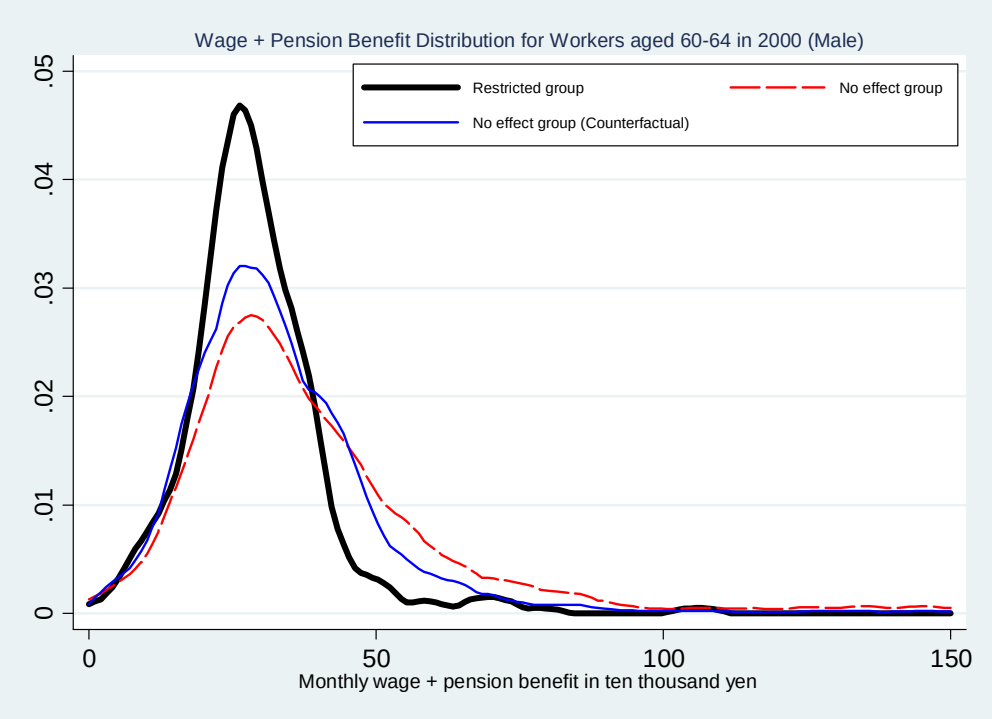

No work group
Restricted group
No effect group
Percent
0
5
10
15
20
0
20
22
34
40
60
Monthly wage + pension benefits in ten thousand yen

Figure 2 Decomposition analysis of wage and pension benefits

(A) Male



(B) Female