ESRI Discussion Paper Series No.381

Social security earnings test and labour supply: Evidence from Japan

Naoki Motohashi

June 2023

Economic and Social Research Institute
Cabinet Office
Tokyo, Japan

Social security earnings test and labour supply: Evidence from Japan*

Naoki Motohashi

Abstract

The social security earnings test reduces pension benefits if the individual's income exceeds a given threshold. Does it reduce the labour supply of older workers? The literature analyses the impact on employment, but little research investigates the influence on working hours in Japan. I analyse the impact of the Japanese social security earnings test by studying two pension reforms in the 2000s. The results suggest four points: (1) for male workers aged 60-64, there is a positive impact on the employment rate, the proportion of full-time workers, and working hours, (2) a non-robust impact on female behaviour, (3) no robust effect on those aged over 65, and (4) the impact is stronger for low-income households and non-regular workers. These findings suggest that policy makers should consider these heterogeneous impacts and reform the system to minimize the deadweight loss in the future.

JEL Classification Numbers: H55, J08, J14, J21, J26
Key words: Earnings test; Employment; Japan

* I thank Kai Liu and Masayuki Yagasaki for helpful comments. The views expressed in this study are those of the author and do not necessarily reflect the position of the Government of Japan.

1. Introduction

How can labour policy promote the employment of older people, and reduce the benefits for the financial stability of the pensions at the same time? These conflicting policy objectives have been a controversial issue for policy makers for decades in Japan. The social security earnings test reduces pension benefits if the individual's income exceeds a given threshold. While many OECD countries have eliminated the earnings test over the last 30 years, the Japanese earnings test still exists. In April 2022, the threshold for 60-64 year-olds has been raised to the same level as over 65s with a view to promoting labour participation among older people. However, policy makers should take note of the impact of this system because it is well known that this test affects the behaviour of older workers and can create substantial deadweight loss (Friedberg 1998).

There are two major arguments relating to earnings tests which are still controversial in Japan. One is that high implicit tax is a work disincentive. Another is that the social security system requires more people to contribute towards their pensions from the perspective of fiscal sustainability. Since an earning-test system leads to very high marginal tax rates for older workers (Baker and Benjamin 1999), some people claim that this system should be eliminated in order to facilitate old people's labour participant in an ageing society. More recently, there has been a profound increase in interest in how to facilitate older people to work to keep the labour supply in Japan. On the other hand, however, it is also a policy issue to increase the number of contributors because the unfunded pension finance has been deteriorating due to the low birth rate and longevity. Others also insist that affluent older workers should give up some of their benefits and contribute to the pension[1][2].

It is important for policy makers to understand how the earnings test affects labour supply for two reasons. First, the earnings test still covers individuals aged over 60. Therefore, understanding how labour supply responds to the earnings test will be helpful to evaluate future reform proposals. Second, the earnings test provides us with an opportunity to examine how the elderly respond to taxes on earnings and, more generally, to changes in wages (see Haider and Loughran 2008).

The literature on this topic has been growing in Japan. However, many of the previous studies

[1] It is also argued that pension benefits should be based on past contributions, but the earnings test distorts this principle.
[2] In addition, in an unfunded pension system, elimination of the earnings test also theoretically implies an income transfer from future generations to the affluent older workers of today. From this perspective, the reform of the earnings test causes a huge debate.

have focused only on the change in the employment rate, and few empirical studies have examined the shift of the budget line in pension reform. Therefore, this paper focuses not only on the employment rate but also on hours worked, and analyses the two pension reforms in the 2000s as quasi-experiments. I find four key implications: (1) The reform of the earnings test has a positive impact on the employment rate (especially for regular workers), the proportion of full-time workers working more than 43 hours per week [3], and working hours for males aged 60-64, (2) the impact on women aged 60-64 is not robust, (3) no robust effect on both genders over 65, and (4) the impact seems stronger for low-income households and non-regular workers.

The paper proceeds as follows. First, Section 2 discusses the international and Japanese literature relating to earnings tests' impacts. Second, I present some theoretical background in Section 3. Third, I describe the institutional details of the Japanese earnings-test system in Section 4. After that, Section 5 explains the detailed empirical framework and the data used in this paper, and Section 6 and Section 7 provide the main regression results of the two pension reforms in 2000 and 2004, respectively. Finally, Section 8 covers the conclusions from the study.

2. Literature Review

Since it is controversial whether an earnings test has a negative impact on the employment rate and hours worked and how much the impact is, this section reviews the existing literature concerning the effect on labour supply to clarify the stance of this paper. In general, there are two approaches to estimating the impact of earnings tests. One is to develop structural models (see Friedberg 2000). The other is to utilise reduced-form models such as the difference-in-differences (DID) approach (see Gruber and Orszag 2003, Song 2004, Song and Manchester 2007, Haider and Loughran 2008). This topic has been researched for decades, and the empirical evidence from both reduced-form approaches and structural model approaches is voluminous.

Since earnings tests had been operated in a number of developed countries, a large body of empirical research exists, especially in the United States. Burtless and Moffitt (1985) is one of the first important studies on the subject, which examined a liner budget constraint approach with a simple life cycle model and estimated the effect of the U.S. Social Security system, including the earnings test, on the labour supply of the aged. They showed that the earnings test had a relatively

[3] Though full-time workers are defined as those working more than 40 hours per week, in this paper, I define full-time workers as those who work more than 43 hours due to the data limitations.

small effect on both the age of retirement and on postretirement hours of work in the United States. While Haider and Loughran (2008) show a substantial impact on hours worked for males, a number of studies such as Gruber and Orszag (2003), and Song and Manchester (2007) failed to find a robust influence on labour supply in the United States.

On the other hand, several studies provide significant evidence from outside the United States. Disney and Smith (2002) provide evidence from the U.K. earnings test. They take the abolition of the U.K. earnings test in 1989 as a quasi-experiment, and use the DID method. They consider the unaffected cohort as a control group to control for the potential effect of macro factors, and find that the abolition of the test raised working hours of older male workers by around 4 hours a week. The impact was smaller for females. Baker and Benjamin (1999) examine the sequential elimination of the earnings test in Canada in the 1970s. The outstanding aspect of their research is that they focus on the reform of the two different pension plans: one is the Quebec Pension Plan, which covers workers in the province of Quebec, and the other is the Canada Pension Plan, which covers the other provinces. This unique structure of the pension system enabled them to separate the control group from the treatment group not by age but by geography. Their result indicates a large increase in hours worked and a large shift from part-time to full-time work.

In addition, as Gelber et al. (2021) points out, while several U.S. studies using the DID find little evidence for an effect on the employment rate, some recent papers employing structural models show a meaningful impact. Friedberg and Webb (2009) examine the impact of not only actual but also anticipated earnings test parameters to assess the influence of the past and anticipated future rules on current employment. They predict that the elimination of the U.S. earnings test in 2000 raised employment by around 2 percentage points at ages 66–69 and 3.5 percentage points at age 65.

In conclusion, the impact of earnings tests on labour supply is still controversial and these studies suggest that the labour customs and culture of each country may also affect the behaviour of older workers. Therefore, it is essential to consider both internal validity and external validity when discussing the policy impact of an earnings test reform. For those reasons, Japanese policy should be based on evidence from Japan, and more literature should be accumulated. Japanese policy makers should examine the wide variety of literature to estimate the impact of the Japanese earning test for future reforms in Japan.

After reviewing the international evidence, I move to look at the previous research in Japan.

The literature has also been increasing in Japan from the late 1990s, but it focused mainly on the influence on labour supply and few studies have discussed the impact on working hours. Abe (1998) examined the effect of the reform in 1989, which aimed to facilitate the labour supply of older workers, on males using the micro data by the DID. She found little evidence of encouraging employment, and doubted the policy effect. In the 2000s more literature has been provided utilising the micro data with both reduced-form approach and structural model approach. While Yamada (2003) failed to find evidence of labour decrease, a large literature has shown the impact on older workers aged 60-64 with both structural models (see Iwamoto 2000, Higuchi and Yamamoto 2002, Ishii and Kurosawa 2009, and the Cabinet Office Director General 2018) and reduced form models (see Otake and Yamaga 2003, and Ishii and Kurosawa 2009). Higuchi and Yamamoto (2002) develop a structural model and report that the Japanese earnings test has a negative impact on the labour supply of workers aged 60-64. Ishii and Kurosawa (2009) also employ both a reduced-from model and a structural model, and find an impact on male workers aged 60-64, but fail to show a significant effect on those over 65. Hamada (2008) examines the impact of reform in 2004 and finds that the abolition of a 20% pension cut mitigated employment disincentive effects for 60-64 workers. More recently, the Cabinet Office Director General (2018) utilises panel data from 2005 to 2015 and examines the influence on employment choice using a structural model. While this paper reports a significant impact on early 60s workers, it fails to find any meaningful impact on workers over 65.

In conclusion, I note three key points. First, many of Japanese previous studies point out a significant negative impact of the earnings test on labour supply of workers aged 60-64, while they find little evidence of influence on labour supply for those over 65. Second, although previous literature focuses mainly on employment disincentive effects and changes in the employment rate, there is little literature discussing the influence on hours worked. The third point is that few Japanese studies focus on changes in the budget line caused by reforms. Therefore, in this study I carefully analyse changes in the budget line to predict the theoretical impact of earnings-test reforms. After that, I conduct two case studies to examine the impact on both the 60-64 cohort and the 65-69 cohort, and examine the change in hours using the DID approach.

3. Theoretical Background

This section reviews the background theory of the earnings test[4]. Until recently, many advanced countries had earnings tests. However, these systems have been eliminated in multiple countries over the last 30 years. For instance, the United Kingdom used to have an earnings test in which men aged 65-69 and women aged 60-64 earning above a certain threshold had to pay a high implicit tax, and this system was abolished in 1989.The key idea of an earnings test is simple; if earnings are above a threshold, benefits are taxed away. Therefore, when this system is abolished, two different effects will influence behaviour; the income effect and the substitution effect. “The income effect will motivate the worker to consume more leisure and work fewer hours; the substitution effect induces the worker to consume fewer leisure hours and work more hours” (Borjas 2013: 79).

Figure 3-1 demonstrates a simple diagram of the earnings test. Now we assume a worker who continues to work after the pension age. If he does not work at all and receives the pension, he is at point A. If the sum of his pension and income is lower than the threshold (between points A and B), his pension benefit is not cut and there is no disincentive to work. However, if the sum is above the threshold, his pension is cut and thus the slope of his budget line will be gradual (between points B and C), which implies an implicit tax for working. Finally, he receives no pension benefits if his or her income is too high (between points C and D). Therefore, from a basic labour supply theory, this earnings test has a huge impact on the work decisions of older people because it generates a non-linear static budget constraint with a convex kink. Therefore, as Blundell et al. (2000), and Eissa and Leibman (1996) pointed out, an earning test plays a similar role as the Working Family Tax Credit in the United Kingdom. and the Earned Income Tax Credit in the United States and leads to a substantial disincentive to work.

The abolition of the earnings rule means that the budget constraint turns linear (the dotted line in Figure 3-1), and the marginal tax rate decreases for those who earn above the threshold. If the older people are between points A and B before the reform, their labour supply (working hours) will not be affected by the earnings test. On the other hand, workers between points B and C are affected by this test, and will change their labour supply when the test is eliminated. In this case, it depends on the balance between the income effect and the substitution effect whether the

[4] Disney and Smith (2002) provide a detailed explanation of basics of earnings tests in the United Kingdom and Friedberg (1998) shows an illustration of the U.S. earning test. Furthermore, several textbooks also explain the theoretical details. See Borjas (2013) Chapter2.

earnings test increases hours worked or not. In contrast, workers above C will reduce working hours because only the income effect is in operation, inducing the workers to consume more leisure hours. Therefore, it is also important to examine the impact by income level in this paper.

4. Institutional Details

This section explains the Japanese social security earnings test system (Zaisyokurourei Nenkin Seido) and its history. The Japanese earnings test is a system under which employees' pension payments are partially or fully suspended for those pensioners whose total monthly wages and employees' pension exceed a certain amount. Since older workers who earn more than the threshold have their pensions reduced, the earnings tests are criticised for discouraging work.

There are two opposing views on this system and it is still a controversial issue. One is that it is necessary to ask high-income earners to give up part of their employees' pension and ease the burden of the younger generations, because the financial sustainability of the unfunded pension system has been worsening under the situation of declining birth rate and ageing population in Japan. The other one is that this system distorts the labour supply of older people and there exists strong policy request that more older people should remain in the labour market to keep the labour supply. The policy makers at the time have tried to amend this system to balance these two conflicting claims. Therefore, understanding the history of the Japanese earnings test is important before evaluating the policy impact.

4.1. Earnings test in Japan

The Japanese public pension system was developed in 1954 in much the same form as it is today. Namely, the National Pension (Kokumin Nenkin) is the universal floor and the Employees' Pension Insurance is the second floor. Though it was necessary to be retired to receive the National Pension at that time, there were many older people who worked for low wages and needed to continue to work to afford the cost of living. Under such circumstances, the Japanese earnings test was introduced in 1965 to allow older workers over 65s to receive employees' pension while continuing to work. The amount was 80% of the employees' pension that would have been received if they had retired. The earnings test for over 65s had been abolished in 1985, but it was reintroduced in the 2000s. For people aged 60-64, a 20% pension cut was introduced in 1994 in

order to simplify the system and ensure that the sum of wages and pensions increases gently. This pension cut was applied once people started working, and it generated substantial work disincentives. As Table 4-1 summarises, the Japanese earnings test has been amended several times. In this paper, I take two pension reforms as quasi-experiments. In 2000, the earnings test was reintroduced for 65-69 year olds, and the 20% pension cut was abolished in 2004.

Until March 2022, the threshold level was different according to age[5]. The threshold for 60-64 years old is 280,000 yen (per month), and that for over 65 years old is 470,000 yen. If the sum of the older worker's income and employees' pension is over the threshold, the pension is cut at a rate of 50% of excess wages. Figure 4-1 shows a simple image of the earning system assuming that the pension is equal to 100,000 yen. According to the Ministry of Health, Labour and Welfare, the number of people covered by the scheme is about 670,000 and the total amount of suspended payments is about 480 billion yen for 60-64years old, and that for over 65 years old is 410,000 people and 410 billion yen, respectively[6].

4.2. Reform in 2000

First, I discuss the detail of the reform in 2000, which I analyse as a quasi-experiment. In 2000, the earnings test was reintroduced for workers aged 65-69 in order to reduce the burden on the working population in pension financing. Figure 4-2 shows the change in the budget line for a worker with a employees' pension of 100,000. The red dotted line represents the budget line before the reform and the red solid line is that after the reform. The employees' pension is fully paid until the sum of the wage and employees' pension reaches 370,000. However, if the total amount is over the threshold, the pension is cut at a rate of 50% of excess wages and it creates a convex kink and a non-convex kink in the budget constraint. Therefore, this reform was expected to generate disincentive to work for workers above points A, and decrease their employment rate and working hours.

4.3. Reform in 2004

Next, I explain the reform in 2004. In this reform, the suspension of the pension payments was

[5] From April 2022, the threshold for 60-64s has been raised to the same level as over 65s (470,000 yen). But I explain the system before April 2022 for a clearer understanding of the two reforms in the 2000s.
[6] See the Ministry of Health, Labour and Welfare (2019).

eased for 60-64 workers[7]. Figure 4-3 shows the budget line of an older worker with a employees' pension of 100,000. The red dotted line represents the budget constraint before the reform and the budget line after the reform is the red solid line. Before the reform, the pension was cut by 20% regardless of the wage amount. By abolishing this pension cut, the budget line shifted upwards. In addition, the threshold (280,000 yen) was not changed. Since the slope of the budget line between points A and B, and that between points A' and B' were not changed, only the income effect can affect and reduce the labour supply. Therefore, from this budget line analysis, it is expected that the employment rate will rise because the abolition will increase the incentive for work, but average hours worked may decrease due to the income effect.

5. Empirical Approach

5.1. Empirical Framework

Section 5 shows the detailed empirical framework. Applying the estimation methods from Disney and Smith (2002), Gruber and Orszag (2003), and Song (2004), I analyse the effect on employment and hours worked using the difference-in-differences (DID) approach with repeated cross section data. In the DID approach, I compare mean changes in the employment or working hours of treatment groups before and after the reform and corresponding changes for control groups. As the control group, I use the older cohorts because the younger cohorts may have spill over effects when they reach the age of the treatment group[8].

First, I analyse the effect of earning tests on the employment of older workers. The goal of this estimation is to provide evidence to the following question: whether an earning-tests reform facilitates the employment of older workers? This paper uses the following estimation equation:

$$Y_{it} = \alpha_1 Treatment_i + \alpha_2 POST_t + \delta Treatment_i POST_t + X_{it}\beta + \varepsilon_{it} \quad (1)$$

In (1), the 'i' and 't' subscripts denote the individuals and years, respectively. The dependent variable is the dummy of whether the individual is employed or not. X'_{it} is a vector of individual

[7] This reform came into force in April 2005.
[8] Angrist and Krueger (1999) point out that two conditions must be satisfied to identify the effect of the reform using the DID: the composition of the groups must be stable before and after the reform, and both groups have the same macro trends. The spill over effect can be a threat of the first condition (Disney and Smith 2002).

attributes, including some dummies (education, children, spouse, industry, firm-size, etc). With education dummies, I divided the sample into three-parts, graduate from junior high school, graduate from high school, and graduate from 2-year college, specialized high school, university, or graduate school. The children dummy shows whether an individual has a child under 15 years old. $Treatment_i$ represents the treatment group dummy, and $POST_t$ is the time dummy. δ is my research interest and the average treatment effect (differential change in labour participant).

Second, to clarify the research question: how an earning-tests reform affects the working behaviour of older workers, I analyse the effect of earning tests on the proportion of full-time workers and hours worked. In this paper, I use two dependent variables. One is the dummy of whether the individual works more than 43 hours per week. The other is average hours worked per week.

5.2. Data

In this paper, I used the Employment Status Survey (ESS), for 1997, 2002 and 2007. ESS is cross-section data of approximately 440,000 households (over 15 years old), and it investigates employment and non-employment status on 1st October in order to obtain basic data on the structure of employment, job transfers and wishes regarding employment in Japan. Using data for 3 years, the models in 5.1. are estimated to examine the effects of the earning tests reform on labour supply in Japan. This paper focuses on employed or non-employed people, and excludes self-employed workers, public servants, and executives, because the employees' pension does not cover self-employed workers and public servants, and these workers' decision process for retirement is different from that of employed workers. Finally, since ESS collects only trimmed figure of hours worked (-15, 15-19, 20-21, 22-29, 30-34, 35-42, 43-45, 46-48, 49-59, 60- hours), I calculated the hours worked using the means as a proxy.

6. Estimation Results 1 -2000 Reform-

This section focuses on the Japanese pension reform in 2000 (reintroduction of an earnings test for 65-69). My aim is to answer the following question: does an earnings test affect the behaviour of older workers aged 65-69? I estimate the impact of this old-age pension system for active employees on labour supply by using the DID. The treatment group is people aged 65-69, and the

control group is people aged 70-74. Since younger generation (60-64) were also affected by this reform in the future, I use the older cohort as a control group. In this analysis, I use ESS1997 and ESS2002 as the pre-reform and post-reform data, respectively.

Table 6-1 provides the descriptive statistics for employed and non-employed workers aged 65-74 in 1997 and 2002, respectively. In Section 6 and 7, I use three dependent variables. Employed means whether individuals are employed or not, and I analyse the change in labour supply of older people. Long hours is a dummy variable which is equal to one if the individual works more than 43 hours per week. Hours worked is average working hours per week.

For the characteristics of a household or individual, I use several variables. Urban is equal to 1 if the individuals live in Tokyo, Nagoya, or Osaka area. If an individual has a child under 15, Under15 is 1. Working spouse dummies and Spouse over 60 dummies show whether the spouse works or not and whether the partner is over 60 years old or not, respectively. Nemployed is the number of employed in a household except the individual. As personal characteristics, High school and College means graduating from high school, and graduate from 2-year college, specialized high school, university, or graduate school, respectively. I also employ working information. The regular dummy represents whether he or she is a regular worker[9] or not. The sector dummies explain in which sector an individual works. The large size dummy means that individuals work for a large company (more than 300 employees).

Next, the descriptive statistics of dependent variables for the treatment group (65-69) and the control group (70-74) are shown in Table 6-2. In addition, this table reports the simple difference in differences of changes in the dependent variables. The result for male employment rate is statistically negative and that of hours worked for males is statistically positive, the other results are not robust. These facts suggest that the reform in 2000 may have influenced the male employment and working hours. The former is consistent with the theoretical prediction discussed in Section 4.3. while the latter is not.

6.1. Effect on employment

Table 6-3 shows the effect of the earning-test reform in 2000 on the employment of older workers. Column (1) and (2) show the results by gender with OLS, respectively. The coefficient of the DID

[9] Regular workers mean full-time workers, and non-regular workers include part-time workers and temporal workers.

is my research interest. The coefficients are statistically significant but positive for males. On the other hand, I also use a probit model, Column (3) and (4) report non-robust results. Theoretically, this reform increased disincentive for work, and would decrease the employment rate. However, I cannot find robust evidence of a negative impact on the labour participation rate for those aged 65-69 in my analysis. While this result is not consistent with Disney and Smith (2002), it is consistent with the previous Japanese literature, which also fail to find a robust policy impact for late 60s workers in Japan (see Hamada 2008, Ishii and Kurosawa 2009, and the Cabinet Office Director General 2018).

6.2. Effect on hours worked

Second, to test the research question: whether the earnings test reform affects older workers' working hours, I estimate two types of equation examining the rate of full-time workers, and average working hours per week. Table 6-4 shows the effect of the pension reform on the proportion of full-time workers, and Table 6-5 reveals the impact on the hours worked per week, respectively. As in Section 6.1, I fail to find evidence that the reintroduction of the earnings test changed the hours worked of late 60s.

One possible explanation for these findings is that over 65 years old is beyond the Japanese legal retirement age. The legal retirement age based on the Act on Stabilization of Employment of Elderly Persons has been revised during the last three decades. The legal retirement age was set to 60 from 1998, and it was raised to 65 in 2013. As the Cabinet Office Director General (2018) points out, while many people in Japan continue to work after retirement age, reaching the retirement age is an opportunity to switch employment status. Therefore, over 65 workers are likely to continue to work for non-financial reason[10] and are not affected by the pension reform.

7. Estimation Results 2 -2004 Reform-

Following Section 6, in Section 7, I focus on the Japanese pension reform in 2004 (abolition of 20% pension-cut for 60-64s), which is expected to mitigate disincentive effects for older workers aged 60-64 but may decrease average work hours theoretically. The treatment group is individuals aged 60-64, and the control group is the older cohort, people aged 65-69. I use ESS2002 and

[10] For example, to feel part of a community.

ESS2007 as the pre-reform and post-reform data, respectively.

Table 7-1 represents the descriptive statistics for employed and non-employed workers aged 60-69 in 2002 and 2007, respectively, and the descriptive statistics of dependent variables for each cohort are shown in Table 7-2. For males, all dependent variables show statistically significant changes, while only the employment rate effect is robust for females. These results imply that the impact of this reform could be lager for male workers than female workers.

7.1. Effect on employment

First, Table 7-3 shows the effect of the earnings test reform in 2004 on the employment of older workers. Column (1)-(2) show the results by gender with OLS, respectively. This estimate result shows that the abolition of the 20% pension-cut for 60-64s increased the labour participant rate for male and female workers by approximately 1.2 percentage point. On the other hand, however, Column (3)-(4) present the results with a probit model, but it reports a significantly positive effect for males only[11]. This result is consistent with the previous literature in Japan, which reports a robust effect on male employment rate. In conclusion, this research finds a statistically significant effect on male employment rates for those aged 60-64 but not for those aged over 65, which is consistent with Hamada (2008), Ishii and Kurosawa (2009), and the Cabinet Office Director General (2018).

7.2. Effect on hours worked

Table 7-4 shows the effect on the proportion of full-time workers. Only the coefficients for males (Column (1) and (3)) are statistically significant, indicating that I cannot find evidence that the pension reform affects the full-time worker rate for females. This result is not consistent with Disney and Smith (2002), which finds a positive impact on the proportion of workers who work more than standard working hours (40 hours per week) for both genders. Why were females not affected? One possible reason is a conventional idea of family. Since both the treatment group and control group members were born before and just after World War II, a strong tendency towards a gender role division of labour, so called "men at work and women at home", existed. Therefore,

[11] Appendix A also shows marginal effect. It reports that the 2004 reform increased the male employment rate by approximately 2.6 percentage point.

the financial incentive of the reform might not influence these people's behaviour.

The impacts on the hours worked per week are shown in Table 7-5. While the coefficient of the DID is strongly statistically significant for men (+1.6 hours), we find only weak evidence for women (+0.7 hour). This propensity, a stronger influence on males than on females, is similar to the literature. While I report a smaller change in hour worked than Disney and Smith (2002), which finds a 4.2 hours-increase for males and 2.4 hours for females, I conclude that the influence of the Japanese earnings test on males is not weak because the reform in 2004 was just an amendment and did not eliminate this system[12]. In addition, the results from Tables 7-3, 7-4 and 7-5 suggest that not only the number of full-time workers but also that of short-time workers increased, and short-time workers increased their working hours, implying that working more hours can help lower income earners escape poverty.

On the other hand, however, this result is not consistent with the theoretical prediction discussed in Section 4. 3.. One possible explanation is that low-income earners can work longer until they reach the threshold. Since this reform increased the employment rate, it also raised the number of low-income workers, for example non-regular workers. Therefore, this reform increased the low-income workers whose slope of the budget line is more gradual, and they are like to work more, leading to an increase in average working hours. This counter-intuitive result can also be explained by the increase in regular workers. In general, the average of work hours of regular workers is longer than that of non-regular workers. For instance, the average working hours of male regular workers aged 60-64 is 44.3 while that of non-regular workers is 36.8. Therefore, if this hypothesis is correct and the 2004 reform prevented workers from retiring and kept them in the labour market as regular workers, it would increase in average working hours. In fact, Table 7-1, in fact, tells us that the rate of regular workers is increasing for both genders.

7.3. Further analyses

Finally, I perform further analyses for the 60-64s to answer the following question: which demographic groups are more impacted by the earnings test, because the discussion in Section 3 and Section 4 suggests that the impact of the earnings-test reform can be different according to the characteristics.

[12] Disney and Smith (2002) use the abolition of the earnings test in the United Kingdom.

First, I focus on low-income households. As Borjas (2013) notes, it is possible for high income earners that the income effect dominates the substitution effects and thus they reduce working hours. Moreover, it is necessary for policy makers to examine the behaviour of low-income older people to prevent them from poverty after retirement. Since the average income of the older households is nearly 4 million yen, I test the impact for households whose income is less than 4 million yen. For males, the coefficients in Table 7-6 are stronger and more robust than those in Table 7-3. These results suggest that lower-income households are likely to react more to the earnings test. Therefore, easing the earnings-test system can stimulate low-income households to work more and reduce the risk of poverty.

Next, I examine the impact on working hours for non-regular workers to discuss the hypotheses in Section 7.2.. In addition, it is worth to focus on non-regular workers, because they can decide their working hours more flexibly than regular workers[13]. While Table 7-7 reports an insignificant effect for female non-regular workers, it shows a greater impact for male non-regular workers than the average, implying that male non-regular workers reacted more to the 2004 reform in terms of working hours because it increased working incentive[14]. Furthermore, Appendix C reports the result of employment status with a multinominal logit model. This model suggests that the 2004 reform increased the probability of regular work by 3.4 percentage points. Therefore, these estimation results support the two hypotheses discussed in Section 7.2..

8. Conclusion

The aim of this paper was to empirically evaluate the impact of the earnings test in Japan using the two reforms in the 2000s as quasi-experiments. Specifically, I have explored whether older workers change their working behaviour such as employment rate and hours worked as a response to the change in the earnings-test system. This paper estimated three variations of the models which analysed the impact on the employment rate, the proportion of full-time workers, and average weekly working hours. The results of this study suggest four points: (1) a positive impact on the employment rate (especially for regular workers), the proportion of full-time workers, and working

[13] I also estimate a model for the rate of full-time workers, but only the result for males is weakly statistically positive.
[14] Moreover, Appendix B shows the results of working hours by regularity. Comparing with the results in Table 7-5, Table B indicates larger impact for non-regular workers, while that for regular workers is nonsignificant.

hours (especially for non-regular workers) for males in their early 60s, (2) a non-robust impact on the female employment rate for those aged 60-65, (3) no robust effect on either gender aged over 65, and (4) the impact seems stronger for low-income workers and non-regular workers.

The findings of this paper offer several important practical implications for Japanese policy makers. First, this paper found similar results to the Japanese previous literature: the strongest impacts are on male workers aged 60-64. Since the earnings test for those aged 60-64 continued after the reform in 2022, policy makers should analysis the impact of the remaining system. Second, my findings reveal that low-income households and non-regular workers are affected more by the test, so they will receive the most benefit from easing this system. Therefore, one implication of this study is that the impact of the pension reform is heterogeneous, and policy makers should take note of the heterogeneous impact on older workers in order to reduce the deadweight loss. They should also consider that the elimination of the test can increase the welfare of low-income households and prevent them from poverty after their retirement by facilitating employment. Since the ageing of the population is regarded as one of the causes of widening income inequality in Japan, it is crucial that labour policy for older workers should be amended in order not to discourage low-income people from working, as the solution to eliminating inequality among the elderly.

It is important to note that ESS includes only trimmed figures for working hours and income, so it is difficult to analyse or predict the detailed changes in working hours by income bracket. There is also another way in which earnings tests affect the labour supply of older people. An earnings test might affect the timing of the benefits receipt. Gruber and Orszag (2003) find that the abolition of the earnings test accelerates benefits take-ups among the eligible population. But due to the data limitation, this paper cannot address this point.

In spite of these limitations, I believe that the conclusion can provide several policy implications and the evidence could be applied for EBPM in Japan. The Act on Stabilization of Employment of Elderly Persons was amended and ‘measures to ensure job security', such as extending the retirement age to 70, became obligatory for efforts from April 2021. Additionally, the healthy life expectancy has been extended by more than 2 years from 2001 to 2016[15]. The Ministry of Health, Labour and Welfare (2020) points out that healthy life expectancy is increasing

[15] According to the Ministry of Health, Labour and Welfare (2020), the male healthy life expectancy increased from 69.40 to72.14, and that of female rose from 72.65 to 74.79.

while physical functioning in old age is rejuvenated. These facts imply that more old people will continue to work and be affected by the earnings test in the future. Therefore, further amendment and discussion will be required in the near future. I hope that the results of these analyses will be applied to the future reform of the earning test, and help better policy-making in practice.

References

Abe, Y. (1998) '1980-1990 Nendai No Danseikoureisya No Roudoukyoukyuu To Zaisyokuroureinennkinnseido', *JCER economic journal*, 36, pp.50-82 (In Japanese)

Angrist, J. and Krueger, A. (1999) 'Empirical strategies in labor economics', in Ashenfelter, O. and Card, D. (ed.) *Handbook of Labor Economics*. 3: Elsevier Science B.V., pp. 1277-365

Baker, M. and Benjamin, D. (1999) 'How do retirement tests affect the labour supply of older men?', *Journal of Public Economics*, 71, pp. 27-51

Blundell, R., Duncan, A., McCrae, J. and Meghir, C. (2000) 'The Labour Market Impact of the Working Families' Tax Credit', *Fiscal Studies*, 21, pp.75-104

Borjas, G.J., (2013) *Labor economics*, sixth edition, McGraw-Hill Ser. in Economics

Burtless, G. and Moffitt, R. (1985) 'The joint choice of retirement age and post-retirement hours of work', *Journal of Labor Economics*, 3, pp. 209-36

Cabinet Office Director General (2018) '60-Dai No Roudoukyoukyuu Ha Donoyouni Kimarunoka?', Available at: https://www5.cao.go.jp/keizai3/2018/09seisakukadai16-0.pdf (Accessed 1 July 2022) (In Japanese)

Disney, R. and Smith, S. (2002) 'The Labour Supply Effect of the Abolition of the Earnings Rule for Older Workers in the United Kingdom', *The Economic Journal*, *112*(478), pp.136–152

Eissa, N. and Liebman, J. (1996) 'Labor Supply Response to the Earned Income Tax Credit', *The Quarterly Journal of Economics*, 111 (2), pp.605-637

Friedberg, L. (1998) 'The social security earnings test and labor supply of older men', in Poterba, J.(ed.) *Tax Policy and the Economy*, 12. Cambridge MA: MIT Press. pp. 121-50

Friedberg, L. (2000) 'The Labor Supply Effects of the Social Security Earnings Test', *The Review of Economics and Statistics*, 82 (1), pp.48–63

Friedberg, L., and Webb, A. (2009) 'New Evidence on the Labor Supply Effects of the Social Security Earnings Test', in Moffitt, R. (ed.) *Tax Policy and the Economy*, Chicago: University of Chicago Press. pp. 1-35

Gelber, A, M., Jones, D., Sacks, D, W., and Song, J. (2021) 'Using Nonlinear Budget Sets to Estimate Extensive Margin Responses: Method and Evidence from the Earnings Test', *American Economic Journal: Applied Economics*, 13 (4), pp.150-193

Gruber, J, and Orszag, P. (2003) 'Does the Social Security Earnings Test Affect Labor Supply and Benefit Receipt?', *National Tax Journal*, 56 (4), pp.755–773

Haider, S, J., and Loughran, D, S. (2008) 'The Effect of the Social Security Earnings Test on Male Labor Supply: New Evidence from Survey and Administrative Data', *Journal of Human Resources*, 43 (1), pp.57–87

Hamada, K. (2008) 'Zaisyokuroureinenkin, Kounenreikoyoukeizokukyufu Ga Koureisya No Keizokukoyoukibou Ni Oyobosu Eikyo', *Roudouseisakukenkyu Houkokusyo*, 100, pp.98-112 (in Japanese)

Higuchi, Y., and Yamamoto, I. (2002) 'Wagakuni Danseikoureisya No Roudoukyoukyukoudou Mechanism-Nenkin・Chingin Seido No Koukabunnseki To Koureisyasyuugyou No Syouraizou', *Monetary and Economic Studies*, 21 (2), pp.31-77 (in Japanese)

Ishii, K., and Kurosawa, M. (2009) 'Nenkinseidokaisei Ga Danseikoureisya No Roudoukyoukyuukoudou Ni Ataeru Eikyou No Bunseki', *the Japan Institute of Labour*, 589, pp.43-64 (in Japanese)

Iwamoto, Y. (2000) 'Zaisyokuroureinennkin To Koureisya No Syuugyoukoudou', *The quarterly of social security research*, 35 (4), pp.364-376 (in Japanese)

Ministry of Health, Labour and Welfare (2011) 'Zaisyokuroureinenkinseido No Minaoshi Ni Tsuite', Available at: https://www.mhlw.go.jp/content/12601000/000555792.pdf (Accessed 20 June 2022) (In Japanese)

Ministry of Health, Labour and Welfare (2019) 'Zaisyokuroureinenkinseido No Minaoshi', Available at: https://www.mhlw.go.jp/content/12601000/000555792.pdf (Accessed 20 June 2022) (In Japanese)

Ministry of Health, Labour and Welfare (2020) 'Annual Health, Labour and Welfare Report 2020', Available at: https://www.mhlw.go.jp/stf/wp/hakusyo/kousei/19/index.html (Accessed 11 July 2022) (In Japanese)

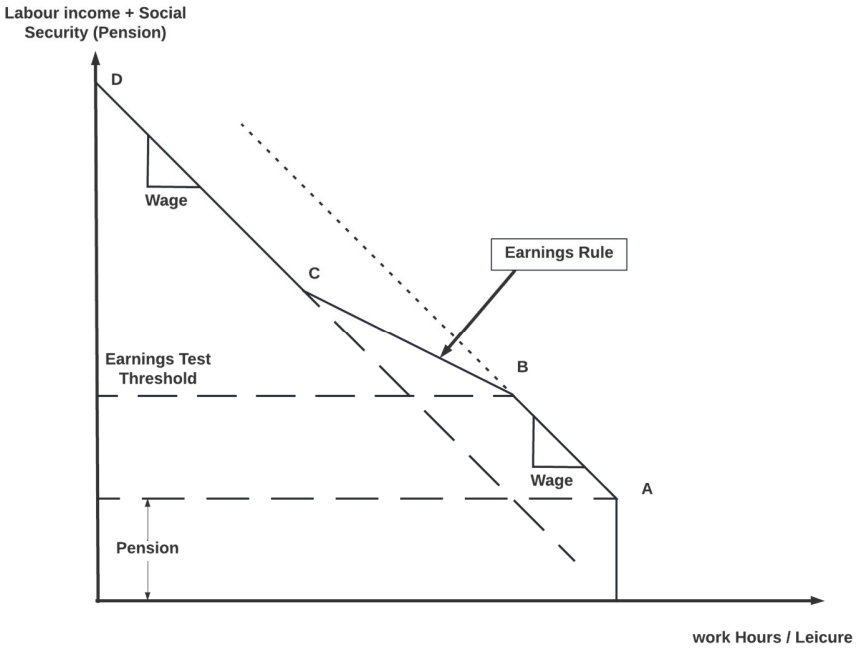
Otake, F., and Yamaga,H. (2003) 'Zaisyokuroureinenkinseido To Danseikoureisya No Roudoukyoukyuu' National Institute of Population and Social Security Research (ed.) Sentaku No Jidai No Syakaihosyou. University of Tokyo Press (In Japanese)

Song, J, G. (2004) 'Evaluating the Initial Impact of Eliminating the Retirement Earnings Test', *Social Security Bulletin*, 65 (1), pp.1–15

Song, J, G., and Manchester, J. (2007) 'New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000', *Journal of Public Economics*, 91 (3–4), pp.669–700

Yamada,A. (2012) 'The labour market behaviour of older people: analysing the impact of the reformed "Earning test" and new "Early withdrawal" rule', *Keio journal of economics*, 104(4), pp.587-605 (in Japanese)

Figure 3-1: Earnings test



Note: Figure 3-1 represents a simple diagram of the earnings test.

Table 4-1 Main amendment history of the Japanese earnings test

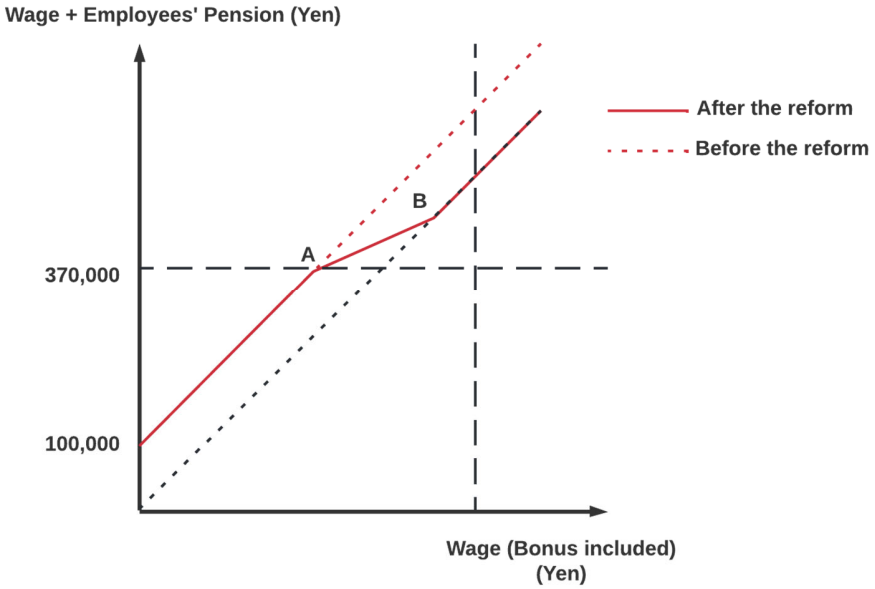
	60-64	65-69	70-
1965		Introduction of the earnings test (Payment rate = 80%)	
1969	Introduction of the earnings test (Four payment rates (80, 60, 40, and 20%))		
1975	Three payment rates (80, 50, and 20%)		
1985		Abolition of the earnings test	
1989	Seven payment rates (20-80%)		
1994	Amended to smoothen the sum of wages and pensions Introduction of 20% pension cut		
2000		Reintroduction of the earnings test (Insurance contribution is required)	
2004	Abolition of 20% pension cut		Reintroduction of the earnings test (Insurance contribution is not required)
2022	Raising the threshold		

Source: Social Security Council (Pensions Subcommittee) (11th October 2011) the Ministry of Health, Labour and Welfare

Figure 4-1 Basics of the Japanese earnings test

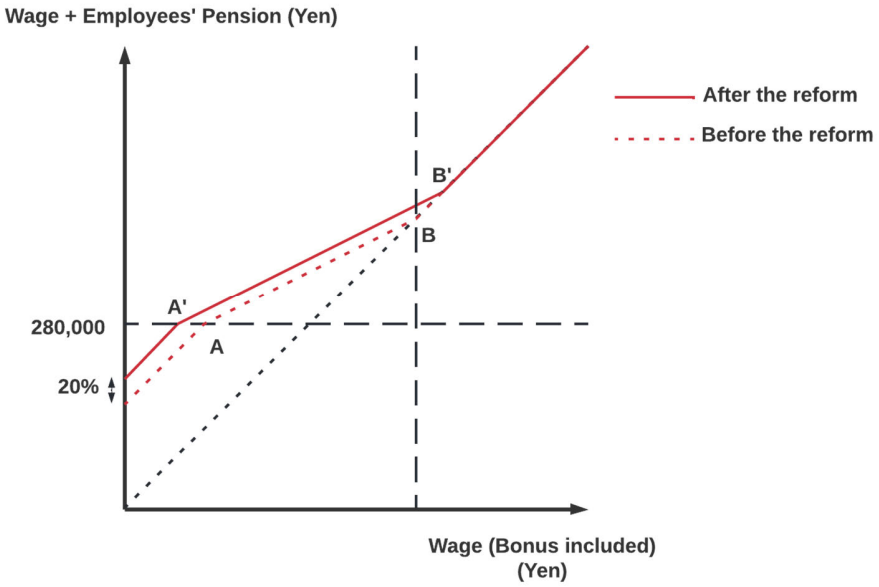
Source: Social Security Council (Pensions Subcommittee) (9th October 2019) the Ministry of Health, Labour and Welfare

Figure 4-2 Image of the reform in 2000
(Reintroduction of the earnings test for 65-69)



Source: Social Security Council (Pensions Subcommittee) (11th October 2011) the Ministry of Health, Labour and Welfare

Figure 4-3 Image of the reform in 2004
(Abolition of 20% pension cut for 60-64)



Source: Social Security Council (Pensions Subcommittee) (11th October 2011) the Ministry of Health, Labour and Welfare

Table 6-1 Descriptive statistics 1

(a) Men

	1997 Male (65-74)		2002 Male (65-74)	
VARIABLES	mean	sd	mean	sd
Employed	0.243	0.429	0.190	0.393
Long hours	0.347	0.476	0.342	0.474
Hours worked	38.09	10.36	36.78	11.90
Urban	0.288	0.453	0.279	0.448
Under15	0.155	0.362	0.122	0.327
Spouse	0.887	0.316	0.883	0.321
Working spouse	0.263	0.440	0.227	0.419
Spouse over 60	0.762	0.426	0.820	0.384
Nemployed	0.766	0.941	0.650	0.873
High school	0.331	0.471	0.358	0.480
College	0.133	0.339	0.138	0.345
Regular	0.112	0.315	0.0669	0.250
Primary sector	0.0103	0.101	0.00975	0.0983
Secondary sector	0.107	0.309	0.0710	0.257
Large size	0.0222	0.147	0.0172	0.130
Observations	29,860		34,249	

(B)Women

	1997 Female (65-74)		2002 Female (65-74)	
VARIABLES	mean	sd	mean	sd
Employed	0.0851	0.279	0.0804	0.272
Long hours	0.267	0.442	0.257	0.437
Hours worked	34.38	12.13	32.82	13.07
Urban	0.267	0.442	0.254	0.435
Under15	0.175	0.380	0.151	0.358
Spouse	0.595	0.491	0.641	0.480
Working spouse	0.384	0.486	0.324	0.468
Spouse over 60	0.739	0.439	0.798	0.401
Nemployed	0.957	1.064	0.815	0.999
High school	0.366	0.482	0.400	0.490
College	0.0447	0.207	0.0506	0.219
Regular	0.0296	0.170	0.0240	0.153
Primary sector	0.00436	0.0659	0.00463	0.0679
Secondary sector	0.0229	0.150	0.0187	0.135
Large size	0.00602	0.0773	0.00548	0.0738
Observations	44,715		48,350	

Table 6-2 Descriptive statistics 2

	Employment						
	Pre (1997)			Post (2002)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 65-69	16,766	0.325	0.469	18,009	0.269	0.444	-0.0201***
Men 70-74	13,094	0.138	0.345	16,240	0.103	0.303	(0.00637)
Women 65-69	23,618	0.118	0.323	24,697	0.115	0.319	0.000303
Women 70-74	21,097	0.0482	0.214	23,653	0.0445	0.206	(0.00359)

	Hours worked >=43						
	Pre (1997)			Post (2002)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 65-69	4,403	0.357	0.479	3,872	0.355	0.478	0.0121
Men 70-74	1,394	0.317	0.466	1,242	0.303	0.460	(0.0213)
Women 65-69	2,131	0.266	0.442	2,235	0.247	0.432	-0.0340
Women 70-74	790	0.270	0.444	807	0.285	0.452	(0.0257)

	Hours worked						
	Pre (1997)			Post (2002)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 65-69	4,403	38.45	10.17	3,872	37.34	11.68	0.843*
Men 70-74	1,394	36.98	10.88	1,242	35.04	12.40	(0.496)
Women 65-69	2,131	34.28	12.10	2,235	32.62	12.88	-0.372
Women 70-74	790	34.66	12.22	807	33.37	13.58	(0.738)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6-3 Effect on employment

VARIABLES	(1) Employment Male OLS	(2) Employment Female OLS	(3) Employment Male Probit	(4) Employment Female Probit
Post	-0.00648**	0.00235	-0.0409	0.0536*
	(0.00314)	(0.00186)	(0.0283)	(0.0288)
Treatment	0.0657***	0.0387***	0.474***	0.504***
	(0.00314)	(0.00188)	(0.0258)	(0.0258)
DID	0.0102**	0.00293	0.0527	-0.00405
	(0.00424)	(0.00258)	(0.0349)	(0.0348)
Constant	0.0178***	0.00831***	-2.003***	-2.282***
	(0.00526)	(0.00311)	(0.0429)	(0.0410)
Observations	63,957	92,913	54,174	89,722
R-squared	0.581	0.495		
Control group	70-74	70-74	70-74	70-74

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 6-3 shows the effect of the earning-test reform in 2000 on employment. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 6-4 Effect on the proportion of full-time workers

VARIABLES	(1) Work>=43 Male	(2) Work>=43 Female
Post	0.0242	0.0341*
	(0.0179)	(0.0207)
Treatment	0.0394***	0.0247
	(0.0141)	(0.0173)
DID	0.00555	-0.0282
	(0.0205)	(0.0241)
Constant	0.269***	0.140***
	(0.0241)	(0.0277)
Observations	10,887	5,948
R-squared	0.082	0.132
Control group	70-74	70-74

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 6-4 shows the effect of the earning-test reform in 2000 on the proportion of full-time workers. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 6-5 Effect on hours worked

	(1)	(2)
	Hours worked	Hours worked
VARIABLES	Male	Female
Post	-0.815**	-0.492
	(0.404)	(0.562)
Treatment	1.348***	0.617
	(0.317)	(0.470)
DID	0.728	-0.210
	(0.462)	(0.655)
Constant	34.46***	29.53***
	(0.545)	(0.752)
Observations	10,887	5,948
R-squared	0.146	0.224
Control group	70-74	70-74

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 6-5 shows the effect of the earning-test reform in 2000 on hours worked. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 7-1 Descriptive statistics 1

(a) Men

	2002 Male (60-69)		2007 Male (60-69)	
VARIABLES	mean	sd	mean	sd
Employed	0.386	0.487	0.447	0.497
Long hours	0.419	0.493	0.383	0.486
Hours worked	39.80	10.91	38.80	11.66
Urban	0.309	0.462	0.316	0.465
Under15	0.0955	0.294	0.0714	0.258
Spouse	0.881	0.324	0.861	0.346
Working spouse	0.297	0.457	0.300	0.458
Spouse over 60	0.654	0.476	0.687	0.464
Nemployed	0.761	0.873	0.760	0.876
High school	0.406	0.491	0.475	0.499
College	0.145	0.352	0.168	0.374
Regular	0.173	0.379	0.179	0.383
Primary sector	0.0121	0.109	0.0128	0.112
Secondary sector	0.153	0.360	0.164	0.370
Large size	0.0475	0.213	0.0795	0.271
Observations	36,803		40,496	

(B)Women

	2002 Female (60-69)		2007 Female (60-69)	
VARIABLES	mean	sd	mean	sd
Employed	0.186	0.389	0.250	0.433
Long hours	0.268	0.443	0.251	0.433
Hours worked	34.01	12.42	33.33	12.90
Urban	0.287	0.452	0.305	0.460
Under15	0.134	0.340	0.0989	0.299
Spouse	0.735	0.442	0.763	0.425
Working spouse	0.377	0.485	0.384	0.486
Spouse over 60	0.854	0.353	0.888	0.315
Nemployed	0.833	0.909	0.840	0.889
High school	0.426	0.494	0.542	0.498
College	0.0702	0.255	0.0807	0.272
Regular	0.0556	0.229	0.0756	0.264
Primary sector	0.00762	0.0870	0.0151	0.122
Secondary sector	0.0482	0.214	0.0530	0.224
Large size	0.0172	0.130	0.0305	0.172
Observations	49,065		53,024	

Table 7-2 Descriptive statistics 2

	Employment						
	Pre (2002)			Post (2007)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 60-64	18,794	0.497	0.500	20,457	0.586	0.493	0.0524***
Men 65-69	18,009	0.269	0.444	20,039	0.305	0.461	(0.00685)
Women 60-64	24,368	0.258	0.438	26,842	0.331	0.471	0.0192***
Women 65-69	24,697	0.115	0.319	26,182	0.168	0.374	(0.00508)

	Hours worked >=43						
	Pre (2002)			Post (2007)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 60-64	8,264	0.449	0.497	10,892	0.424	0.494	0.0355***
Men 65-69	3,872	0.355	0.478	5,083	0.294	0.456	(0.0126)
Women 60-64	5,318	0.277	0.447	7,753	0.256	0.437	-0.0117
Women 65-69	2,235	0.247	0.432	3,629	0.239	0.426	(0.0141)

	Hours worked						
	Pre (2002)			Post (2007)			Difference in Differences
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.	
Men 60-64	8,264	40.85	10.42	10,892	40.43	10.90	1.796***
Men 65-69	3,872	37.54	11.58	5,083	35.33	12.45	(0.288)
Women 60-64	5,318	34.47	12.22	5,318	33.85	12.63	0.0490
Women 65-69	2,235	32.92	12.81	2,235	32.24	13.39	(0.409)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7-3 Effect on employment

VARIABLES	(1) Employment Male OLS	(2) Employment Female OLS	(3) Employment Male Probit	(4) Employment Female Probit
Post	0.0319***	0.0299***	0.157***	0.218***
	(0.00358)	(0.00273)	(0.0187)	(0.0180)
Treatment	0.0802***	0.0761***	0.338***	0.471***
	(0.00380)	(0.00280)	(0.0200)	(0.0176)
DID	0.0118**	0.0119***	0.0688***	-0.0239
	(0.00501)	(0.00384)	(0.0263)	(0.0234)
Constant	0.0913***	0.0750***	-1.554***	-1.550***
	(0.00572)	(0.00474)	(0.0307)	(0.0288)
Observations	76,583	101,334	62,523	91,929
R-squared	0.509	0.455		
Control group	65-69	65-69	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7-3 shows the effect of the earnings test reform in 2004 on employment. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 7-4 Effect on the proportion of full-time workers

VARIABLES	(1) Work>=43 Male OLS	(2) Work>=43 Female OLS	(3) Work>=43 Male Probit	(4) Work>=43 Female Probit
DT	-0.0265***	-0.0203*	-0.0879***	-0.0778*
	(0.00986)	(0.0109)	(0.0290)	(0.0398)
treatment	0.0657***	0.0228**	0.188***	0.0823**
	(0.00927)	(0.0102)	(0.0267)	(0.0366)
DID	0.0296**	0.00636	0.0989***	0.0272
	(0.0119)	(0.0130)	(0.0347)	(0.0471)
Constant	0.322***	0.131***	-0.482***	-1.127***
	(0.0147)	(0.0154)	(0.0421)	(0.0553)
Observations	27,846	18,777	27,846	18,777
R-squared	0.121	0.158		
Control group	65-69	65-69	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7-4 shows the effect of the earnings test reform in 2004 on the proportion of full-time workers. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 7-5 Effect on hours worked

VARIABLES	(1) Hours worked Male	(2) Hours worked Female
Post	-1.387***	-0.951***
	(0.221)	(0.299)
Treatment	2.239***	1.284***
	(0.208)	(0.279)
DID	1.570***	0.674*
	(0.267)	(0.357)
Constant	35.71***	29.19***
	(0.329)	(0.421)
Observations	27,846	18,777
R-squared	0.178	0.252
Control group	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7-5 shows the effect of the earnings test reform in 2004 on hours worked. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 7-6 Effect on employment rate for lower income households

VARIABLES	(1) Employment Male OLS	(2) Employment Female OLS	(3) Employment Male Probit	(4) Employment Female Probit
Post	0.0322***	0.0280***	0.187***	0.201***
	(0.00449)	(0.00364)	(0.0260)	(0.0234)
Treatment	0.0585***	0.0750***	0.294***	0.459***
	(0.00501)	(0.00383)	(0.0288)	(0.0234)
DID	0.0182***	0.0105**	0.0724*	-0.0262
	(0.00662)	(0.00529)	(0.0372)	(0.0313)
Constant	0.0814***	0.0988***	-1.559***	-1.418***
	(0.00731)	(0.00937)	(0.0418)	(0.0541)
Observations	38,523	53,837	33,323	50,389
R-squared	0.507	0.399		
Control group	65-69	65-69	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7-6 shows the effect on employment rate for lower income households whose income is less than 4 million yen. In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Table 7-7 Effect on hours worked for non-regular workers

	(1)	(2)
	Hours worked	Hours worked
VARIABLES	Male	Female
Post	-1.799***	-0.833**
	(0.307)	(0.370)
Treatment	2.882***	1.476***
	(0.310)	(0.347)
DID	1.887***	0.528
	(0.388)	(0.443)
Constant	35.45***	28.19***
	(0.484)	(0.524)
Observations	14,977	12,480
R-squared	0.054	0.063
Control group	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: Table 7-7 shows the effect on hours worked for non-regular workers. In addition to the above variables, explanatory variables include urban, children, spouse, education, industry, work-spouse dummies, and number of employed in the household except the individuals.

Appendix A Marginal effect on employment rate

This appendix reports both the coefficients and the marginal effects on the employment rate for the probit model. Table A shows that the 2004 reform increased the male employment rate by approximately 2.6 percentage points, while it also reports non-significant results for the female employment rate.

Table A Coefficient and marginal effect on employment rate

	(1) Employment Male Probit	(2) Employment Male Probit	(3) Employment Female Probit	(4) Employment Female Probit
VARIABLES	Coefficient	Marginal Effect	Coefficient	Marginal Effect
Post	0.157***	0.0598***	0.218***	0.0461***
	(0.0187)	(0.00709)	(0.0180)	(0.00382)
Treatment	0.338***	0.129***	0.471***	0.101***
	(0.0200)	(0.00764)	(0.0176)	(0.00393)
DID	0.0688***	0.0264***	-0.0239	-0.00502
	(0.0263)	(0.0101)	(0.0234)	(0.00489)
Constant	-1.554***		-1.550***	
	(0.0307)		(0.0288)	
Observations	62,523	62,523	91,929	91,929
Control group				

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: In addition to the above variables, explanatory variables include urban, children, spouse, education, regular-work, industry, work-spouse dummies, and number of employed in the household except the individuals.

Appendix B Hour worked by regular and non-regular workers

This appendix reports the result of working hours for males by regular and non-regular workers. Comparing with the results in Table 7-5, Table B shows larger impact for non-regular workers. On the other hand, that for regular workers is insignificant.

Table B Effect on hours worked by regularity

	(1)	(2)
	Hours worked	Hours worked
VARIABLES	Non regular Male	Regular Male
Post	-1.799***	-0.344
	(0.307)	(0.312)
Treatment	2.882***	1.456***
	(0.310)	(0.267)
DID	1.887***	0.539
	(0.388)	(0.361)
Constant	35.45***	44.38***
	(0.484)	(0.415)
Observations	14,977	12,869
R-squared	0.054	0.029
Control group	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: In addition to the above variables, explanatory variables include urban, children, spouse, education, industry, work-spouse dummies, and number of employed in the household except the individuals.

Appendix C Multinominal Logit

This appendix reports the result of employment status with a multinominal logit model. The basement status is non-employed. Column (1)-(2) report the coefficients, and Column (3)-(5) show the marginal effects. This result suggests that the 2004 reform increased the probability of regular work by 3.4 percentage points, while the probabilities of non-employed and non-regular work are not robust.

Table C Effect on employment status

	(1)	(2)	(3)	(4)	(5)
	Employed Non-regular	Employed Regular	Non-employed	Employed Non-regular	Employed Regular
VARIABLES	Male	Male	Male	Male	Male
	Coefficient	Coefficient	Marginal Effect	Marginal Effect	Marginal Effect
Post	0.300***	0.0292	-0.0334	0.0704***	-0.0370
	(0.0327)	(0.0415)	(0.0904)	(0.00807)	(0.0895)
Treatment	0.619***	1.196***	-0.145	-0.0170	0.162***
	(0.0336)	(0.0381)	(0.381)	(0.346)	(0.0365)
DID	0.0934**	0.221***	-0.0243	-0.00925	0.0335***
	(0.0436)	(0.0509)	(0.0675)	(0.0662)	(0.0102)
Constant	-2.666***	-3.193***			
	(0.0520)	(0.0591)			
Observations	76,583	76,583	76,583	76,583	76,583
Control group	65-69	65-69	65-69	65-69	65-69

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Note: In addition to the above variables, explanatory variables include urban, children, spouse, education, industry, work-spouse dummies, and number of employed in the household except the individuals.