博士論文

Essays in Empirical Labor Economics

（実証労働経済学に関する研究）

Yoshinori Nishimura

西村　仁憲

I am indebted to Hidehiko Ichimura for his advice and encouragement. I am also grateful to Takanori Adachi, Michihito Ando, Eric Bonsang, Hisao Endo, Taiyo Fukai, Gang Chen, Terence Cheng, Denise Doiron, Junya Hamaaki, Konan Hara, Tomohiro Hara, Kosuke Hirose, Toshiaki Iizuka, Jose Iparraguirre, Yasushi Iwamoto, Nobuyuki Izumida, Ryu Kambayashi, Daiji Kawaguchi, Nobuyoshi Kikuchi, Yukinobu Kitamura, Ayako Kondo, Wataru Kureishi, Shiko Maruyama, Kiho Muroga, Ryo Nakajima, Haruko Noguchi, Asako Ohinata, Hideo Owan, Tadashi Sakai, Shinpei Sano, Michio Suzuki, Mari Tanaka, Ryuichi Tanaka, Takahiro Toriyabe, Masanori Tsuruoka, Kengo Yasui and the participants at the conferences and seminars where my studies are presented for their helpful comments. I sincerely thank Eric French, Hirokazu Matsuyama, Stefanie Behncke, Dhaval Dave, and Mathilde Godard for their advice and providing computer codes. Editing assistance has been provided by Philip C. MacLellan. All remaining errors are my own. I sincerely thank Hisataka Anezaki and Tetsuya Iwamoto for explaining the institutional background of public elderly care services and commenting on our analysis. I also sincerely thank Masato Hisatake, Hiroyuki Motegi and Masato Oikawa for their support and encouragement. Without their help, I could not have completed this dissetation. Finally, I wish to thank my parents for their supports and encouragement. This research is supported by Grant-in-Aid for JSPS Fellows, Grant Number 2510637. This work was also supported by JSPS KAKENHI Grant Number 15H05692 (Principal Investigator: Ichimura Hidehiko). The use of data in this paper is approved by Ministry of Health, Labour and Welfare under this grant.

Data Acknowlegement

The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

The ELSA (English Longitudinal Study of Ageing) is funded by the US National Institute on Ageing and a consortium of UK government departments coordinated by the Office for National Statistics.

This paper uses data from SHARE Waves 1, 2, 3 (SHARELIFE), 4 and 5 (DOIs: 10.6103 /SHARE.w1.260, 10.6103/SHARE.w2.260, 10.6103/SHARE.w3.100, 10.6103/SHARE.w4.111, 10.6103/

SHARE.w5.100), see Börsch-Supan et al. (2013) for methodological details. The SHARE data collection has been primarily funded by the European Commission through FP5 (QLK6-CT-2001-00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005

-028857, SHARELIFE: CIT4-CT-2006-028812) and FP7 (SHARE-PREP: N° 211909, SHARE-LEAP: N° 227822, SHARE M4: N° 261982). Additional funding from the German Ministry of Education and Research, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG00

5842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064$) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

The Japanese Study of Aging and Retirement (JSTAR) was conducted by the Research Institute of Economy, Trade and Industry (RIETI), Hitotsubashi University, and the University of Tokyo.

The Longitudinal Survey of Middle-aged and Elderly Persons is provided by the Ministry of Health, Labour and Welfare.

The Preference Parameters Study was conducted by the Osaka University Institute of Social and Economic Research.

The data for this secondary analysis, "Chusho kigyo rodo jijo jittai chosa, National Federation of Small Business Associations," was provided by the Social Science Japan Data Archive, Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo.

Table of Contents

Chapter 1: Introduction

Due to rising life expectancies and declining birthrates associated with economic development, many industrialized countries are now facing the problem of an aging population. In 2015, there were 900 million people over 60 years of age worldwide, and this number is expected to continue to grow rapidly. As a country's population ages, the cost of social security and welfare increases, eroding the country's budget, and so numerous developed countries have introduced retirement-related policies such as pension system reform in order to reduce the cost of social security and social welfare to a sustainable level. Pension reforms in developed countries are mostly targeted at delaying retirement, and the United States, United Kingdom and Korea, for example, have decided to increase the age of pension eligibility, while Japan has already done so. The relationship between social security and retirement in developed countries has attracted a fair amount of attention in economics (Gruber and Wise (1998)). This dissertation includes studies analyzing retirement and some topics related to retirement. This dissertation consists of four studies.

In Chapter 2, we analyze whether government intervention on firms' employment policies have an effect on the employment of the elderly. As a result of the pensionable age increasing in Japan, this policy distinguishes between the mandatory retirement age and the pensionable age. The Japanese government has obliged firms to employ elderly workers until they reach the pensionable age. According to literature, the labor force participation rate of elderly male workers increased just after the implementation of this policy. However, according to this paper's results, there is no effect on the employment of the elderly workers after the introduction of the policy. Consequently, this paper discusses why the government intervention in the demand side of the elderly labor market had no effect on elderly employment. According to this discussion, it is possible that firms have avoided the cost of employing the elderly by using measures that, while following the letter of the law, do not fully support the policy aims.

In Chapter 3, we analyze the effect of informal elderly care on caregiver labor supply. Since the Japanese government intervenes on the supply side of the elderly care market and market entry of nursing home suppliers is regulated, this analysis utilizes exogenous variations from the supply side of government intervention on the elderly care market. Owing to such intervention and regulation, public nursing home capacity exogenously changes for caregivers, which we use to estimate the effect of informal elderly care on labor supply. To the best of our knowledge, no study has thus far utilized exogenous institutional variation as an instrument to estimate this effect. Analysis results reveal that the effect of informal elderly care on female labor force participation is negative. By contrast, male labor force participation is not affected by such care, since, in Japan, females spend more time on informal care than males. The increase in nursing home capacity is thus effective for decreasing the female burden of informal care. This chapter is based on Nishimura and Oikawa (2017).

In Chapter 4, we analyze the effect of the Japanese social security system on retirement behavior. We simulate that some counterfactual reforms on the Japanese social security system will inuence the Japanese male elderly's labor supply. If the eligibility age of receiving basic pension is changed to age 70, the labor participation rate of the Japanese male elderly increases by about 4.7 percent between age 63 and age 69 in average. If the full amount of basic pension is cut in half, the reform increases the labor participation rate of the Japanese male elderly by about 1.4 percent between age 63 and age 75 in average. We estimate the structural parameters of the utility function of the Japanese households. We find that the estimated Japanese consumption weight is close to the estimated value in the U.S.

In Chapter 5, we analyze the reasons for differences in the estimated effect of retirement on health in previous studies. We investigate these differences by focusing on the analysis methods used by these studies. Using various health indexes, numerous researchers have examined the effects of retirement on health. However, there are no unified views on the impact of retirement on various health indexes. Consequently, we show that the choice of analysis method is one of the key factors in explaining why the estimated results of the effect of retirement on health differ. Moreover, we reestimate the effect of retirement on health by using a fixed analysis method controlling for individual heterogeneity and endogeneity of the retirement behavior. We analyze the effect of retirement on health parameters, such as cognitive function, self-report of health, activities of daily living (ADL), depression, and body mass index in eight countries. We find that the effects of retirement on self-report of health, depression, and ADL are positive in many of these countries. This chapter is based on Nishimura et al. (2018).

References

Gruber, J. and Wise, D. A. (1998). Social Security and Retirement: An International Comparison. *American Economic Review, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association*, 88(2):158–163.

Nishimura, Y. and Oikawa, M. (2017). Effects of Informal Elderly Care on Labor Supply: Exploitation of Government Intervention on the Supply Side of Elderly Care Market. *The University of York, Health, Econometrics and Data Group Working Papers WP17/02.*

Nishimura, Y., Oikawa, M., and Motegi, H. (2018). What Explains the Difference in the Effect of Retirement on Health?: Evidence from Global Aging Data. *Journal of Economic Surveys, forthcoming.*

Chapter 2: Did Government Intervention on Firm's Employment Policies Have an Effect on the Employment of Elderly Workers?

1 Introduction

Retirement related policies, such as a reform of the pension system have become important in developed countries as to sustain social security systems. Many developed countries have faced the same problems of decreasing birthrate and ageing populations. As a population ages, the cost of social security and social welfare increases, eroding the country's budget. As such, numerous developed countries have reformed their pension systems to reduce the cost of social security and social welfare. Many developed countries, such as the United States, the United Kingdom, and Korea have already decided to increase pension eligibility age over the next decades. Japan has already increased pension eligibility age. Pension reforms in developed countries are expected to influence retirement. As Gruber and Wise (1998) discuss, the relationship between the social security system and retirement in developed countries has attracted a lot of attention in economics. In many developed countries, regulations about the mandatory retirement system have also been reconsidered when reforming the pension systems, especially after 2000. For example, the UK, Germany, and France have reformed the law that regulates mandatory retirement age. However, in the US, the mandatory retirement system has been abolished in the 1980s.

In the US, there are studies that provide direct evidence on the effect of the abolition of mandatory retirement age,[1] which is discussed in this paper (Neumark and Stock (1999), Ashenfelter and Card (2002), von Wachter (2002) and Adams (2003)). Ashenfelter and Card (2002) analyze the labor market for university professors. According to their results, the employment of workers protected by the law increases. Except in the US, there is not enough evidence with respect to the effect of reforming regulations on the mandatory retirement age, although some developed countries have reformed regulations regarding the mandatory retirement system. In fact, the results in this paper are different compared to the result in the US just after the introduction of the policy. Below, I discuss why this is the case.

In Japan, the government has changed the basic pension eligibility age from 60 to 65 so as to decrease the payment amount for public pensions. However, many firms have set their mandatory retirement age at around 60 and, as a result, many elderly reach the mandatory retirement age before they start receiving their public pension. The Japanese government has recently encouraged firms to re-employ elderly people after reaching the mandatory retirement age until they arrive at the basic pensionable age (flat-rate part).[2] This regulation is called the Elderly Employment Stabilization Law (EESL). Kondo and Shigeoka (2016) were the first to analyze this policy,[3] estimating the probability of being a salaried worker at age 60-65 and comparing the 1945 and 1946 birth year cohorts to establish the effect of the EESL immediately after the implementation of the policy. They found that the 1946 birth year cohort was more likely to consist of salaried workers at ages 60 and 61 by 2.4 and 3.2 percent, respectively. While this effect seems small, the study suffers from a data limitation as I subsequently explain. The goal of this paper, then, is to estimate the effect of the EESL on the employment of the elderly and discuss how firm have reacted to this policy after implementation. According to the results, there are no significant positive effects on the

[1] Since Lazear (1979), theoretical research that answers why there is a mandatory retirement has developed. Examples are the related studies such as Lazear (1981), Burkhauser and Ouinn (1983), Lazear and Moore (1984) and Lang (1989).

[2] The government allowed firms that used a restrictive reemployment system not to remove it.

[3] Clark and Ogawa (1992) estimated the effect of the change in the mandatory retirement policy on the wage profile before the EESL.

employment of the elderly immediately after the introduction of the policy, which is discussed in subsequent sections. In this paper, I mainly focus on cohorts born from November, 1945[4] to March, 1947. According to the Ministry of Health, Labour and Welfare, there is an exemption to the implementation of the EESL before 2013, and there are no clear statements with respect to wage contracts when a firm engages in a contract with a worker who wants to continue work after the mandatory retirement age. Additionally, there is an important exception: before 2013, a firm could restrict the workers offered reemployment by accepting the agreement from a labor union. This is an "escape route" from additional costs, which firms could use. As explained, most firms react to this policy by introducing a reemployment system, without abolishing the mandatory retirement system or increasing the mandatory retirement age, which means that many firms choose a reaction that enables them to use these "escape routes." This point is further discussed in the subsequent sections.

There are numerous related studies that analyze government intervention in the labor market. However, the studies directly analyzing the effect of changing the mandatory retirement policy on the employment of the elderly are limited, and are discussed in the literature review section. I also provide evidence by showing what happened after the implementation of the government intervention in the demand side of the elderly labor market. The remainder of this paper is organized as follows: in section 2, I discuss the effect of the EESL and review literature; section 3 describes the data; in section 4, I explain the estimation procedure; section 5 reports the results; and section 6 concludes the paper.

[4]The Longitudinal Survey of Middle-aged and Elderly Persons used for this analysis includes the respondents with the birth date from 1945 November.

2 Discussion and Literature Review

2.1 What is the EESL?

With respect to the EESL, Kondo and Shigeoka (2016) explain its details. Briefly, the Japanese pension program is divided into two parts: the basic pension (flat rate part) and the income-related pension (wage proportional part). The pension eligibility ages of these two programs are different. This paper uses the pension eligibility age presented in Motegi et al. (2016). Please see Table 1. In Japan, the pension eligibility age has gradually increased. For employees in private companies or the public sector, the pension including the basic pension and the income-related pension, which is called the Employees' Pension Insurance or the Mutual Aid Insurance, are provided. For self-employed workers, only the basics pension, which is called the National Pension Insurance, is provided by the government.

The EESL is a law which obliges a firm to increase the mandatory retirement age, omit the mandatory retirement system, or give a reemployment offer and employ workers reaching the mandatory retirement age until they arrive at the basic pensionable age (flat rate) after 2006. Depending on the birth year of elderly workers, the pensionable age increases. The mandatory retirement age is around 60 in Japan. As a result, for example, the elderly born in 1945 arrive at the mandatory retirement age before they arrive at the basic pensionable age (flat rate) (age 63) if the mandatory retirement age is age 60. The government prepared this law to fill a gap between the pensionable age (flat rate) and the mandatory retirement age. Figure 1 shows this fact. The year in this figure is the birth year (e.g., 1947, 1948). For example, with respect to workers born between 1944 and 1945, there is a gap between the pensionable age (flat rate) and the of age 60. The blue line shows the age when a worker starts receiving pension (flat rate part). With respect to the workers born after 1946, the government obliges firms to increase the mandatory retirement age, abolish the mandatory retirement system, or make a reemployment offer and employ workers arriving at the mandatory retirement age until they arrive at the pensionable age (flat rate part). This is, in summary, the EESL concept.

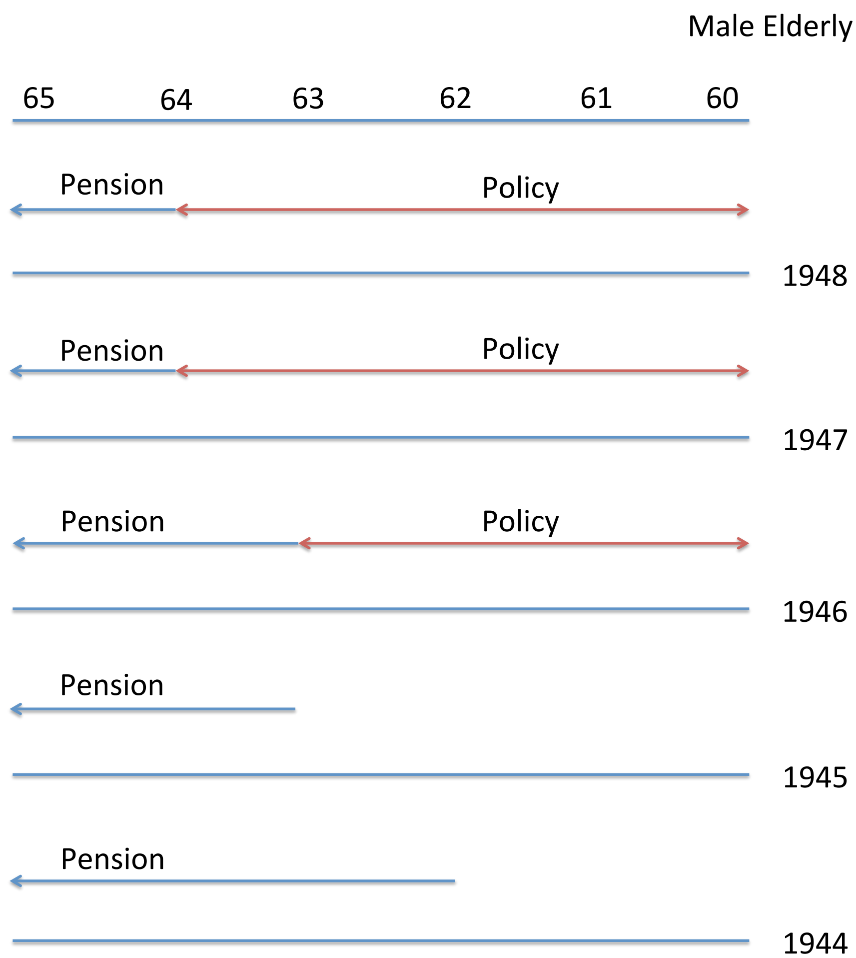
However, in the EESL, there is an important exception. Before 2013, a firm could restrict the workers who can get a reemployment offer by accepting the agreement from a labor union. This exception has been omitted after 2013, and a firm is obliged to employ all workers who want to continue to work in the firm after the mandatory retirement age. Additionally, there is no clear statement in the law with respect to wage contracts when a firm makes a reemployment offer to a worker reaching at the mandatory retirement age. As explained in section 6 , many firms introduce the reemployment system without increasing the mandatory retirement age or omitting it. When a firm engages in a reemployment contract with a worker, they discuss a decreased wage rate with the workers who intend to work in the firm after the mandatory retirement age, as the law does not concretely mention anything with respect to decreasing wage rates for these workers.

Table 1: Public Pension Reform in Japan

		Pensionable age			
	Public pension reform year	Employees' pension		Mutual aid pension	
Birth Cohort		Flat-rate part	Wage proportional part	Flat-rate part	Wage proportional part
		Men		**Men & Women**	
-1941.4.1	-	60	60	60	60
1941.4.2-1943.4.1	2001	61	60	61	60
1943.4.2-1945.4.1	2004	62	60	62	60
1945.4.2-1947.4.1	2007	63	60	63	60
1947.4.2-1949.4.1	2010	64	60	64	60
1949.4.2-1953.4.1	2013	65	60	65	60
1953.4.2-1955.4.1	2013	65	61	65	61
1955.4.2-1957.4.1	2016	65	62	65	62
1957.4.2-1959.4.1	2019	65	63	65	63
1959.4.2-1961.4.1	2022	65	64	65	64
1961.4.2-	2025	65	65	65	65
		Women			
-1932.4.1	-	55	55		
1932.4.2-1934.4.1	1987	56	56		
1934.4.2-1936.4.1	1990	57	57		
1936.4.2-1937.4.1	1993	58	58		
1937.4.2-1938.4.1	1993	58	58		
1938.4.2-1940.4.1	1996	59	59		
1940.4.2-1946.4.1	2001	60	60		
1946.4.2-1948.4.1	2006	61	60		
1948.4.2-1950.4.1	2009	62	60		
1950.4.2-1952.4.1	2012	63	60		
1952.4.2-1954.4.1	2015	64	60		
1954.4.2-1958.4.1	2018	65	60		
1958.4.2-1960.4.1	2018	65	61		
1960.4.2-1962.4.1	2021	65	62		
1962.4.2-1964.4.1	2024	65	63		
1964.4.2-1965.4.1	2027	65	64		
1965.4.2-	2030	65	65		

Source: Ministry of Health, Labor and Welfare

Figure 1: The Elderly Employment Stabilization Law (Birth Years 1945-48)



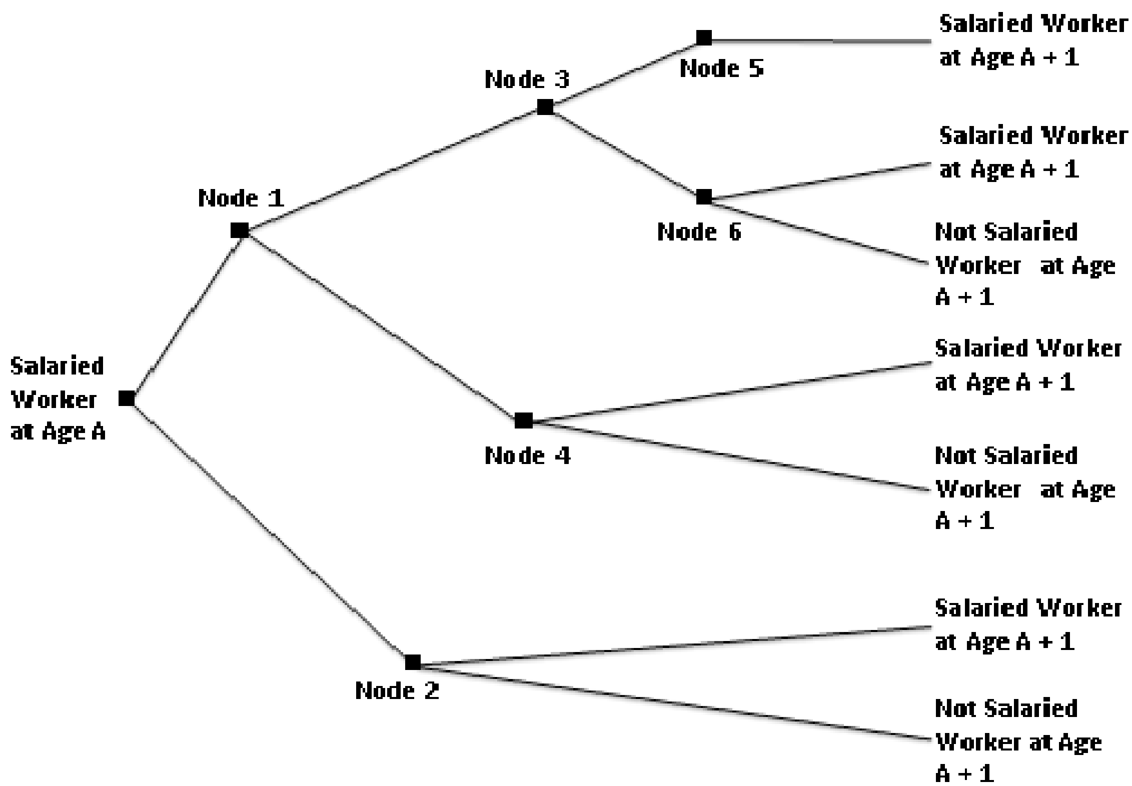
2.2 The effect of EESL

In this section, I discuss what happened after the implementation of the EESL. This policy makes firms change their employment system. As a result, it also works as a restriction for firms. If we were to precisely understand the impact of this policy, we have to analyze the channels which influence the outcome of whether a worker works or not. For example, one is whether firms increase the number of offers to reemploy workers (Channel (2)). The other is whether firms rescind or increase the mandatory retirement age (Channel (1)). Another is whether workers accept the offer or not (Channel (3)). If we only consider whether a worker works or not after the mandatory retirement age, we cannot distinguish between the three channels. The results in the literature (Kondo and Shigeoka (2016)) consider the combined effects on each channel. In this paper, to clarify and understand the EESL effect, I discuss the factors that decide the impact of this policy with respect to the labor participation rate after the implementation of the EESL. By using this framework, we can better understand the estimations of both the literature and this paper and better interpret the result.

The retirement path of a worker after the mandatory retirement age is shown in Figure 2. Assuming that only one cohort exists, I explain the meaning of each node in Figure 2:

- Node 1: A worker i faces the mandatory retirement age.
- Node 2: A worker i does **not** face the mandatory retirement age.
- Node 3: A worker i receives a re-employment offer at age A.
- Node 4: A worker i does **not** receive a reemployment offer at age A.
- Node 5: A worker i **accepts** a reemployment offer at age A.
- Node 6: A worker i **rejects** a reemployment offer at age A.

Figure 2: Retirement Path of Workers after the Mandatory Retirement Age



Subsequently, I define the following sets:

$$A^{T}_{\text{salaried}} = \left\{ i \middle| \ i \text{ is a salaried worker at age } T. \ \right\}$$

$$A_{\text{node } k} = \left\{ i \middle| \ i \text{ is on node } k. \ \right\}$$

Then, I consider the meaning of the following probability:

$$\Pr\left\{ i \in A^{A+1}_{\text{salaried}} \middle| \ i \in A^{A}_{\text{salaried}} \ \right\} \tag{1}$$

For simplicity, I define the probability as follows:

$$\Pr\left\{ A^{A+1}_{\text{salaried}} \middle| \ A^{A}_{\text{salaried}} \ \right\} = \Pr\left\{ i \in A^{A+1}_{\text{salaried}} \middle| \ i \in A^{A}_{\text{salaried}} \ \right\} \tag{2}$$

I also discuss the policy effects of the EESL by using the following expression, thus showing that there are three important paths through which this policy influences workers and firms:

$$\begin{aligned}
&\Pr\left\{ A^{A+1}_{\text{salaried}} \middle| \ A^{A}_{\text{salaried}} \ \right\} \\
&= \Pr\left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node } 5} \right\} \underbrace{\Pr\left\{ A_{\text{node } 5} \middle| A_{\text{node } 3} \right\}}_{\text{Effect Channel (1)}} \underbrace{\Pr\left\{ A_{\text{node } 3} \middle| A_{\text{node } 1} \right\}}_{\text{Effect Channel (2)}} \underbrace{\Pr\left\{ A_{\text{node } 1} \middle| A^{A}_{\text{salaried}} \right\}}_{\text{Effect Channel (3)}} \\
&+ \Pr\left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node } 6} \right\} \Bigg(1 - \underbrace{\Pr\left\{ A_{\text{node } 5} \middle| A_{\text{node } 3} \right\}}_{\text{Effect Channel (1)}}\Bigg) \underbrace{\Pr\left\{ A_{\text{node } 3} \middle| A_{\text{node } 1} \right\}}_{\text{Effect Channel (2)}} \underbrace{\Pr\left\{ A_{\text{node } 1} \middle| A^{A}_{\text{salaried}} \right\}}_{\text{Effect Channel (3)}} \\
&+ \Pr\left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node } 4} \right\} \Bigg(1 - \underbrace{\Pr\left\{ A_{\text{node } 3} \middle| A_{\text{node } 1} \right\}}_{\text{Effect Channel (2)}}\Bigg) \underbrace{\Pr\left\{ A_{\text{node } 1} \middle| A^{A}_{\text{salaried}} \right\}}_{\text{Effect Channel (3)}}
\end{aligned}$$

$$+ \Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 2}}\right\}\Bigg(1-\underbrace{\Pr\left\{A_{\text{node 1}}\middle|A^{A}_{\text{salaried}}\right\}}_{\text{Effect Channel (3)}}\Bigg)$$

I define the above expression for one cohort. To consider the difference of (1) between two cohorts, I introduce another cohort. To simplify the discussion, consider that there are only two cohorts, $C1$ and $C2$. Additionally, there is only one mandatory retirement age, A. Assume that the policy is introduced after cohort $C1$ faces the mandatory retirement age and some individuals do not face the mandatory retirement age because some firms did not introduce the mandatory retirement system. I analyze the difference between cohort $C1$ and cohort $C2$. However, I omit the discussion on the difference of the following terms of (3) to focus on the effects on the demand side of the labor market.[5]

$$\Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 5}}\right\}, \Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 6}}\right\}, \Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 4}}\right\}, \Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 2}}\right\} \quad (3)$$

By the way, $\Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 5}}\right\}=1$ because a salaried worker will become a salaried worker when he/she accepts a reemployment offer, the first effect on Channel (1) represents the path of the effect of acceptance of reemployment by workers. Some firms may decrease wages to reduce employment cost when they reemploy workers after the mandatory retirement age. If the amount of the offered wage is very low when workers are reemployed, these workers may reject the offer. As a result, the acceptance rate may decrease. The second effect on Channel (2) represents the path of the effect that a firm prepares an office where workers are able to work after the mandatory retirement age. Channel (3) represents the path where some firms rescind or increase the mandatory retirement age. Effect Channel (3) represents the path that

[5] The effect on the four terms of (3) is caused by the difference in the characteristics of workers who arrive at nodes 2, 4, 5, and 6. In fact, when workers arrive at nodes 2, 4, 5, and 6, they decide whether they will continue being salaried or not. If workers arrive at node 6, they have to apply to another firm. Whether they become a salaried worker at age $A+1$ or not depends on the state variables (e.g. pension eligibility in the next period) which workers have on nodes 2, 4, 5, and 6. With the introduction of the EESL, the distribution of the characteristics of workers on nodes 2, 4, 5, and 6 changes. However, these influences are not clear.

some firms rescind or increase the mandatory retirement age after the EESL. Then, I define $\Delta \Pr \left\{ A^{A+1}_{\text{salaried}} \middle| \ A^{A}_{\text{salaried}} \ \right\}$ as the difference of $\Pr \left\{ A^{A+1}_{\text{salaried}} \middle| \ A^{A}_{\text{salaried}} \ \right\}$ between $C1$ and $C2$. Finally, I can derive the following relationship by the definition of probability measure.

$$\Delta \Pr \left\{ (A^{A+1}_{\text{salaried}})^c \middle| \ A^{A}_{\text{salaried}} \ \right\} = -\Delta \Pr \left\{ A^{A+1}_{\text{salaried}} \middle| \ A^{A}_{\text{salaried}} \ \right\} \tag{4}$$

I discuss the relationship between (4) and the results in the literature. Kondo and Shigeoka (2016) estimate $\hat{\beta}_{61} - \hat{\beta}_{60} = 0.032 - 0.024 = 0.008$. Let δ_1 and δ_2 be the factors included in $\beta_{61} - \beta_{60}$. I explain these in the next section, along with the relationship $\beta_{61} - \beta_{60} = -\Delta \Pr \left\{ (A^{61}_{\text{salaried}})^c \middle| \ A^{60}_{\text{salaried}} \ \right\} + \delta_1 + \delta_2 = \Delta \Pr \left\{ A^{61}_{\text{salaried}} \middle| \ A^{60}_{\text{salaried}} \ \right\} + \delta_1 + \delta_2$ using relationship (4).

2.3 Literature Review

2.3.1 Literature Estimates

Kondo and Shigeoka (2016) use a dummy variable of being a salaried worker. The outcome is influenced by effects from multiple channels, explained in detail subsequently. They used the following outcome:

$$y_i = \begin{cases} 1 & \text{if } i \text{ is a salaried worker at survey year.} \\ 0 & \text{if } i \text{ is not a salaried worker at survey year.} \end{cases} \tag{5}$$

They analyzed two cohorts, which had the same pension eligibility age. If I consider this environment, they utilize an environment where the difference of the following probabilities $\Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 5}} \right\}, \Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 6}} \right\}, \Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 4}} \right\}, \Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 2}} \right\}$ between the two cohorts is small. As I explained, $\Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 5}} \right\} = 1$. For example, let us consider $\Pr \left\{ A^{A+1}_{\text{salaried}} \middle| A_{\text{node 2}} \right\}$. If two workers (workers 1, 2) with different pension eligibility ages arrive at node 2, it is possible that the decisions of these workers are different conditional on demographics. If the pension eligibility age of one cohort (worker 1) is age $A+1$ and the other (worker 2) is age $A+2$, worker 2 is more willing to work conditional on demographics. I can discuss the terms $\Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 6}}\right\}$, $\Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 4}}\right\}$, $\Pr\left\{A^{A+1}_{\text{salaried}}\middle|A_{\text{node 2}}\right\}$ can be analyzed in the same manner.

Age_i is a vector of age dummy variables. $Age_{it} = 1$ means that the dummy variables, except the age t dummy variable, are zero in the vector Age_i and the age t dummy variable is equal to one. Kondo and Shigeoka (2016) estimate the following parameter:[6] $T_i = 1$ if the birth year of i is 1946.

$$
\begin{aligned}
\beta_{61} - \beta_{60} = &\Big(\Pr[y_i = 1|T_i = 1, X_i = x, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 1, X_i = x, Age_{i60} = 1]\Big) \\
&- \Big(\Pr[y_i = 1|T_i = 0, X_i = x, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 0, X_i = x, Age_{i60} = 1]\Big)
\end{aligned}
$$

Then, I can rewrite this parameter as follows. Here, $\Pr[y_i = 1|X_i = x, T_i = 1, Age_{i60} = 1] = \alpha_{60} + \beta_{60} + \gamma + \delta' x$.

$$
\begin{aligned}
&\Big(\Pr[y_i = 1|T_i = 1, X_i = x, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 1, X_i = x, Age_{i60} = 1]\Big) \\
&\quad - \Big(\Pr[y_i = 1|T_i = 0, X_i = x, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 0, X_i = x, Age_{i60} = 1]\Big) \\
&\qquad = \underbrace{\underbrace{\Big[\Big(\Pr[y_i = 1|T_i = 1, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 1, Age_{i60} = 1]\Big)}_{\text{Part 1}} \\
&\qquad\quad - \Big(\Pr[y_i = 1|T_i = 0, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 0, Age_{i60} = 1]\Big)\Big]} \\
&\qquad\quad - \Big[\Big(\int \delta' x dF(x|T_i = 1, Age_{i61} = 1) - \int \delta' x dF(x|T_i = 1, Age_{i60} = 1)\Big) \\
&\qquad\qquad - \Big(\int \delta' x dF(x|T_i = 0, Age_{i61} = 1) - \int \delta' x dF(x|T_i = 0, Age_{i60} = 1)\Big)\Big]
\end{aligned}
$$

[6] They assume $E[\epsilon_i|X_i = x, T_i = t, Age_i = a] = 0$

In Part 1, y_i^t can be defined as:

$$y_i^t = \begin{cases} 1 & \text{if } i \text{ is a salaried worker at age } t. \\ 0 & \text{if } i \text{ is not a salaried worker at at age } t. \end{cases} \tag{6}$$

Part 1 can be rewritten as follows:

$$\begin{aligned}
&\underbrace{\Big(\Pr[y_i = 1|T_i = 1, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 1, Age_{i60} = 1]\Big)}_{\text{Part 1}} \\
&\underline{-\Big(\Pr[y_i = 1|T_i = 0, Age_{i61} = 1] - \Pr[y_i = 1|T_i = 0, Age_{i60} = 1]\Big)} \\
&= \underbrace{-\Big(\Pr[y_i^{61} = 0, y_i^{60} = 1|T_i = 1] - \Pr[y_i^{61} = 0, y_i^{60} = 1|T_i = 0]\Big)}_{\text{Part 2 (This part equals } -\Delta \Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \; A^{60}_{\text{salaried}} \right\})} \\
&\quad + \Big(\Pr[y_i^{61} = 1, y_i^{60} = 0|T_i = 1] - \Pr[y_i^{61} = 1, y_i^{60} = 0|T_i = 0]\Big)
\end{aligned}$$

Let me assume that the population of one cohort is fixed. If the mandatory retirement age is 60, Part 2 of the following expression means $\Delta \Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \; A^{60}_{\text{salaried}} \right\}$. This is the difference in $\Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \; A^{60}_{\text{salaried}} \right\}$ between cohort 1945 and cohort 1946.
It is possible that the influence of the following parts is small if I considering the meaning of each part.

- $$\begin{aligned}
-\delta_1 = &\Big(\int \delta' x dF(x|T_i = 1, Age_{i61} = 1) - \int \delta' x dF(x|T_i = 1, Age_{i60} = 1)\Big) \\
&-\Big(\int \delta' x dF(x|T_i = 0, Age_{i61} = 1) - \int \delta' x dF(x|T_i = 0, Age_{i60} = 1)\Big)
\end{aligned}$$

  Remark: The difference-in-differences of conditional expectation about $\delta' x$ between age 61 and age 60. Kondo and Shigeoka (2016) use region dummies and the unemployment rate as control variables.

- $\delta_2 = \Pr[y_i^{61} = 1, y_i^{60} = 0 | T_i = 1] - \Pr[y_i^{61} = 1, y_i^{60} = 0 | T_i = 0]$

  Remark: The difference in the probability of being a salaried worker at age 61 while not a salaried worker at age 60.

I derive the relationship $\beta_{61} - \beta_{60} = -\Delta \Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \ A^{60}_{\text{salaried}}\right\} + \delta_1 + \delta_2 = \Delta \Pr\left\{A^{61}_{\text{salaried}} \middle| \ A^{60}_{\text{salaried}}\right\} + \delta_1 + \delta_2$ using equation (4). Kondo and Shigeoka (2016) estimates $\hat{\beta}_{61} - \hat{\beta}_{60} = 0.032 - 0.024 = 0.008$. It is possible that this magnitude is produced by $\Delta \Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \ A^{60}_{\text{salaried}}\right\}$ and small factors δ_1 and δ_2. The estimate of Kondo and Shigeoka (2016) was influenced by multiple channels, from all the effects on the three channels discussed in section 2.2. In addition, the EESL directly influences only $-\Delta \Pr\left\{(A^{61}_{\text{salaried}})^c \middle| \ A^{60}_{\text{salaried}}\right\}$, while the estimated coefficients of $\beta_{61} - \beta_{60}$ are influenced by other factors δ_1 and δ_2.

2.3.2 Effect of Government Intervention on the Elderly Labor Market in the US

According to the literature on the US, since the 1980s, studies about retirement have been published continuously (e.g., Fields and Mitchell (1984), Alan and Thomas (1986) and Slade (1987)). With respect to mandatory retirement in the US, Neumark (2003) explains its history and the relevant literature.6 Some studies have focused on government intervention in the supply side of the labor market (e.g., Staubli and Zweimüller (2013) and Neumark and Song (2013)). I here discuss the results on the US related to this paper. Since around 1990, in the US, many studies have provided evidence with respect to how a firm discriminates workers based on their age. (e.g., Hutchens (1988), Hirsch et al. (2000) and Adams (2002)). Johnson and Neumark (1997) analyze the consequences of age discrimination in the workplace. Lahey (2008) analyzes the effect of the age discrimination law on the labor market. The following four studies directly analyze the abolition of the mandatory retirement system.

- Neumark and Stock (1999)

  - After the implementation of age discrimination laws, the labor force participation of workers protected by age discrimination laws increases.
  - With respect to other workers which age discrimination laws do not protect, the effect is not clear.
  - They indicate that age discrimination laws steepen age-earning profiles for workers entering the labor market.

- Ashenfelter and Card (2002)
  - A special exemption from the 1986 Age Discrimination Act allowed colleges and universities not to abolish compulsory retirement at age 70 until 1994.
  - After the abolition of mandatory retirement, the retirement rates at 70 and 71 fell by two thirds after 1994.

- von Wachter (2002)
  - Overall, the labor force of workers 65 and older increases by 10 percent to 20 percent after the end of mandatory retirement. Neither job tenure nor wage of older workers were affected.

- Adams (2003)
  - This study analyzes the effect of age discrimination laws on employment, hiring, and retirement.
  - With respect to employment, the labor force participation rate increases for the workers which the laws protect.
  - However, there is no clear effect with respect to the workers which the laws do not protect.
  - With respect to hiring and retirement, there is no effect.

2.4 Paper Objectives and Results

This paper analyzes the effect of government intervention in the demand side of the elderly labor market on the employment of the elderly. According to the literature review, with respect to the workers which age discrimination laws protect, the labor force participation rate increases. However, as I discuss in the subsequent sections, the results are different for Japan, and I further discuss the reasons for this.

Finally, I analyze why the results are different from those of Kondo and Shigeoka (2016), who find that the effect of the EESL is significant, but weak, while we find a lack of significant effects. There are two main reasons. First, this study focuses only on analyzed groups directly influenced by the EESL, while the Kondo and Shigeoka (2016) estimates are influenced by effects unrelated to the EESL, as discussed in section 2.2. Second, there is a difference in the estimation procedure. This study eliminates any potential unobserved heterogeneity and controls for important demographics, and finds that there is no significant effect if we control for and eliminate these factors that cause bias in the coefficient.

3 Data

The main analysis sample for this study was compiled from the Longitudinal Survey of Middle-aged and Elderly Persons (LSMEP) provided by the Ministry of Health, Labour and Welfare. [7] The LSMEP provides panel data on family structure, employment status, and social activities of a cohort of middle-aged and elderly men and women nationwide who were aged between 50 and 59 at the end of October 2005 and with birth dates from November 1945 to October 1955. In this paper, the dataset from 2005 to 2010 was used, with samples with birth date from November 1945 to March 1947. While the survey provides information about health, labor and welfare measures, it does not include information about the residential area,

[7]See the website at (http://www.mhlw.go.jp/english/database/db-ls/ls.html) for details on the Longitudinal Survey of Middle-aged and Elderly Persons.

and so the analyzed periods were set to be as short as possible so as to avoid an observation from experiencing a change in the residential area. As explained in detail in section 4, one group with the birth dates from November 1945 to March 1946 was compared with a second group with birth dates from April 1946 to March 1947 in order to isolate the groups immediately before and after the introduction of the policy. Table 2 shows the descriptive statistics of the Longitudinal Survey of Middle-aged and Elderly Persons for the first wave of male respondents with birth dates from November 1945 to March 1946 and the second wave from April 1946 to March 1947.

Additionally, a secondary data set, the Preference Parameters Study (PPS) provided by the Osaka University Institute of Social and Economic Research,[8], was used to conduct the same analysis as a validity check on the main results. The PPS is mainly conducted for calculating parameters of preferences defining utility functions, that is, time preference, risk aversion, habit formation, externality. The panel survey has been conducted every year since 2004. The surveyed individuals are men and women aged 20-69. This survey is conducted by a self-administered placement method. In this paper, I use the dataset from 2003 to 2013, with only the samples whose birth year is between 1941 and 1950. The response rate is 71.1 percent in 2003. This panel data are suitable for this study because the data include the labor force participation around age 60 with respect to the observations born between 1941 and 1950. For this secondary analysis, I compared one group born between 1941 and 1945 with another born between 1946 and 1950. Table 3 shows the descriptive statistics of the Preference Parameters Study for first wave (1941-45 birth year) and sixth wave (1946-50 birth year) male respondents.

In Japan, there is a dataset focusing only on surveying the elderly whose name is the Japanese Study of Aging and Retirement (JSTAR),[9] which is a panel survey of elderly people

[8]See the website at (http://www.iser.osaka-u.ac.jp/surveydata/engpanelsummary.html) for details on the Preference Parameters Study.

[9]See the website at (http://www.rieti.go.jp/en/projects/jstar/) for details on the JSTAR.

aged 50 or older conducted by the Research Institute of Economy, Trade and Industry, Hitotsubashi University, and, more recently, the University of Tokyo. However, the JSTAR has been conducted since 2007, which means that the labor participation information before age 60 is not available for the elderly whose birth year is around 1945. As a result, I use the Preference Parameters Study. This dataset is the most suitable panel data for this study. In section 5, which presents main results in this paper, I use the Preference Parameters Survey. However, I use the JSTAR in section 6 to discuss the results. I explain what data from the JSTAR I use in section 6.

Finally, in section 6, I use another dataset which is the Fact-finding Survey on the Work Conditions among Small and Medium-sized Enterprises (Chusho kigyo rodo jijo jittai chosa) conducted by the National Federation of Small Business Associations.[10] The surveyed firms are drawn from the firms whose number of employees is below 300. This survey is conducted by a self-administered placement method, resulting in repeated cross-section data. With respect to firms whose number of employees is above 300, public repeated cross section data are not available. In this survey, information about the mandatory retirement policy among small and medium-sized enterprises is available. Additionally, there is no panel data of Japanese firms at present. With respect to the Study of Employment in Small Companies, I also explain which data I use in section 6.

[10]See the website at (https://ssjda.iss.u-tokyo.ac.jp/Direct/gaiyo.phpeid0407langeng) for details on the Fact-finding Survey on the Work Conditions among Small and Medium-sized Enterprises.

Table 2: Summary statistics (Longitudinal Survey of Middle-aged and Elderly Persons, Male Elderly by Birth Cohort)

	1945.11-1946.3	1st wave		1946.4-1947.3	2nd wave	
Variable	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Age	59.166	0.372	452	59.594	0.491	1568
Living with a Child	0.912	0.283	433	0.723	0.448	1442
Living with a Parent	0.367	0.483	433	0.328	0.47	1568
Married	0.744	0.437	441	0.756	0.429	1465
More Than University	0.212	0.409	416	0.222	0.416	1444
Work	0.898	0.303	440	0.863	0.344	1463
Self-Employment	0.218	0.413	395	0.258	0.438	1262
Own House	0.847	0.36	438	0.862	0.345	1458
Financial Asset (8 million yen (80 thousand dollars) < =)	0.33	0.471	342	0.316	0.465	1249
Financial Asset (0 million yen < =, < 8 million yen)	0.389	0.488	342	0.392	0.488	1249
Financial Asset (-3 million yen < =, < 0 million yen)	0.108	0.311	342	0.11	0.314	1249
Financial Asset (< -3 million yen (-30 thousand dollars))	0.173	0.378	342	0.181	0.385	1249

Table 3: Summary statistics (Preference Parameters Study, Male Elderly by Birth Cohort)

	1941-1945	1st wave		1946-1950	6th wave	
Variable	**Mean**	**Std. Dev.**	**N**	**Mean**	**Std. Dev.**	**N**
Age	60.057	1.349	422	59.813	1.362	524
Living with a Parent	0.106	0.312	47	0.199	0.4	196
The Number of Children	2.043	0.942	46	2.01	1.015	193
Married	0.936	0.247	47	0.898	0.303	196
Less Than High School	0.723	0.452	47	0.228	0.42	189
Junior College or Vocational School	0.106	0.312	47	0.529	0.5	189
More Than University	0.17	0.38	47	0.243	0.43	189
Work	0.766	0.428	47	0.85	0.358	187
Self-Employment	0.485	0.508	33	0.406	0.493	155
Own House	0.894	0.312	47	0.939	0.24	196
Residential Asset (< 10 million yen (100 thousand dollars))	0.105	0.311	38	0.271	0.446	177
Residential Asset (10 million yen < =, < 50 million yen)	0.737	0.446	38	0.548	0.499	177
Residential Asset (50 million yen < =)	0.158	0.37	38	0.181	0.386	177
Financial Asset (< 7.5 million yen (75 thousand dollars))	0.472	0.506	36	0.429	0.497	156
Financial Asset (7.5 million yen < =, < 50 million yen)	0.444	0.504	36	0.487	0.501	156
Financial Asset (50 million yen < =)	0.083	0.28	36	0.083	0.277	156
Live in Hokkaido Area	0.021	0.146	47	0.056	0.231	196
Live in Tohoku Area	0.149	0.36	47	0.071	0.258	196
Live in Kanto Area	0.277	0.452	47	0.306	0.462	196
Live in Koshinetsu Area	0.021	0.146	47	0.061	0.24	196
Live in Hokuriku Area	0.085	0.282	47	0.026	0.158	196
Live in Tokai Area	0.128	0.337	47	0.122	0.329	196
Live in Kinki Area	0.149	0.36	47	0.173	0.38	196
Live in Cyugoku Area	0.085	0.282	47	0.071	0.258	196
Live in Shikoku Area	0.021	0.146	47	0.036	0.186	196
Live in Kyusyu Area	0.064	0.247	47	0.077	0.267	196

4 Estimation Procedure

In this section, I explain the difference-in-differences type estimation procedure used in this study. The analysis sample consists of observations from the LSMEP dataset with birth dates between November 1945 and March 1947. From the analysis sample, a control group and treatment group are constructed. As the EESL was introduced in April 2006, the control group with birth dates from November 1945 to March 1946 who turned age 60 before April 2006 are not protected by the EESL. The treatment group born from April 1946 to March 1947 turned 60 just after April 2006 and so are protected by the EESL. These two groups are compared using a difference-in-differences type method by estimating the equation below. As information about residential area is not available in the provided data, this analysis does not control for residential characteristics.

$$y_{it} = \beta_0 + \lambda_t + \sum_{k=0}^{5} \alpha_k 1\{age_{it} = 60 + k\} 1\{birthdate_i \in \{\text{1946 April}, \text{1946 May}, ..., \text{1947 March}\}\} \quad (7)$$
$$+\gamma' x_{it} + a_i + \epsilon_{it}$$

where y_{it} is an indicator equal to one when a respondent works at period t. λ_t is a time fixed effect. a_i is an individual fixed effect. x_{it} are control variables at period t. x_{it} include the respondent's age, family structure, whether a respondent arrives at their basic pensionable age (flat rate part) and the amount of assets.

In addition to the main analysis, an additional analysis was conducted with a secondary dataset, the Preference Parameters Study, using observations with birth years between 1941 and 1950. The following equation was estimated using a difference-in-differences type method.

$$y_{it} = \beta_0 + \lambda_t + \sum_{k=0}^{5} \alpha_k 1\{age_{it} = 60 + k\} 1\{1946 \leq birthyear_i \leq 1950\} \qquad (8)$$
$$+\gamma' x_{it} + a_i + \epsilon_{it}$$

This analysis estimates the difference in labor force participation after age 60 between those born between 1941 and 1945 and between 1946 and 1950. The coefficient of interest is β_1.

I estimate equations (7) and (8) by separating the observations into the following groups. The results are reported in the next section. The analysis focuses on groups directly affected by the EESL, with Groups 1 and 2 from the LSMEP those working full-time at a firm at age 60 and those with mandatory retirement at age 60, respectively. Group 3 from the PPS are non-self-employed workers during the first period of the analyzed periods:

- Group 1 (LSMEP): elderly males born from November 1945 to March 1946 who are working full-time at a firm during the the first wave of the survey (October 2005) versus elderly males born from April 1946 to March 1947 who are working full-time at a firm during the second wave;

- Group 2 (LSMEP): elderly males born from November 1945 to March 1946 who are working full-time at a firm during the first wave of the survey and who are subject to mandatory retirement at age 60 versus elderly males born from April 1946 to March 1947 who are working full-time at a firm during the second wave and are subject to mandatory retirement at age 60;

- Group 3 (PPS): elderly males born from 1941 to 1945 who are both working and not self-employed during the survey first wave versus elderly males born from 1946 to 1950 who are both working and not self-employed during the sixth wave of the survey.

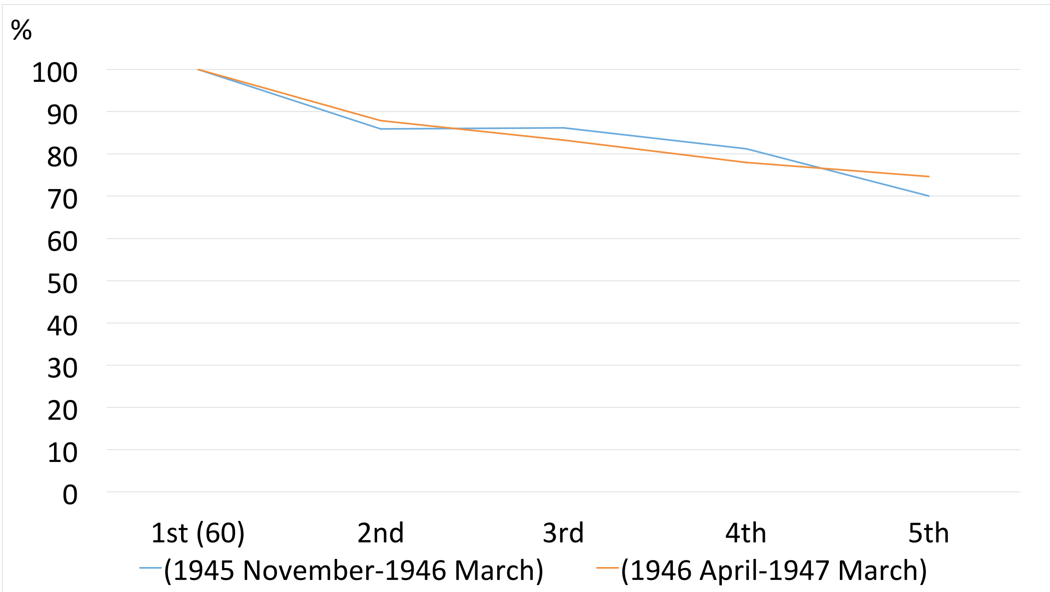
Figure 3 (a) shows the labor force participation trend over time of the control group (those men born from November 1945 to March 1946 and working full-time at a firm during the first wave) and the treatment group (those men born from April 1946 to March 1947 and working full-time at a firm during the second wave). “1st (60)” indicates the fiscal year in which each group becomes age 60. According to Figure 3 (a), the labor force participation of the treatment group is higher than that of the control group at the 2nd. However, it seems that the labor force participation of the control group is higher than that of treatment group after the 2nd wave. Figure 3 (b) shows the same treatment and control groups as in Figure 3 (a) except that only those subject to mandatory retirement at age 60 are included. We can see the same tendency as in Figure 3 (a).

Figures 4 (a) and 4 (b) show a comparison of the labor force participation rate of the cohort born from November 1945 to March 1946 (the control group) and several cohorts of treatment groups born after April 1947. According to Figures 4 (a) and 4 (b), those born from April 1950 to March 1951 have a higher labor force participation rate compared with those born from November 1945 to March 1946, which suggests that there may exist an EESL effect for those born after April 1947.

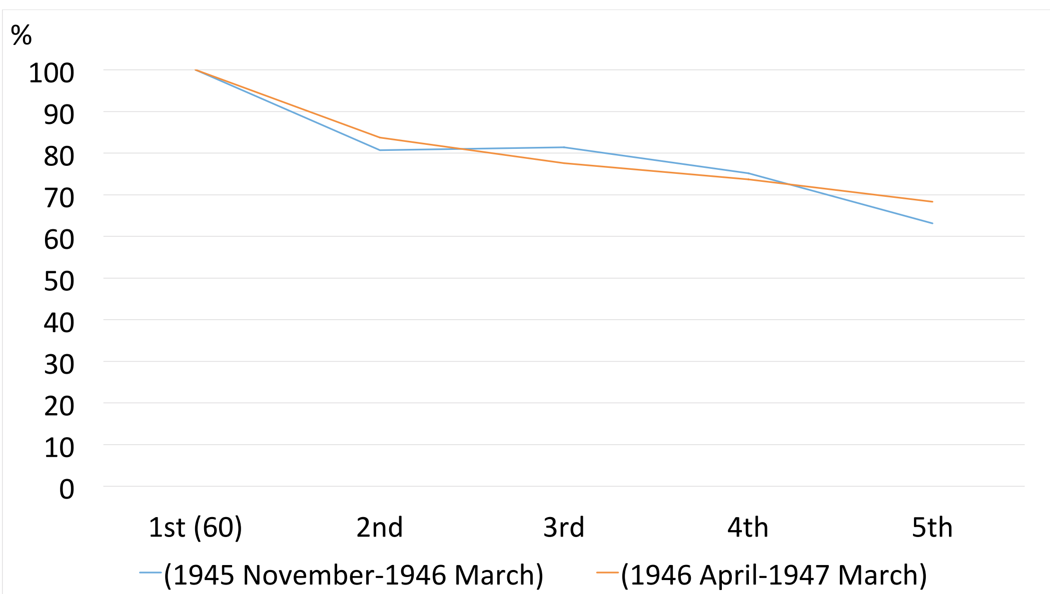
Next, I discuss the labor force participation trend of Group 3 using PPS data. Figure 5 compares the participation rate of the control group, men born from 1941-1945 who were both working and not self-employed during the first wave of the survey, and the treatment group, men born from 1946-1950 who were both working and not self-employed during the sixth wave. We observe that the labor force participation trend is similar for the two groups before age 60 but begins to differ after age 60, which appears that the treatment group has been directly influenced by the EESL policy. This effect is verified in our estimations reported below. Subsequently, I controlled for factors such as respondent demographics and the business cycle, which is the main target of analysis in this paper because government intervention seems to directly influence it.

This paper has certain limitations. With respect to the LSMEP dataset, residential information is not available in the provided data, as discussed above. With respect to the PPS dataset, the Preference Parameters Study asks respondents only respondents' birth year. As a result, the exact age at the time of the interview is unknown. Additionally, whether a respondent arrives at the basic pensionable age (flat rate part) is also unknown. In this paper, I set age = survey year - birth year, and the basic pensionable age (flat rate part) is based on Table 1. However, the birth month is unknown, so the age of pension eligibility is set by birth year A, which is equal to that of people whose birth date is between $A.4.2$ and $A + 1.4.1$. An additional limitation is that the education level of some observations is imputed. When the education level is available in one wave, the same education level is imputed in other waves when the data is not available.

Figure 3: Labor Force Participation



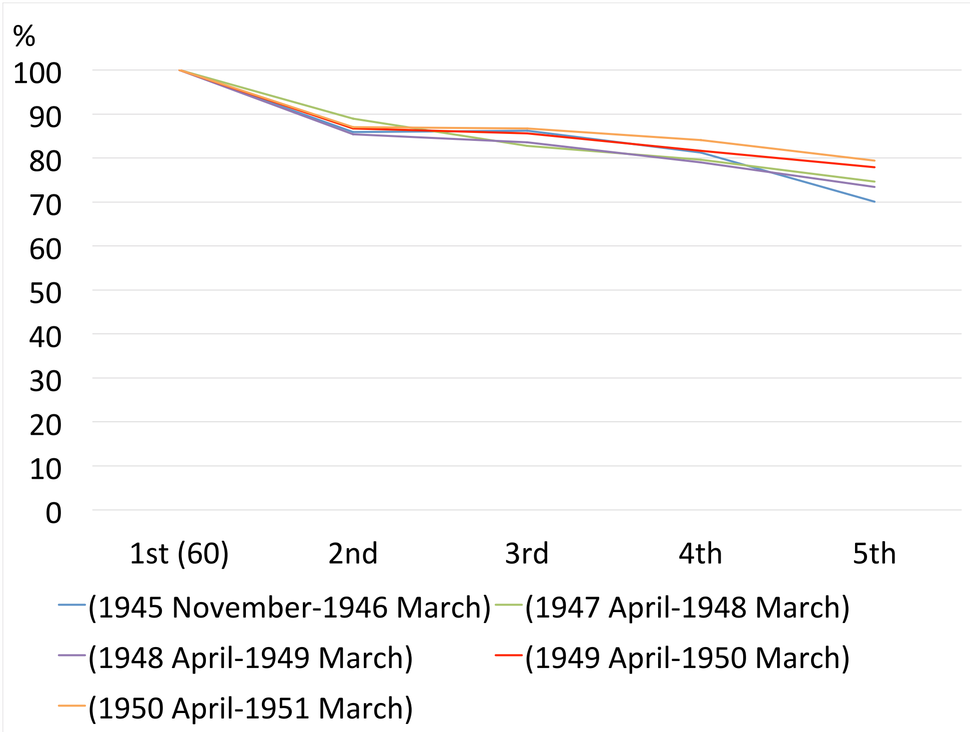
(a) Male, Full-Time Workers



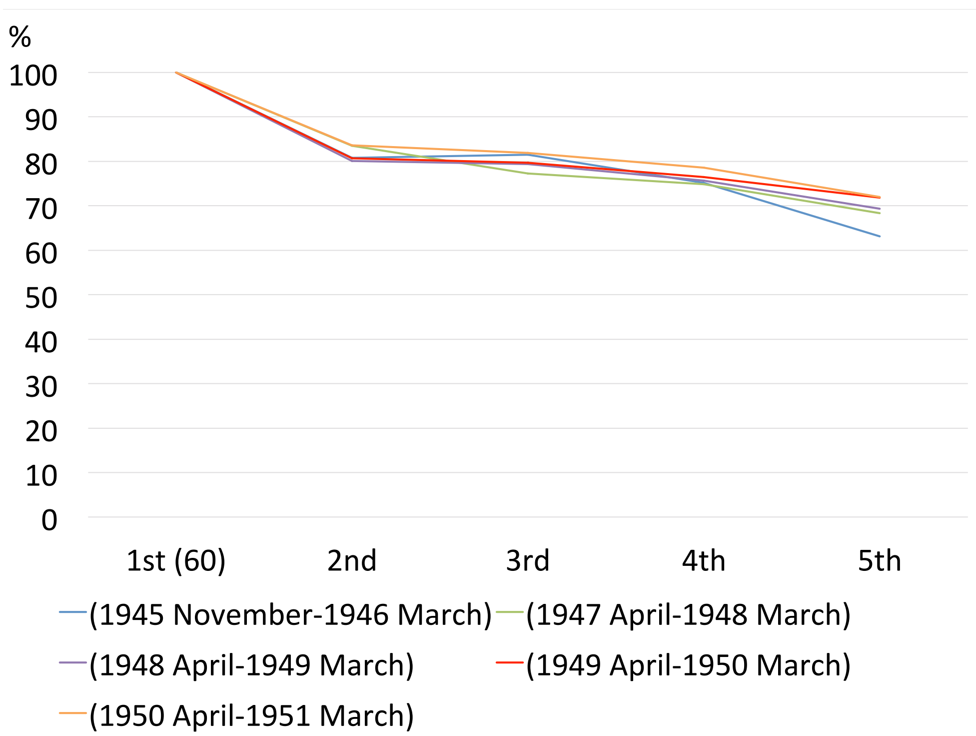
(b) Male, Full-Time Workers with Mandatory Retirement Age of 60

Source: The Longitudinal Survey of Middle-aged and Elderly Persons

Figure 4: Labor Force Participation



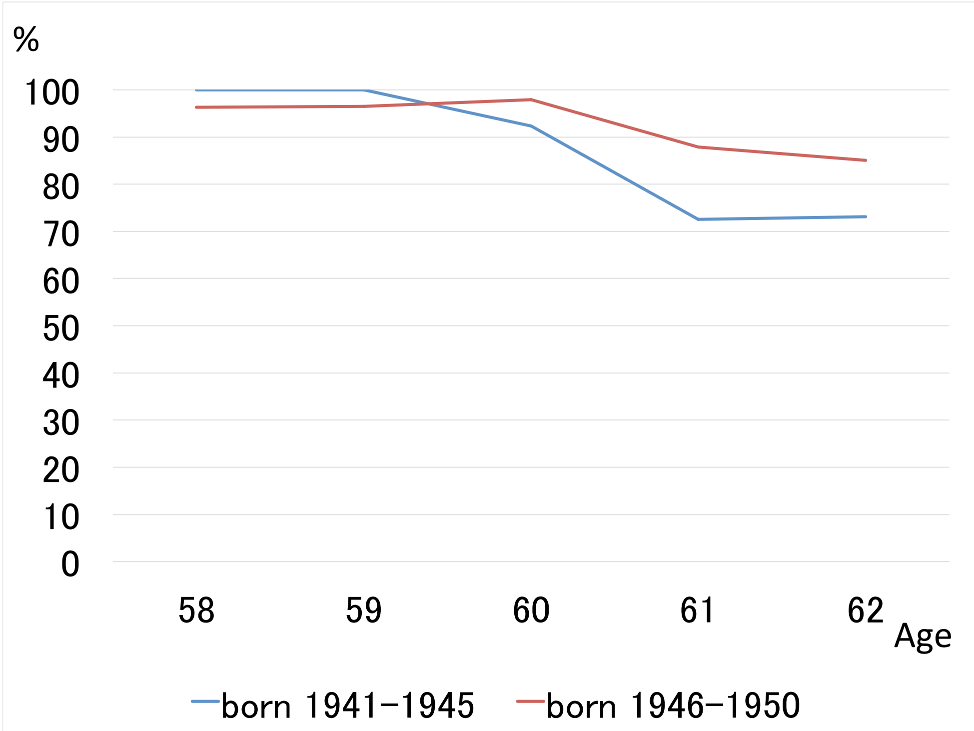
(a) Male, Full-Time Workers



(b) Male, Full-Time Workers with Mandatory Retirement Age of 60

Source: The Longitudinal Survey of Middle-aged and Elderly Persons

Figure 5: Labor Force Participation (Male Workers Not Self-Employed)



Source: The Preference Parameters Study

5 Results

This section discusses the results. First we discuss the cohorts in Group 1 of the LSMEP. According to Table 4, the coefficient at age 60 ($1\{age_{it} = 60\}$ $1\{birthdate_i$ between 1946.4 and 1947.3$\}$) is positively significant in the OLS estimation at ages 61 (0.0437), 64 (0.0906) and 65 (0.0940). However, the coefficients are not significant in the FE estimation. According to Table 4, there exists an upper bias in the OLS coefficients. With respect to the coefficients at ages 61 and 62, these coefficients are negative (age 61: -0.0079, age 62: -0.0206). It appears that there is no effect on labor force participation immediately after the introduction of the EESL. In the FE result, the coefficient is positive only at ages 63 and 64, which seems to not be related to the EESL policy.

Next, we report the results of our estimates of the cohorts of Group 2 of the LSMEP. This analysis focuses only on the samples facing the mandatory retirement at age 60, which means that this group is directly protected by the EESL policy. According to Table 5, in the OLS result, the coefficients at ages 61, 64 and 65 are significantly positive. As in Table 4, all FE coefficients are not significantly positive. The coefficients are negative except at age 64. In this analysis, it seems that there also exists an upper bias in the OLS coefficients. For these cohorts, too, it appears that there is no effect of the EESL on labor force participation immediately after the introduction of the EESL.

Table 6 shows the results of estimates of the cohorts in Group 3 of the PPS. These elderly men also appear to be influenced by the EESL in the same manner as the above groups as shown in Tables 4 and 5. As we can observe, the coefficients of the OLS estimation at ages 60, 61 and 62 are significantly positive. However, the coefficients of the FE estimation are not significant. It should be noted, though, that the sample size is small and, accordingly, the standard errors of the coefficients in the fixed effects specification are quite large. However, the absolute value of the coefficients are comparatively large, which suggests that the EESL may indeed have an effect on the labor force participation rate in the groups born after March

1947.

To put the above results into perspective, I compare the results of this study with those of Kondo and Shigeoka (2016) who, using repeated cross sectional data and not controlling for educational characteristics or other demographic variables, find that the labor force participation rate of salaried workers born in 1946 is significantly larger than that of salaried workers born in 1945 at ages 60 and 61 by 2.4 and 3.2 percent, respectively. This impact is considered small. In this paper, I cannot confirm that there exists any significant increase in the labor participation rate immediately after the implementation of the EESL.

Table 4: Labor Force Participation (Male, Full-Time Workers)

	(1) OLS	(2) FE
$1\{age_{it} = 60\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0239	-0.0058
	(0.0160)	(0.0149)
$1\{age_{it} = 61\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0437*	-0.0079
	(0.0233)	(0.0260)
$1\{age_{it} = 62\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0402	-0.0206
	(0.0307)	(0.0315)
$1\{age_{it} = 63\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0612*	0.0045
	(0.0348)	(0.0378)
$1\{age_{it} = 64\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0906***	0.0203
	(0.0329)	(0.0396)
$1\{age_{it} = 65\}1\{birthdate_i \in \{1946 \text{ April}, 1946 \text{ May}, ..., 1947 \text{ March}\}\}$	0.0940**	-0.0104
	(0.0425)	(0.0471)
N	5267	5267

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

age, family structure, education level, asset level dummies, a dummy of arriving at the basic pensionable age, wave dummies are also included in the estimation model.

The Result from The Longitudinal Survey of Middle-aged and Elderly Persons

Table 5: Labor Force Participation (Male, Full-Time Workers with Mandatory Retirement Age of 60)

	(1) OLS	(2) FE
$1\{age_{it} = 60\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.0352	-0.0157
	(0.0237)	(0.0205)
$1\{age_{it} = 61\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.0550*	-0.0218
	(0.0329)	(0.0373)
$1\{age_{it} = 62\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.0678	-0.0214
	(0.0436)	(0.0451)
$1\{age_{it} = 63\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.0582	-0.0319
	(0.0453)	(0.0502)
$1\{age_{it} = 64\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.1168***	0.0055
	(0.0425)	(0.0515)
$1\{age_{it} = 65\}1\{birthdate_i \in \{1946\text{ April}, 1946\text{ May}, ..., 1947\text{ March}\}\}$	0.1475***	-0.0046
	(0.0549)	(0.0619)
N	3282	3282

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

age, family structure, education level, asset level dummies, a dummy of arriving at the basic pensionable age, wave dummies are also included in the estimation model.

The Result from The Longitudinal Survey of Middle-aged and Elderly Persons

Table 6: Labor Force Participation (Male Workers Not Self-Employed)

	(1) OLS	(2) FE
$1\{age_{it} = 60\}1\{1946 \leq birthyear_i \leq 1950\}$	0.2122**	0.2345
	(0.0861)	(0.1438)
$1\{age_{it} = 61\}1\{1946 \leq birthyear_i \leq 1950\}$	0.1979**	0.1892
	(0.0842)	(0.1408)
$1\{age_{it} = 62\}1\{1946 \leq birthyear_i \leq 1950\}$	0.1716**	0.1687
	(0.0822)	(0.1382)
$1\{age_{it} = 63\}1\{1946 \leq birthyear_i \leq 1950\}$	0.1100	0.1052
	(0.0919)	(0.1282)
$1\{age_{it} = 64\}1\{1946 \leq birthyear_i \leq 1950\}$	0.1286	0.1218
	(0.0915)	(0.1098)
$1\{age_{it} = 65\}1\{1946 \leq birthyear_i \leq 1950\}$	0.0350	-0.0080
	(0.1107)	(0.1132)
N	751	751

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

age, family structure, education level, asset level dummies in the previous period (high, middle), a dummy of arriving at the basic pensionable age, wave dummies, wave × region dummies are also included in the estimation model.

The Result from The Preference Parameters Survey

6 Discussion: What happened after the implementation of EESL?

As previously discussed, no effect of government intervention on the employment of the elderly workers is observed. In this section, I consider why there is no effect of the EESL on the employment of the elderly. To understand the mechanism of this policy effect, we need to consider the channels of this policy effect, as discussed in section 2.2. Consequently, this study must answer the following three questions:

- 1. Did the probability of receiving a reemployment offer increase? (Channel 2)
- 2. Did the number of firms which abolished or increase the mandatory retirement age increase? (Channel 3)
- 3. Did the acceptance rate of reemployment offers decrease due to low wages offer by firms? (Channel 1)

With respect to Channels 2 and 3, the dataset partly implies these facts. On the other hand, with respect to question 3, there is no available data to clarify this point. Table 6 shows whether a firm carries out the employment policy the EESL requires.[11] In 2006 and 2007, almost all firms carries out the necessary employment policy. According to Table 7, most firms obey the EESL by introducing the reemployment system. Subsequently, the first question can be used by using the JSTAR. There is a sharp increase in the ratio of people receiving a reemployment offer after arriving at the mandatory retirement age, as shown in Figure 8. Reemployment offer 1 means the ratio of workers receiving a reemployment offer from the firm where they arrive at the mandatory retirement age. Reemployment offer 2 means the ratio of workers receiving a reemployment offer from the firm where they arrive at the mandatory retirement age or affiliated firms (including Reemployment offer 1). According

[11] See the website at http://www.mhlw.go.jp/stf/houdou/0000101253.html (in Japanese)

to Figure 8, there is a sharp increase in the ratio of workers receiving a reemployment offer after those born in 1946 arrive at the mandatory retirement age. According to Figure 8, the ratio of firms obeying the EESL increases after the workers born in 1946 arrive at the mandatory retirement age. According to this figure (reemployment 2), the ratio of workers receiving the reemployment offer increases by about 10 percent. This approximates the ratio of workers who cannot get the reemployment offer without the EESL. It is possible that the rejection rate for the offers in this group is high. According to Usui et al. (2015), male employees aged 54 gradually move to part-time work or retire after beginning to receive pension. Those who continue to work cannot choose their optimal working hours, although wanting to choose more working hours. Potentially, it is possible that there are some elderly who cannot continue to work, although he/she wants to continue to work if he/she receives a reemployment offer.

With respect to the mandatory retirement age, there is a change in the distribution between 2004 and 2008, at least for firms whose number of employees is below 300. According to Figure 9, the ratio of firms wanting to adopt the mandatory retirement age above 64 in 2004 is larger than that in 2008. Figure 10 shows the distribution of the mandatory retirement age with respect to male full-time workers at age 60 in the LSMEP, and the same tendency is observed.

According to these facts, firms change the employment policy after 2006 by making reemployment offers or increasing the mandatory retirement age. However, according to this paper's results, the employment of the elderly workers does not significantly change after the workers born in 1946 reach 60.[12] The firms have obeyed the government directions; introducing the reemployment system, abolishing the mandatory retirement age, or increasing the mandatory retirement age. The analysis of Channel 1 is important for understanding what happened after the implementation of the EESL, thus providing scope for future work. It

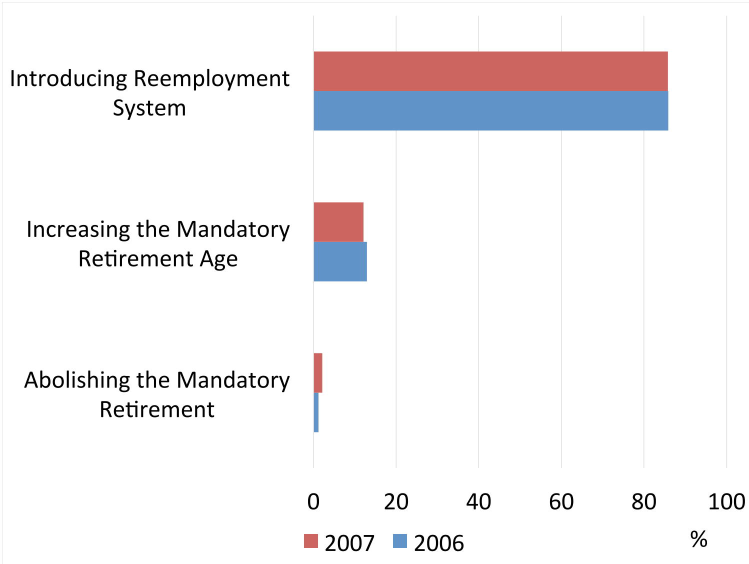
[12] According to the literature's result, there is a significant effect. However, the effect is small.

is possible that the firms tried to reduce the cost of obeying the EESL by decreasing wages after the mandatory retirement age when they engage in a contract with the workers reaching mandatory retirement age. There is no clear statement with respect to wage contracts when a firm gives a reemployment offer to a worker. Kondo (2016) finds a decline in earnings of the elderly workers who reached age 60 after 2006. This evidence is based on only observable wage. The offered wage when making a contract of reemployment after the implementation of the EESL should be thus analyzed. It is possible that some workers reject an offer because the offered wage is too low.

Figure 6: The Ratio of Firms Preparing the Employment Measures for the Elderly

Source: The Employment of the Elderly Workers, Ministry of Health, Labor and Wealfare

Figure 7: The Ratio of the Employment Measures for the Elderly (All Firms Preparing the Employment Measures)

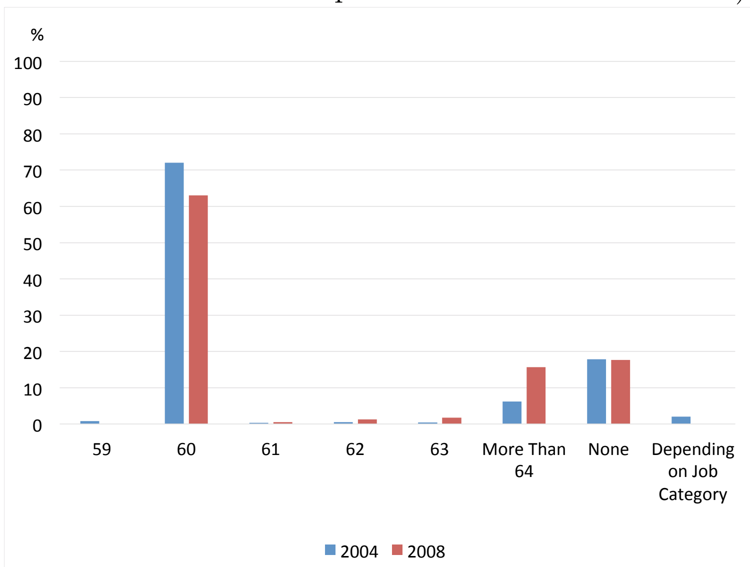


Source: The Employment of the Elderly Workers, Ministry of Health, Labor and Wealfare

Figure 8: The Ratio of Receiving Reemployment Offers

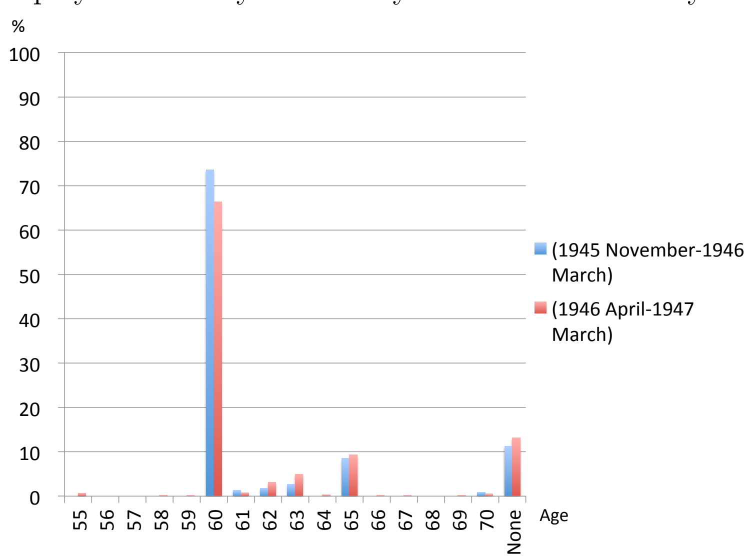
Source: JSTAR

Figure 9: Employment Policy for Elderly Workers in 2004 and 2008: Mandatory Retirement Age (Only Small and Medium Sized Enterprises: Less than 300 workers)



Source: Fact-finding Survey on the Work Conditions among Small and Medium-sized Enterprises (Chusho kigyo rodo jijo jittai chosa)

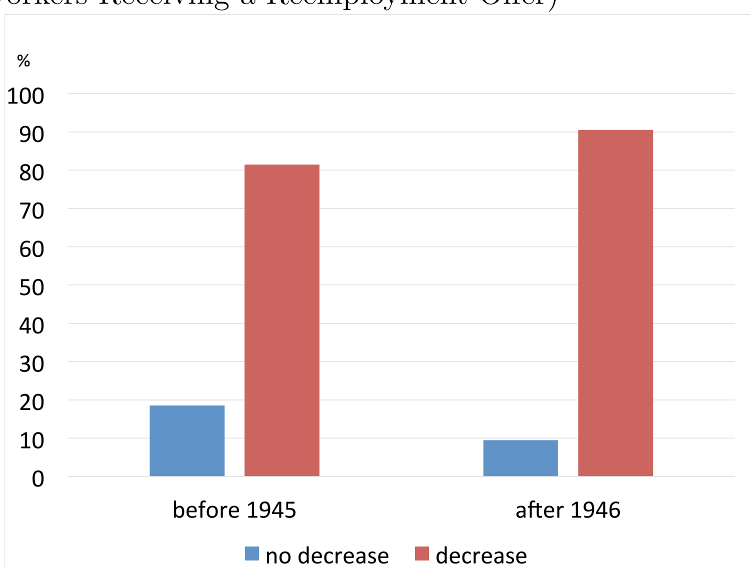
Figure 10: Employment Policy for Elderly Workers: Mandatory Retirement Age



Source: The Longitudinal Survey of Middle-aged and Elderly Persons

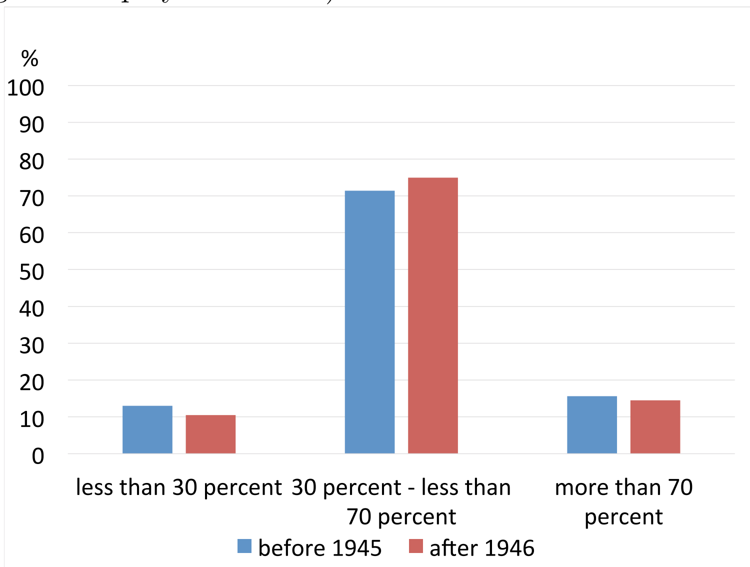
Finally, I identify the changes in the wage contract when a worker receives a reemployment offer from the firm where he/she reaches the mandatory retirement age. Figure 11 shows whether the worker's wage decreases or not after reemployment. This figure also shows the ratio of the worker's wage change after reemployment. According to this figure, the ratio of receiving a decreased wage after reemployment increases by 10 percent after a worker born after 1946 reaches the mandatory retirement age. However, it is unclear whether this is due to the EESL. As such, I compare the workers born in 1945 with those born in 1946. However, the sample size is insufficient with respect to only workers born around 1945 and 1946. Additionally, figure 12 shows the distribution of the wage decrease rate when receiving a reemployment offer. According to this figure, there is an increase in the ratio of the wage decrease rate of between 30 percent and 70 percent. However, this is also not for the dataset which includes only workers born around 1945 and 1946.

Figure 11: The Ratio of Whether Wage Decreases After Reemployment (Reemployment Contract)(Only Workers Receiving a Reemployment Offer)



before 1945: birth year before 1945, after 1946: birth year after 1946
Source: JSTAR

Figure 12: The Ratio of Wage Decrease After Reemployment (Reemployment Contract)(Only Workers Receiving a Reemployment Offer)



before 1945: birth year before 1945, after 1946: birth year after 1946
Source: JSTAR

7 Conclusion

This paper analyzed the effect of government intervention in the demand side of the labor market on the employment of the elderly. However, the results showed that there is no significant effect of the EESL on the employment of the elderly immediately after the implementation of the EESL. According to the discussion in section 6, firms obey the government's directions, thus introducing the reemployment system, abolishing mandatory retirement, or increasing the mandatory retirement age. This suggests that firms attempted to reduce additional costs caused by the government policy by choosing actions that the government does not prohibit. As a result, the number of reemployment offers has increased after the implementation of the EESL.

However, there is no clear statement in the law with respect to wage contracts when a firm makes a reemployment offer to a worker reaching the mandatory retirement age. As explained in section 6 , many firms introduce the reemployment system or increase the mandatory retirement age without abolishing the mandatory retirement age. When a firm makes a reemployment contract with a worker, it discusses a wage decrease rate with the worker who intends to work in the firm after the mandatory retirement age. The law does not concretely enforce a certain wage rate. The following question is important for directly analyzing the reason why the employment of elderly workers has not increased: did the acceptance rate of reemployment offers decrease due to low wages offered by firms? (Channel 1)

Specifically, the effect on Channel 1 is worth mentioning. This study showed there was no positive effect on the employment of the elderly immediately after the implementation of the EESL. However, it is possible that firms might have decreased the offered wage because of the new requirement to provide a re-employment offer. After the mandatory retirement age, firms can offer wage rates that are not strictly regulated and as a result, they have incentives to decrease the offered wage. This is a possible topic for future research. In addition, the preliminary findings of this study suggest that the effect of the EESL with respect to groups born after April 1947 should be analyzed, and this is another possible topic for future research.

References

Adams, S. J. (2002). Passed Over for Promotion Because of Age: An Empirical Analysis of the Consequences. *Journal of Labor Research*, 23(3):447–461.

Adams, S. J. (2003). Age discrimination legislation and the employment of older workers. *Labour Economics*, 11:219–241.

Alan, G. L. and Thomas, S. L. (1986). A Structural Retirement Model. *Econometrica*, 54(3):555–584.

Ashenfelter, O. and Card, D. (2002). Did the Elimination of Mandatory Retirement Affect Faculty Retirement? *The American Economic Review*, 92(4):957–980.

Burkhauser, R. V. and Ouinn, J. F. (1983). Is Mandatory Retirement Overrated? Evidence From The 1970s. *The Journal of Human Resources*, 18(3):337–358.

Clark, R. L. and Ogawa, N. (1992). The Effect of Mamdatory Retirement on Earnings Profiles in JAPAN. *Industrial and Labor Relations Review*, 45(2):258–266.

Fields, G. S. and Mitchell, O. S. (1984). Economic Determinants of The Optimal Retirement Age: An Empirical Investigation. *The Journal of Human Resources*, 19(2):245–262.

Gruber, J. and Wise, D. A. (1998). Social Security and Retirement: An International Comparison. *American Economic Review, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association*, 88(2):158–163.

Hirsch, B. T., Macpherson, D. A., and Hardy, M. A. (2000). Occupational Age Structure and Access for Older Workers. *Industrial and Labor Relations Review*, 53(3):401–418.

Hurd, M. D. (1990). Research on the Elderly: Economic Status, Retirement, and Consumption and Saving. *Journal of Economic Literature*, 28(2):565–637.

Hutchens, R. (1988). Do Job Opportunities Decline with Age. *Industrial and Labor Relations Review*, 42(1):89–99.

Johnson, R. W. and Neumark, D. (1997). Age Discrimination, Job Separations, and Employment Status of Older Workers. *The Journal of Human Resources*, 32(4):779–811.

Kondo, A. (2016). Effects of increased elderly employment on other workers employment and elderlys earnings in Japan. *IZA Journal of Labor Policy*, 5(2):1–23.

Kondo, A. and Shigeoka, H. (2016). The Effectiveness of Demand-side Government Intervention to Promote Elderly Employment : Evidence from Japan . *Industrial and Labor Relations Review, forthcoming*.

Lahey, J. (2008). State Age Protection Laws and the Age Discrimination in Employment Act. *Journal of Law and Economics*, 51(3):433–460.

Lang, K. (1989). Why was there mandatory retirement? *Journal of Public Economics*, 39:127–136.

Lazear, E. P. (1979). Why Is There Mandatory Retirement? *Journal of Political Economy*, 87(6):1261–1284.

Lazear, E. P. (1981). Agency, Earnings Profiles, Productivity, and Hours Restrictions. *The American Economic Review*, 71(4):606–620.

Lazear, E. P. and Moore, R. L. (1984). Incentives, Productivity, And Labor Contracts. *The Quarterly Journal of Economics*, 99(2):275–296.

Motegi, H., Nishimura, Y., and Terada, K. (2016). Des Retirement Change Lifestyle Habits? *The Japanese Economic Review*, 67(2):169–191.

Neumark, D. (2003). Age Discrimination Legislation in the United States. *Contemporary Economic Policy*, 21(3):297–317.

Neumark, D. and Song, J. (2013). Do stronger age discrimination laws make Social Security reforms more effective? *Journal of Public Economics*, 108:1–16.

Neumark, D. and Stock, W. A. (1999). Age Discrimination Laws and Labor Market Efficiency. *The Journal of Political Economy*, 107(5):1081–1125.

Slade, F. P. (1987). Retirement Status and State Dependence: A Longitudinal Study of

Older Men. *Journal of Labor Economics*, 5(1):90–105.

Staubli, S. and Zweimüller, J. (2013). Does raising the early retirement age increase employment of older workers ? *Journal of Public Economics*, 108:17–32.

Usui, E., Shimizutani, S., and Oshio, T. (2015). Are Japanese Men of Pensionable Age Underemployed or Overemployed? *RIETI Discussion Paper Series*, 15-E-099.

von Wachter, T. (2002). The End of Mandatory Retirement in the US: Effects on Retirement and Implicit Contracts. *CENTER FOR LABOR ECONOMICS UNIVERSITY OF CALIFORNIA, BERKELEY WORKING PAPER*, 5.

Chapter 3: Effects of Informal Elderly Care on Labor Supply: Exploitation of Government Intervention on the Supply Side of Elderly Care Market

1 Introduction

Many developed countries have been facing problems of a decreasing birthrate and an aging population. As population ages, the cost of social security and social welfare increases, eroding the country's budget. As such, numerous developed countries have reformed the social security systems to reduce the cost of social security and social welfare, thus generating a fair amount of attention towards these policy reforms. Countries such as the United States, the United Kingdom, and Korea have decided to increase the pension eligibility age in subsequent decades, while Japan has already increased it. As population ages in developed countries, countries such as Germany and Korea have also been reformed the nursing care system for the elderly. In Germany, a mandatory and universal system of long-term care insurance (LTCI) was implemented in 1995 (Schulz (2010)). The national mandatory elderly LTCI was introduced in Korea in 2008 (Kwon (2009), Won (2013) and Chul et al. (2015)).

With the growing interest in nursing care systems in the United States and Europe, since the 1980s, both demand and supply side of the elderly care market have been analyzed. One important topic in the analysis of the demand side of the elderly care market is the effect of informal care on labor supply. As we explain in section 2, hitherto, related studies in the United States and Europe analyzing the effect of informal care on labor supply have employed family structure and parental health as instrumental variables. As such, they have not utilized institutional change as a natural experiment in estimating the effect of informal care on labor supply. As Van Houtven, Coe, and Skira (2013) point out, some of the instruments employed in literature are weak or their exogeneity is questionable.

In 2000, the Japanese government has also implemented LTCI.[1] In the Japanese care system, there are two important characteristics related to our study. First, there are three types of public nursing homes. Second, the supply of these nursing homes is regulated by the government. The goal of this study is to examine the causal effect of informal care on labor supply, and the analysis utilizes the exogenous variation of government intervention on the supply side of the elderly care market to estimate this effect. Since the supply of public nursing home is regulated by the government, we utilize this exogenous variation for estimating the effect of informal care for the elderly on labor supply. To the best of our knowledge, there is hitherto no study to utilize the exogenous variation of nursing home supply regulated by the government as an instrument to estimate the effect of informal care for the elderly on labor supply. Kondo (2016) utilizes the exogenous variation of nursing home capacity. However, Kondo (2016) does not estimate the effect of informal care on labor supply, and includes directly the capacity of nursing home as an explanatory variable, estimating directly the effect of this capacity on labor supply. In Japan, there are also some studies analyzing the effect of LTCI introduction on labor supply, while they do not directly estimate the effect of informal care on labor supply.[2] According to our results, the effect of informal care for elderly on female labor supply is negative. On the other hand, there is no effect of informal care on male labor supply, since, in Japan, females spending more time on informal care than males spending time on informal care. As such, the government intervention becomes effective for decreasing the female burden of informal care.

The remainder of this paper is organized as follows: section 2 reviews literature; section

[1] Tamiya et al. (2011) explain this system in detail.
[2] For example, Shimizutani et al. (2008), Sugawara and Nakamura (2014), and Fukahori et al. (2015)

3 discusses the data uses; section 4 explains the institutional background and instruments used in this study; section 5 discusses gender differences in providing informal care; section 6 discusses the analysis methods; section 7 presents the results, which are discussed in section 8 discusses; and section 9 concludes this paper and identifies the scope for future research.

2 Literature Review

Since the 1980s, the elderly care market has been analyzed from both supply and demand sides.[3] One of the central topics regarding the demand side of the elderly care market is the effect of informal care on labor supply. Lilly et al. (2007) and Bauer and Sousa-Poza (2015) review studies on the effect of informal care on labor supply in detail,[4] which is beyond the scope of this study.[5]

After 2000, analysis on the effect of informal care on labor supply has also been carried out. The most important issue in these studies is controlling the endogeneity of providing informal care, followed by which instruments the studies should employ. In Table 1, we review which instruments have been employed in the literature after 2000. As Van Houtven et al. (2013) point out, some of the instruments employed in literature are weak or their exogeneity is questioned. Some other studies use other techniques, such as simultaneous equations or dynamic panel data methods, without using the instrumental variables methods. However, the causal influence of exogenous variation on providing informal care cannot is unavailable in these studies. As Table 1 shows, in literature, variables such as parental health and family structure have been used as instrumental variables and no study utilizes institutional

[3]For example, the literature analyzing the supply side of the care market is represented by Nyman (1985, 1988, 1994), Gertler (1989, 1992), Connelly (1992), Norton (1992), Ettner (1993), Cohen and Spector (1996), Grabowski (2001), Grabowski et al. (2008), and Ching et al. (2015).

[4]For example, the related literature includes Wolf and Soldo (1994), Hoerger et al. (1996), Carmichael and Charles (1998, 2003), Heitmueller and Inglis (2007), Carmichael et al. (2010), Lilly et al. (2010), Leigh (2010), Michaud et al. (2010)

[5]Additionally, public health is represented by studies such as Tan (2000), Berecki-gisolf et al. (2008) Hassink and Berg (2011) Trong and Brian (2014). However, we focus on the economics literature.

exogenous variation. Therefore, we propose the estimation procedure to utilize the exogenous variations causal influence on providing informal care.

As previously mentioned, in Japan, the supply side of elderly care market is regulated by the government. Since 2000, the LTCI system has been introduced in Japan. The government has also determine how many public nursing homes to be supplied, thus exogenously controlling the supply of public nursing homes. Additionally, there is an exogenous variation of this supply of public nursing homes depending on municipality. In other words, the availability of formal care is heterogeneous among different municipalities. We utilize this exogenous variation to estimate the effect of informal care on labor supply.

Finally, we introduce the Japanese literature. Since 2000, Japanese researchers have analyzed the effect of informal care on labor supply. However, Shimizutani et al. (2008), Sugawara and Nakamura (2014), Fukahori et al. (2015) and Kondo (2016) do not estimate the direct effect of informal care on labor supply, which Wakabayashi and Donato (2005), Ishii (2015), Yamada and Shimizutani (2015) and Moriwaki (2016) do. Nonetheless, the later do not utilize the exogenous variation caused by the exogenous change in the supply side of the informal care market. Additionally, the magnitude seems inconsistent across. We compare the results of these studies with our results in section A.1.

Table 1: The Instruments Employed in Literature

	Instruments	Memo
Main		
Heitmueller (2007)	•The number of sick and disabled people in the household •The age of the three closest friends of the respondent •The age of the parents and the geographical proximity of parents and friends	
Bolin et al (2008)	•Mother have bad health •Father have bad health •Age of mother •Age of father •Mother lives far away •Father lives far away •Mother deceased •Father deceased •Number of siblings	
Ciani (2012)	•The presence of disabled individuals living in the household •The presence of at least one co-resident individual reporting poor health	
Van Houtven, Coe and Skira (2013)	•Mother ill •Mother in-law ill •Mom died •Dad died •Mother in-law died •Father in-law died •Mother recently widowed •Mother in-law recently widowed	
Meng (2013)	•The four categories of ADL and IADL in which the impaired individual needs help are used as instruments •The variable which indicates whether disabled individuals are present in the household	
Others		
Van Houtven and Norton (2004)	•Proportion of daughters •Distance to the nearest child	•2nd stage dependent variable: the utilization of formal care
Bonsang (2009)	•Number of siblings	•2nd stage dependent variable: the utilization of formal care

3 Data

We use the Japanese Study of Aging and Retirement (JSTAR), [6] which is a panel survey of elderly people aged 50 or older conducted by the Research Institute of Economy, Trade and Industry, Hitotsubashi University, and, more recently, the University of Tokyo. The JSTAR has been conducted since 2007 has survey counterparts in other countries, such as the China Health and Retirement Longitudinal Study (CHARLS), the English Longitudinal Survey on Aging (ELSA), the Health and Retirement Study (HRS) in the US, the Korean Longitudinal Study of Aging (KLoSA), the Longitudinal Aging Study in India (LASI), and the Survey on Health, Aging, and Retirement in Europe (SHARE). Ichimura et al. (2009) explain the details of the JSTAR, such as the sampling design and other detailed information on the survey.

There are three types of JSTAR data, which differ by security level: high, very high, and ultra-high. Our study uses the very high level, which contains the full sample data, including birth month and geographic information, which allows us to identify the nursing home capacity for each municipality. The survey years used in the study are 2007, 2009, 2011 and 2013. The JSTAR includes a rich variety of variables that capture the characteristics of individuals — their economic and health status, family background, and social and work status. In the JSTAR, labor participation, informal care to the parents, respondent demographics, and the place of residence information are available for the elderly. As such, this dataset is a suitable panel data for this study. Generally, we used the Harmonized JSTAR data set. [7] However, when variables were not available in the Harmonized JSTAR, we used the original JSTAR.

[6] See the website at (http://www.rieti.go.jp/en/projects/jstar/) for details on the JSTAR.

[7] The Gateway to Global Aging Data (http://gateway.usc.edu) provides harmonized versions of data from the international aging and retirement studies (e.g., HRS, ELSA, SHARE, and JSTAR). All variables of each dataset aim to have the same items and follow the same naming conventions. The harmonized datasets enable researchers to conduct cross-national comparative studies. The program code for generating the Harmonized JSTAR dataset from the original JSTAR dataset is provided by the Center for Global Aging Research, USC Davis School of Gerontology, and the Center for Economic and Social Research (CESR). Some variables, such as measures of assets and income, are imputed by this code.

Table 2 shows the summary statistics of the data. For this analysis, we impute the asset-level data by replacing missing data with the substituted values of a respondent as explained it in section A.2. We use a similar imputation method to the RAND HRS. (Hurd et al. (2016))

We also use the Population Census of 2005 and 2010 [8] and the Survey of Institutions and Establishments for Long-Term Care for 2007, 2009, 2011, and 2014 to define the instrumental variables for this study. [9] We explain how to use these datasets in section 4.

[8] See the website at (http://www.stat.go.jp/english/data/kokusei/) for details on the Population Census.

[9] See the website at (http://www.mhlw.go.jp/english/database/db-hss/siel-index.html) for details on the Survey of Institutions and Establishments for Long-term Care.

Table 2: Summary Statistics

	(1) **5 cities**		(2) **2 cities**		(3) **3 cities**	
	mean	sd	mean	sd	mean	sd
Demographics						
Age	62.87	7.05	62.99	7.32	62.64	6.86
Age $\geq$ PA	0.58	0.49	0.55	0.50	0.52	0.50
Educ. $\geq$ Univ.	0.12	0.33	0.16	0.36	0.24	0.42
Female	0.50	0.50	0.53	0.50	0.55	0.50
Mariage	0.81	0.39	0.75	0.43	0.78	0.41
N um.of children	2.05	0.97	2.16	1.39	1.70	1.07
Economic variables						
HH income (US$)	41641	32794	44800	38118	61081	53419
Own house	0.77	0.42	0.63	0.48	0.69	0.46
Saving(imputed,US$)	63934	87361	54559	90933	91997	111633
Working status						
Not working for pay	0.43	0.49	0.51	0.50	0.43	0.49
Working hours ≥ 5	0.55	0.50	0.46	0.50	0.52	0.50
Working hours ≥ 10	0.52	0.50	0.44	0.50	0.49	0.50
Working hours ≥ 20	0.48	0.50	0.41	0.49	0.43	0.50
Full time worker (at 1st intw or age 54)	0.49	0.50	0.46	0.50	0.49	0.50
Nursing care and parents' information						
Provide informal care	0.13	0.33	0.11	0.31	0.15	0.35
Formal care utilization (for most severe parent)	0.10	0.30	0.14	0.34	0.14	0.34
NCL (for most severe parent) $\geq$ S1	0.23	0.42	0.24	0.43	0.30	0.46
NCL (for most severe parent) $\geq$ C3	0.13	0.34	0.13	0.34	0.15	0.36
Parents age(for most severe parent)	87.99	5.89	88.55	6.39	88.13	5.52
Year of 1st interview	2007		2009		2011	
Num. of waves	4 waves		3 waves		2 waves	

4 Institutional Background

Since the implementation of the LTCI system in 2000, all Japanese people above 40 have to join the LTCI and are able to receive public care services depending on their age and nursing care level. All those between 40 and 64 can receive public care services with a co-payment ratio of only 10 percent when they have specific diseases due to aging. On the other hand, those above 65 can receive public care services with a co-payment ratio of 10 percent when they "require long-term care." The government assesses the nursing care level for the elderly to decide whether they "require long-term care." As a result, public care services are provided based on the nursing care level as exemplified below for those over 65. Figure 1 shows the process to determine which nursing care level is to be provided. [10]

- Step 1: A family member who finds that an elderly individual in the household has a physical problem can ask the local government to decide the nursing care level.

- Step 2: Depending on the health condition of the elderly and household characteristics, such as the number of adults who can provide informal care, the local government decides the nursing care level, based on which, the choice set of available public care services from which an applicant can choose is determined. For example, the applicant can use a particular nursing home when they have more than nursing care level 1. [11] The following table 3 shows the nursing care level as per Moriwaki (2016). [12] We quote Table 1 from Moriwaki (2016).

  More importantly, there are two judgment procedures (the first and second) to determine the nursing care level. In the first judgment procedure, the computer automatically carries out the first judgment based on standardized information. In the second

[10] We sincerely thank Hisataka Anezaki and Tetsuya Iwamoto for explaining this point.

[11] After 2015, this restriction became effective. Before 2015, the restriction was referring to more than nursing care level 1.

[12] With respect to the decision of nursing care level, see the website at (http://www.mhlw.go.jp/topics/kaigo/nintei/gaiyo2.html) for details. (Ministry of Health, Labour and Welfare)(in Japanese)

Table 3: Care Levels (Table 1 in Moriwaki (2016))

Care Level	Description
Special Elders	Currently independent, needs preventive healthcare
Support Level 1	Having difficulties in standing up, getting up, and/or standing on one foot
Support Level 2 Care Level 1	In addition, having difficulties in walking, washing body, keeping track of the personal finances, and/or clipping nails
Care Level 2	In addition, having difficulties in dressing, moving, and/or decision-making
Care Level 3	In addition, having difficulties in washing face, grooming, tooth-brushing, urination/defecation, and/or use of public transportation
Care Level 4	In addition, having difficulties in eating, and/or communication
Care Level 5	In addition, having difficulties in swallowing, memorizing and/or understanding

procedure, academic experts judge the final nursing care level referring to special report from a doctor. In this report, information about the household of an applicant might be included. The judgment about the nursing care level is influenced by this information on the household, except for the applicant's health status. Additionally, after the nursing care level has been decided, an applicant can apply for a reexamination based on the situation of the applicant's household.

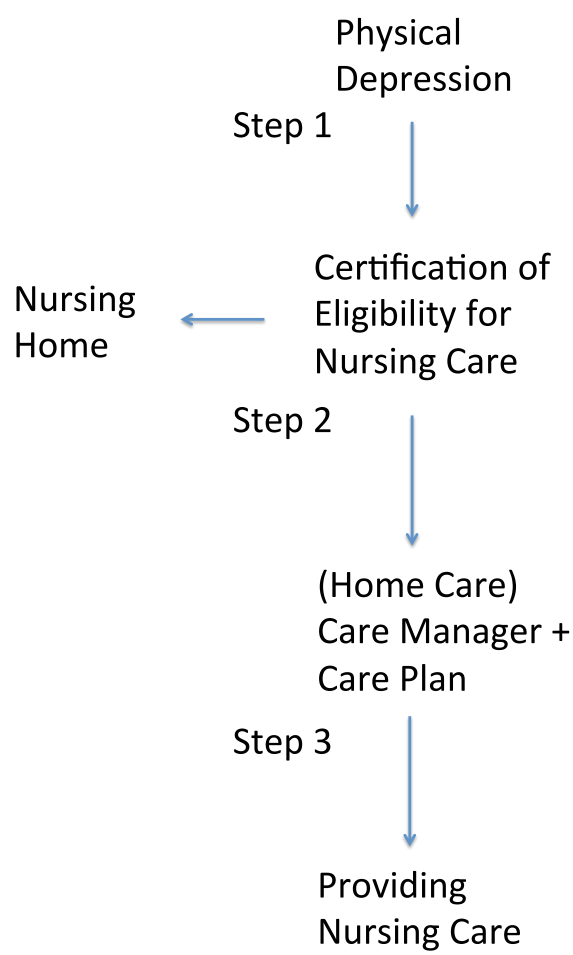
- Step 3: Finally, if an applicant decides to use a home care, they will discuss with a care manager [13] with respect to which care service they will use.

According to the explanation above, in Step 2, an applicant can stay in a public nursing home when their nursing care level is above a certain level. There are three public nursing homes in Japan as per Table 4, [14] Facility Covered by Public Aid Providing Long-Term Care to the Elderly (Tokuyo), Long-Term Care Health Facility (Roken), and Designated Medical Long-Term Care Sanatoriums. In these three public nursing homes, Tokuyo is the most popular nursing home because the price of nursing care is relatively low. As you can observe in Table 4, its utilization rate is almost 100 percent. Basically, most elderly individuals are provided nursing care in Tokuyo or Roken. Additionally, the purpose of each nursing home

[13] A care manager is a specialist who plans the care service that an applicant will use.

[14] See the website at (http://www.mhlw.go.jp/english/database/db-hss/siel-index.html) for details. (Ministry of Health, Labour and Welfare)

Figure 1: The Process to Determine Which Nursing Care is Provided



is different. The allowed length of stay in Tokuyo is unlimited, while in Roken is from three months to one year. The purpose of Roken is to provide the services that help with the rehabilitation of the elderly. The Designated Medical Long-Term Care Sanatoriums are not that common for providing nursing care for the elderly.

Table 4: Three Public Nursing Homes in Japan

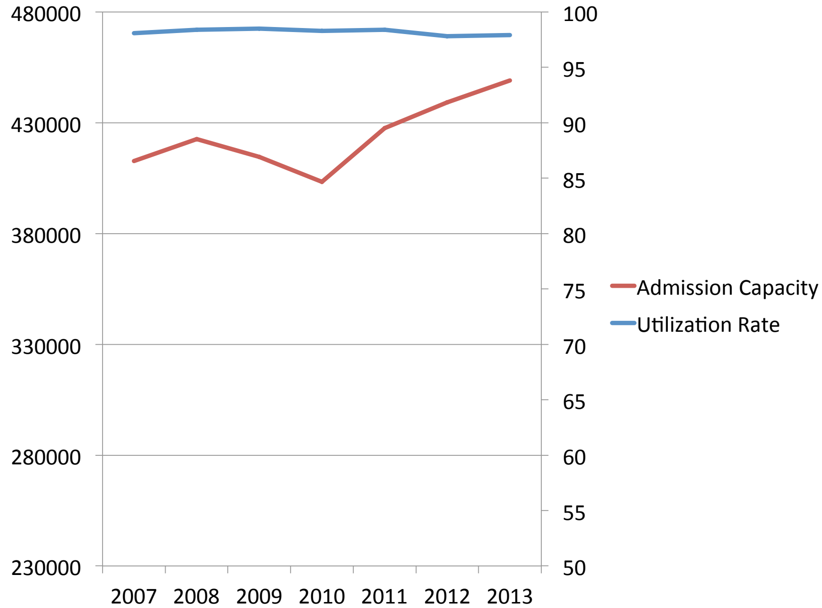
	Facility Covered by Public Aid Providing Long-Term Care to the Elderly (Tokuyo)	Long-Term Care Health Facility (Roken)	Designated Medical Long-Term Care Sanatoriums
Number of Facilities	7065	3857	1318
Admission Capacity	484353	339142	58419
Utilization Rate	97.4	89.2	91.1
Average Nursing Care Level	3.87	3.26	4.38

Source: Survey of Institutions and Establishments for Long-term Care October, 2015

The important point is that these three nursing homes for the elderly are exogenously supplied by the government on the elderly in the demand side of the elderly care market. For example, as you can see in Table 4, the numbers receiving care services in Tokuyo are close to the upper bound of capacity. We thus utilize this exogenous variation of the capacity for controlling the endogeneity of providing informal care. In Figure 2, we show the admission capacity and utilization rate of Tokuyo. Obviously, although admission capacity changes exogenously, the utilization rate does not change (almost 100 percent). The ratio of people who must provide informal care is influenced by the exogenous change of the admission capacity.

In fact, there is an exogenous variation of the admission capacity in different regions and over different periods. In Figure 3, we show the admission capacity per capita for those above 65 for Tokuyo as 100 $\times$ (Capacity of Tokuyo in Each Region)/(Total Population More Than Age 65 in Each Region)) in each region. Here, we use the Population Census 2005 and 2010 and the Survey of Institutions and Establishments for Long-term Care 2007, 2009,

Figure 2: Admission Capacity and Utilization Rate of Tokuyo in Japan



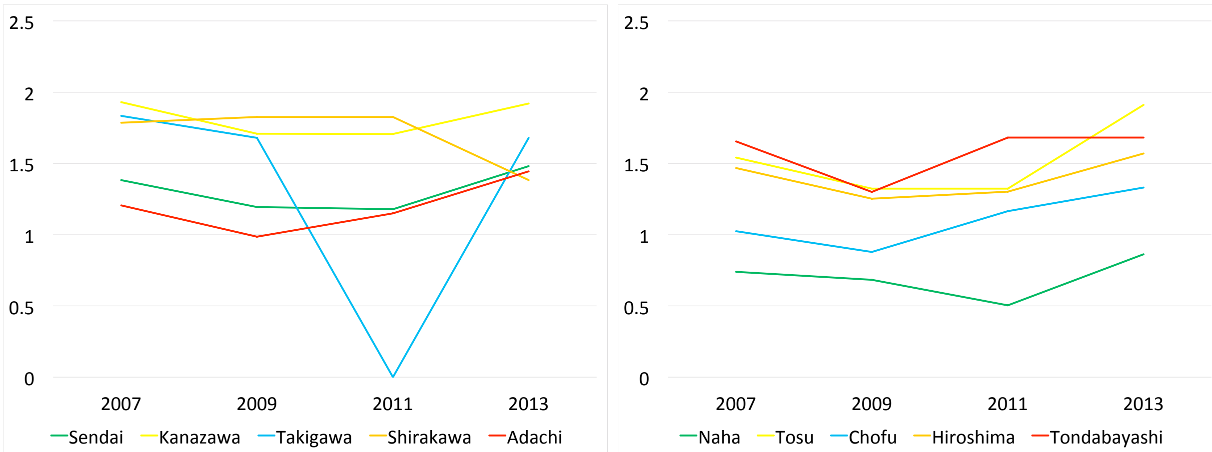
Source: Survey of Institutions and Establishments for Long-term Care October, 2007-2013

2011, and 2014 to build this variable. [15] Here, the variation in the value is exogenous for a caregiver in a household, which we use this variation to control the endogeneity of informal care. Importantly, a household cannot use the nursing home outside the region of residence.

Before 2015, the requirement to apply for admission to Tokuyo is being categorized above nursing care level 1. Moreover, the elderly with a higher nursing care level, who are difficult to give a nursing home care to, were preferably assigned to public nursing homes, although this was not stipulated. We show the utilization rate of formal care by care level in Figure 4. The utilization includes the usage of public and private nursing home. According to Figure 4, the utilization rate increases as the nursing care level increases. In fact, as Table 4 shows, the average nursing care level in Tokuyo was above 3 in 2005. We also use the nursing care level of parents in addition to the exogenous variation of public nursing care home designing the instrumental variable.

[15] We only have 2005 and 2010 population information, and use the information nearest to the surveyed year of capacity.

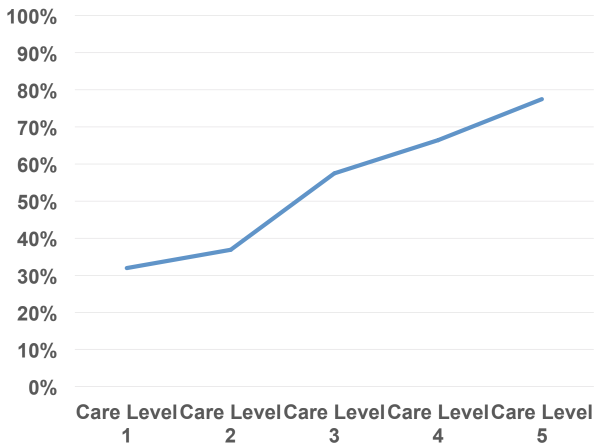
Figure 3: Admission Capacity Per Capita More Than Age 65 of Tokuyo in Japan (Vertical Line: 100 × (Capacity of Tokuyo in Each Region)/(The Total Population More Than Age 65 in Each Region))



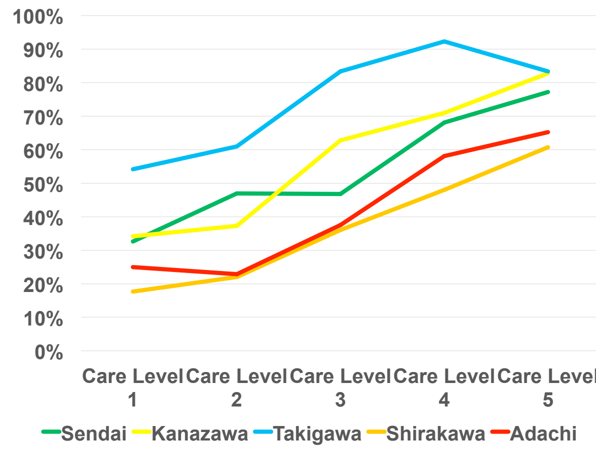
Source: Survey of Institutions and Establishments for Long-term Care October, 2007, 2009, 2011, 2014 and Census 2005 and 2010

According to Figure 5, formal care utilization strongly influences the decision of providing informal care in the household. Figure 5 shows the distribution of who provides informal care in a household with parents certified as being above care level 1. In a household utilizing formal care, the ratio of both male and female members not providing informal care is high. Here, we also use instruments related to government intervention on the supply side of the care market, such as dummy variables indicating the number of parents certified as more than care level 1. The cross term of (parental age) × (dummy variable indicating more than support level 1) is also used. The details are explained in section 6.2.

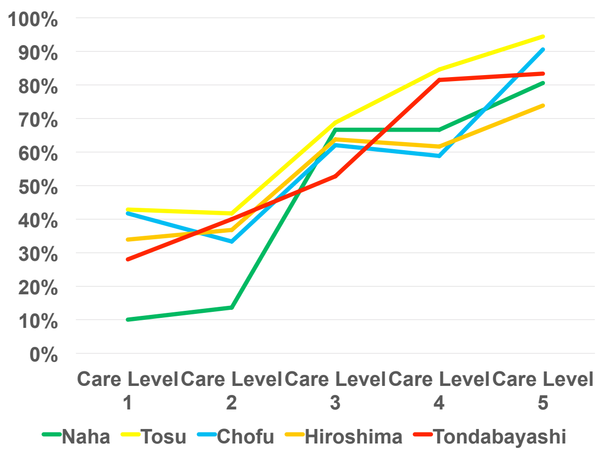
Figure 4: The Utilization Rate of Formal Care by Care Level (Total and By City, City: The Residence of a Respondent)(Horizontal Line: Nursing Care Level of Parents)



(a) Total



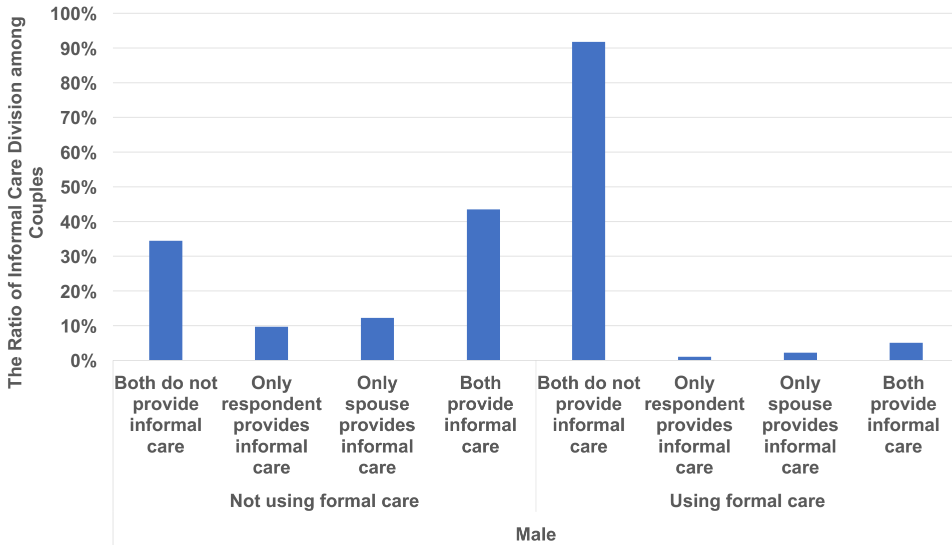
(b) Sendai, Kanazawa, Takikawa, Shirakawa and Adachi



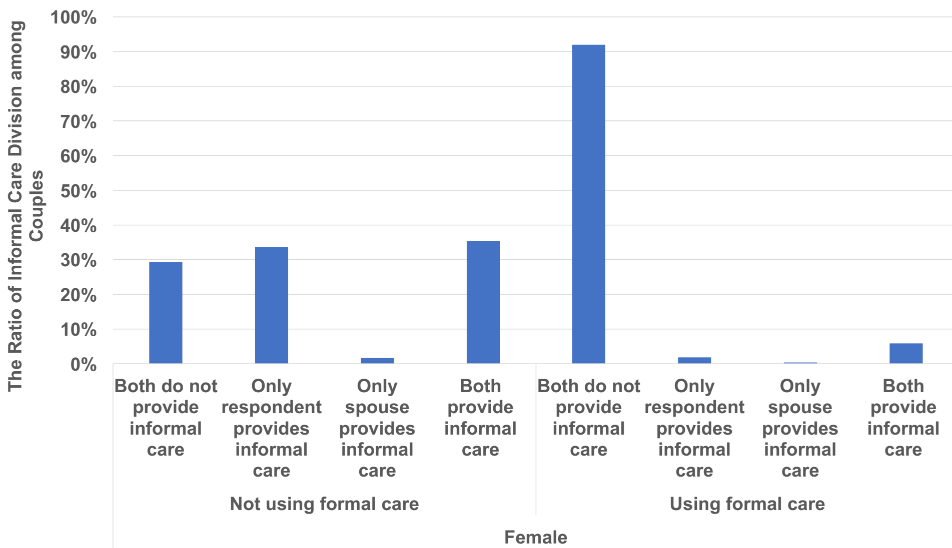
(c) Naha, Tosu, Chofu, Hiroshima and Tondabayashi

Source: JSTAR 2007-2013

Figure 5: Formal care utilization and informal care provision among couples with certified parents



(a) Male



(b) Female

Source: JSTAR 2007-2013

5 Discussion: Gender Difference in the Role of Providing Informal Care

Before we empirically analyze the effect of informal care on labor supply, we must discuss who provides nursing care in a household and difference in the role of providing nursing care between male and female household members, which is critical in Japan, and which we confirm here. According to the discussion in this section, we should consider the heterogeneity of male and female household members when considering the estimated results.

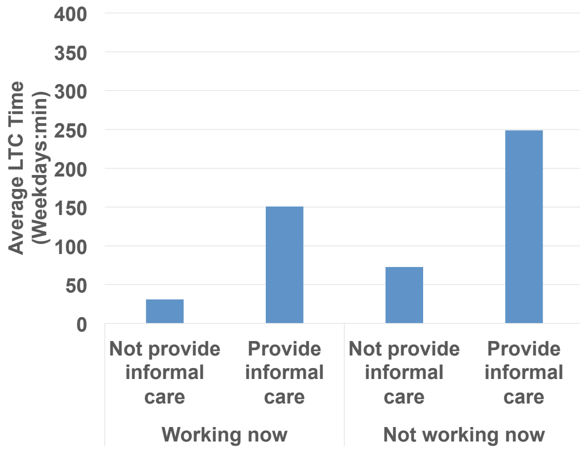
Figure 6 shows long-term care time by gender, which is significantly longer for females than males. Long-term care time for working females is even longer than for males who do not work, which reflects in the estimated results. Next, we focus on long-term care time, depending on whether other household members help or not and whether husbands works or not. According to Figure 6 (c), as expected, long-term care time for females without support is longer than otherwise. Additionally, whether a husband is working or not does not influence long-term care time. Accordingly, when household members have to provide informal care for the elderly, the task is concentrated on a female household member. Figure 6 (e) and (f) shows male household behavior. Figure 6 (e) shows whether male spouses help with providing informal care. Even if the husband is not working for pay, the ratio of the husband helping the wife is 70 percent. On the other hand, the ratio is 60 percent if the husband is working for pay. Overall, husbands are not helping their wives in about 30 percent of households. Figure 6 (f) shows long-term care time for males people when their spouse provides informal care. When not working, the difference in long-term care time is about one hour compared to the case when males do work.

We discuss the relationship between labor force participation rate and informal care. Figure 7 describes the proportion of not working for pay. Basically, the labor force participation rate of males is higher than that of females. In panel (a), the difference in labor force participation rate is about 5 percent between elderly providing informal care and those who are not providing informal care (both female and male). Figure 7 (c), (d), and (e) shows the relationship between the transition of providing informal care and of not working for pay. According to panels (b), (c), (d), and (e), among males, providing informal care in the second interview seems to influence their labor force participation rate. Almost all males work in the first wave. For females, providing informal care in the second interview seems to influence the labor force participation rate, regardless of the working status in the first wave.

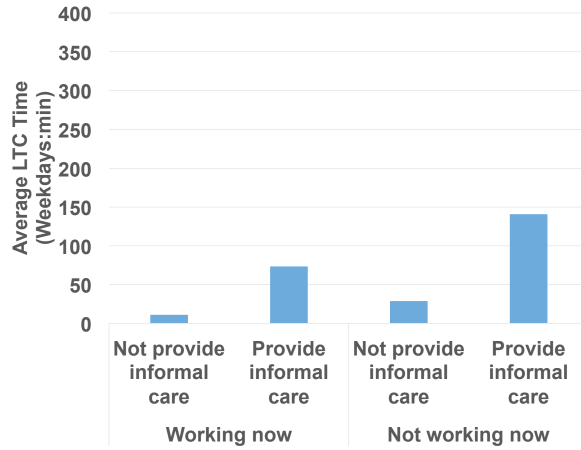
In panel (d), we find that the female elderly continue to work even if they provide nursing care in the second interview. One reason might be that almost all people can use home care services covered by nursing care insurance. Since, in JSTAR, the information with respect to home care services is not available, we use information from the Comprehensive Survey of Living Conditions 2013. [16] Figure 8 shows the long-term care service utilization covered by nursing care insurance when a person who requires nursing care lives in a household. The care service includes home-visiting nursing care services, meal delivery service, and so on. According to Figure 8, most children and their spouses utilize these services when parents require long-term care services. The rate of utilization does not seem to be related to the work status since most dependents can utilize the service.

[16] See the website at (http://www.mhlw.go.jp/english/database/db-hss/cslc-index.html) for details. (Ministry of Health, Labour and Welfare)

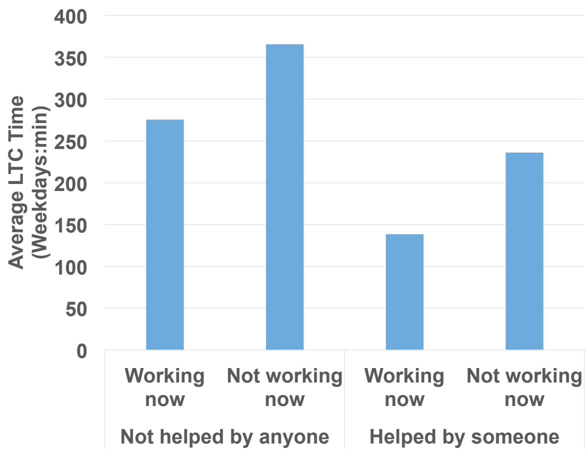
Figure 6: Long Term Care Time



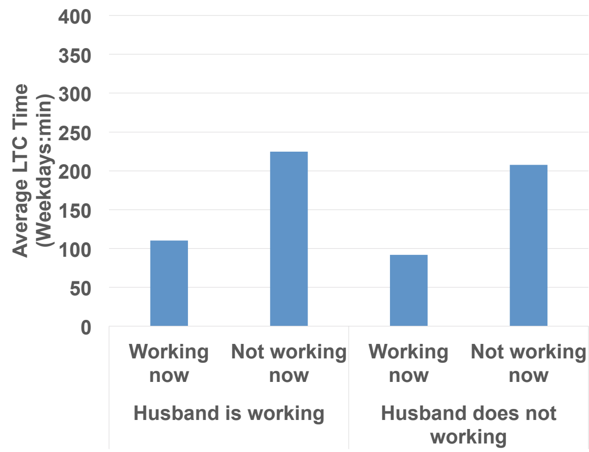
(a) Female



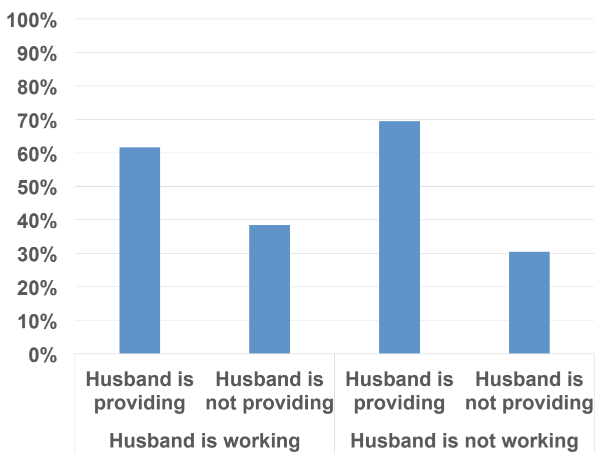
(b) Male



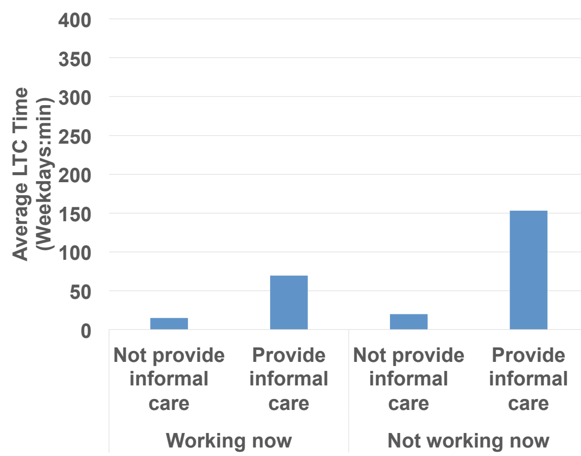
(c) Female by any help



(d) Female by husband' working status



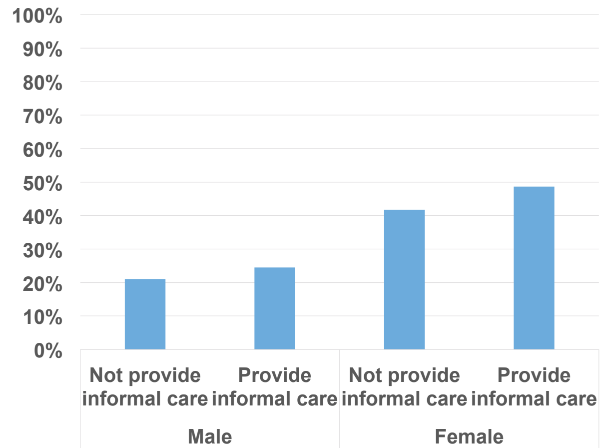
(e) Husband' help for female providing LTC



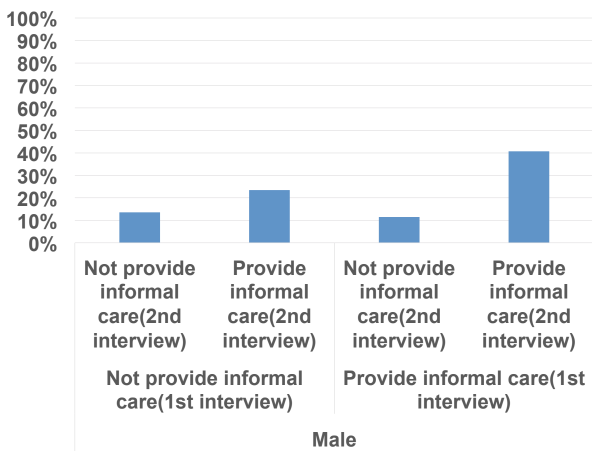
(f) Male who have wife providing LTC

Source: JSTAR 2007-2013

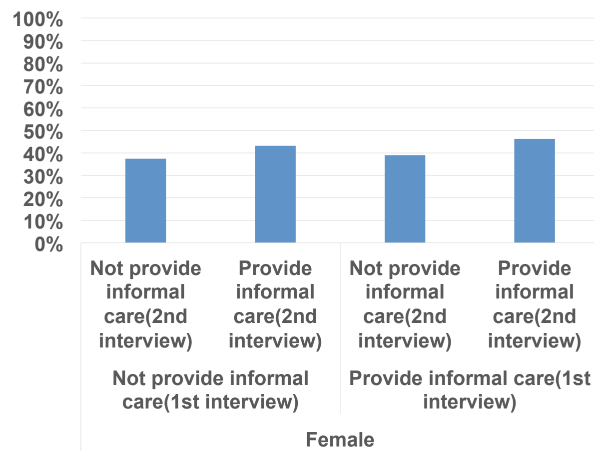
Figure 7: The Proportion of Not Working For Pay



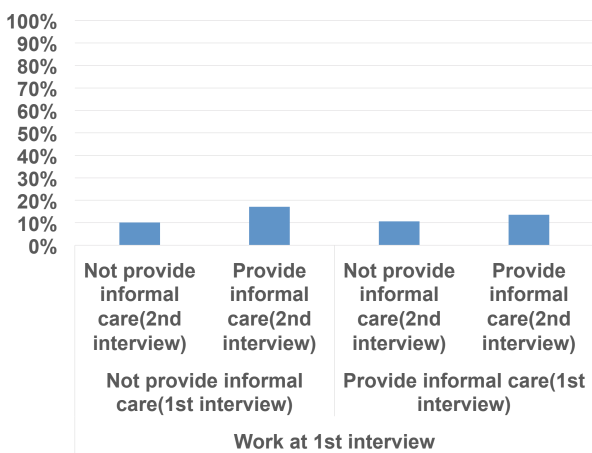
(a) Total



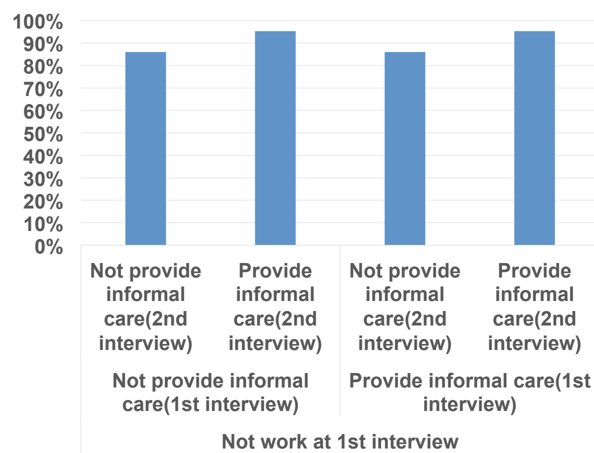
(b) LFP at 2nd INTW by 1st INTW status (Male)



(c) LFP at 2nd INTW by 1st INTW status (Female)



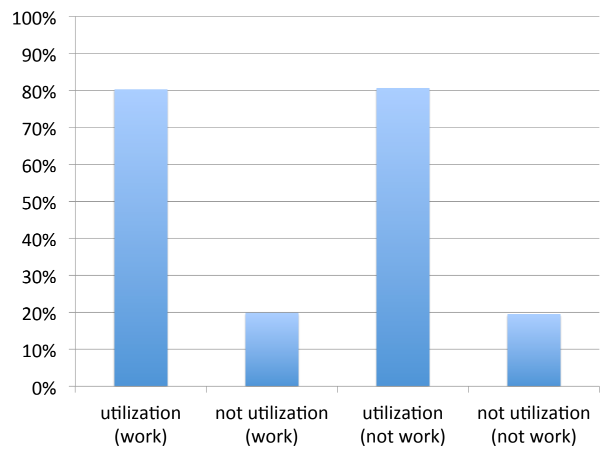
(d) Female working at 1st interview



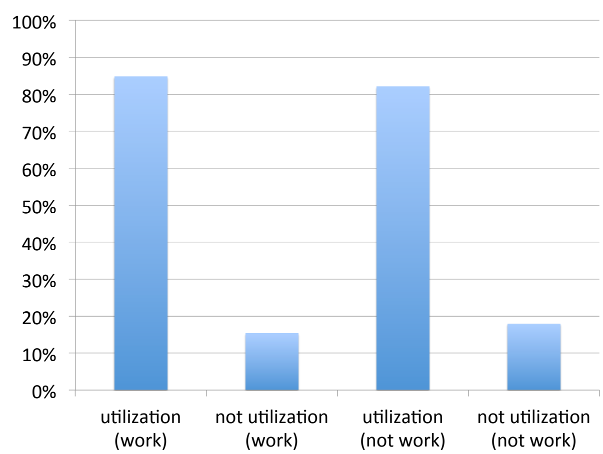
(e) Female not working at 1st interview

Source: JSTAR 2007-2013

Figure 8: The Long Term Care Service Utilization Covered by Nurse Care Insurance When a Person Who Requires Nursing Care Lives in the Household



(a) children (relationship with the member requiring nursing care)



(b) children's spouse (relationship with the member requiring nursing care)

Source: Comprehensive Survey of Living Conditions 2013

6 Analysis Method

6.1 Relationship between Labor Supply and Informal Care

We discuss the division of informal care in a household by using a simple economic model, confirming its relationship with the labor division. By using this model, we will consider the causal relationships (1) between labor supply and informal care, (2) between informal care and formal care utilization, (3) between formal care utilization and spouse informal care, and (4) between spouse informal care and informal care. Figure 9 confirms these relationships. The following is a household collective model, including the division of informal care. [17]

$$\max_{\{c^A,l^A,c^B,l^B,I,\alpha,Apply\}} \mu(w,y,z)u(c^A,l^A)+(1-\mu(w,y,z))u(c^B,l^B)$$

$$s.t.\ c^A+c^B+w^Al^A+w^Bl^B \leq \tag{1}$$

$$y^H+w^A(T^A-\alpha\cdot I)+w^B(T^B-(1-\alpha)\cdot I)-p\cdot\tilde{F}$$

$$CareSum \geq C\cdot 1\{ParentalHealth=Bad\}, \tag{2}$$

$$0\leq T^A-l^A-\alpha\cdot I,\ \ 0\leq T^B-l^B-(1-\alpha)\cdot I, 0\leq I, 0\leq l^j(j=A,B)$$

We add the variables such as $\alpha, I, \tilde{F}$ in the usual collective model. There are two agents in this household (agent A and agent B). The notations are following.

- $c^j(j=A,B)$: consumption, $\tilde{l}^j(j=A,B)$: finally consumed leisure.
- $l^j(j=A,B)$: leisure, I: quantity of informal care.
- $T^j(j=A,B)$: endowment, $w=(w^A,w^B)$: wage vector.
- F: supplied formal care amount from the government
- p: formal care price, C: needed care amount if a parent is not healthy.

[17] With respect to collective household models, please see Vermeulen (2002).

- y_H: household income except wage

$Apply, Availability \in \{0,1\}$. $ParentalHealth \in \{Good, Bad\}$. We define $CareSum, Labor^A, Labor^B, \tilde{F}$ in the following way.

$$CareSum = I + \tilde{F} \tag{3}$$

$$Labor^A = T^A - l^A - \alpha \cdot I \tag{4}$$

$$Labor^B = T^B - l^B - (1-\alpha) \cdot I \tag{5}$$

$$\tilde{F} = F \cdot 1\{Apply = 1\} \cdot 1\{Availability = 1\} \cdot 1\{ParentalHealth = Bad\} \tag{6}$$

$Availability = 1$ if the government supplies formal care to this household. $ParentalHealth = 1$ if one of the parents is not healthy. Equation (4) and (5) show the direct relationship between labor supply and informal care. Equation (3) shows the direct relationship between informal care and formal care utilization, and the direct relationship between formal care utilization and spouse informal care. Finally, the household informal care I is divided into agent A informal care and agent B informal care, which shows the direct relationship between agent A informal care and agent B informal care. When $Availability = 1$ and $ParentalHealth = 1$, a household can utilize formal care according to (6). When $ParentalHealth = 1$, $CareSum \geq C \cdot 1\{ParentalHealth = Bad\}$ is true. In other words, $1\{ParentalHealth = Bad\}$ influences directly household informal care (both agent A and agent B informal care). Summing these relationships, we can describe Figure 9. By the way, the event that $Availability = 1$ and $ParentalHealth = 1$ happens exogenously from the decision making of the household.

Define variables Z_{1it}, Z_{2it} as $Z_{1it} = 1\{Availability = 1\}$ and $Z_{2it} = 1\{ParentalHealth = Bad\}$. Let vector $\tilde{Z}_{3it}$ be other instruments. We use the following equation based on the relationship among labor supply, informal care, formal care utilization, and informal care supply in the household, where $\tilde{X}^j_{it}(j = A, B)$ is an explanatory variable of agent j.

- The Functions:

$$
\begin{aligned}
y^A_{it} &= F_{y^A}(IC^A_{it}, \tilde{X}^A_{it}),\\
y^B_{it} &= F_{y^B}(IC^B_{it}, \tilde{X}^B_{it}),\\
IC^A_{it} &= F_{IC^A}(y^A_{it}, IC^B_{it}, FC_{it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}),\\
IC^B_{it} &= F_{IC^B}(y^B_{it}, IC^A_{it}, FC_{it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^B_{it}),\\
FC_{it} &= F_{FC}(IC^A_{it}, IC^B_{it}, Z_{1it}, Z_{2it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}).
\end{aligned}
$$

- $y^j_{it}(j = A, B)$: labor supply of agent j, $IC^j_{it}(j = A, B)$: informal care supply of agent j, FC_{it}: formal care utilization of household.

We derive the following functions based on this system of equations.

$$
\begin{aligned}
y^A_{it} &= f_{y^A}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}),\\
y^B_{it} &= f_{y^B}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}),\\
IC^A_{it} &= f_{IC^A}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}),\\
IC^B_{it} &= f_{IC^B}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}),\\
FC_{it} &= f_{FC}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}).
\end{aligned}
$$

When we estimate the effect of informal care on labor supply, we use the following functions.

$$
\begin{aligned}
y^j_{it} &= F_{y^j}(IC^j_{it}, \tilde{X}^j_{it})(j = A, B),\\
IC^j_{it} &= f_{IC^j}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it})(j = A, B).
\end{aligned}
$$

6.2 Estimation Method

In this section, we explain how to estimate the effect of informal care for the elderly on labor supply, and estimate the following equations. [18] As discussed in section 4, we utilize the variation of public nursing home capacity by government intervention on the supply side of the elderly care market when estimating the effect of informal care for the elderly on labor supply, where i is the individual number and $j = j(i)$ $(1 \leq j \leq N_R)$ is the region of residence number.

$$y_{it} = \beta_0 + \beta_1 IC_{it} + X'_{it}\delta_1 + \theta_i + \eta_{jt} + \epsilon_{1it} \tag{7}$$

$$IC_{it} = \alpha_0 + \alpha_1 1\{NursingCareLevel_{it} \geq n_1\} \cdot PA_{it} \tag{8}$$

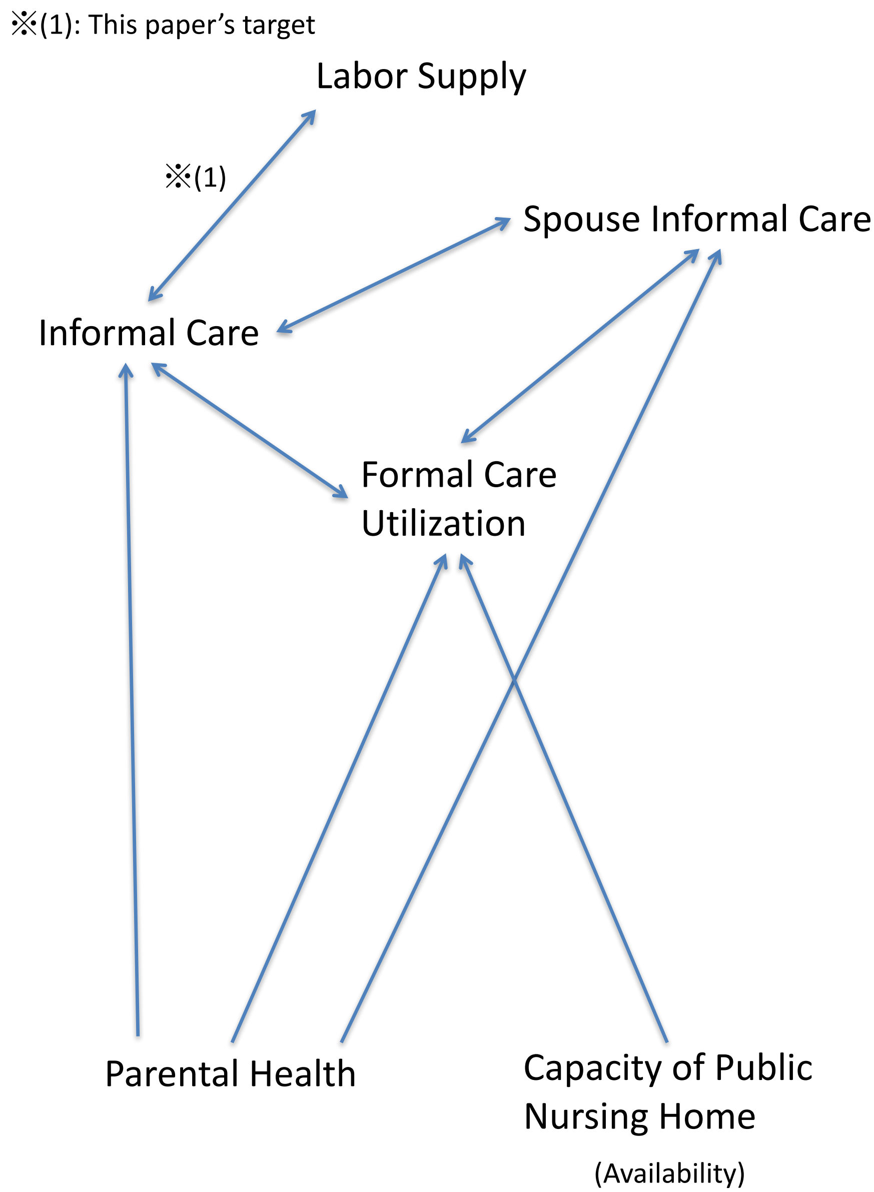
$$+\alpha_2 1\{NursingCareLevel_{it} \geq n_2\} \cdot Capacity_{it} + \tilde{Z}'_{3it}\alpha_3 + X'_{it}\delta_2 + \xi_i + p_{jt} + \epsilon_{2it}$$

We have discussed the causal relationship between informal care, spouse informal care, formal care utilization, and labor supply in section 6.1. We use $1\{NursingCareLevel_{it} \geq n_1\} \cdot PA_{it}$ as a proxy of $ParentalHealth$ and $1\{NursingCareLevel_{it} \geq n_2\} \cdot Capacity_{it}$ as a proxy of $Availability$. The followings are the definition of variables.

- y_{it}: Dummy variable indicating labor participation or working hours per week
- $Capacity_{it}$: $Capacity\ Index_{it} = 100 \times \frac{Capacity\ of\ Tokuyo_{it}}{\#\ of\ the\ people\ Aged\ over\ 65_{it}}$, where $Capacity\ of\ Tokuyo_{it}$: The Capacity of Tokuyo in the residence of respondent at period t, $\#\ of\ the\ people\ Aged\ over\ 65_{it}$: Population above 65 in the residence of respondent at period t. [19]
- $NursingCareLevel_{it}$: The maximum value of nursing care level of parents (only parents in contact with the respondent).

[18] All models are estimated using the STATA module xtivreg2. See Schaffer (2010) for further details.
[19] We only have 2005 and 2010 population information. We use the population nearest to period t.

Figure 9: The Relationship between Labor Supply and Informal Care



- PA_{it}: The age of parent who has maximum nursing care level (equal to zero if all parents are not certified as needing long-term care, only parents in contact with the respondent).

- IC_{it}: Dummy variable, which is equal to 1 if the respondent provides informal care.

- X_{it}: Other control variables, such as family characteristics, household assets and income.

- $\tilde{Z}_{3it}$: Other instruments such as the dummy variables indicating the number of parents certified as above care level one.

- θ_i, ξ_i: Fixed effects.

- η_{jt}, p_{jt}: Year-residence region effects.

- n_1, n_2: Natural numbers indicating an nursing care level.

We assume the following when estimating the effect of informal care for the elderly on labor supply. Let $Z_{1it} = 1\{NursingCareLevel_{it} \geq n_1\} \cdot PA_{it}$ and $Z_{2it} = 1\{NursingCareLevel_{it} \geq n_2\} \cdot Capacity_{it}$. We also define $Time_{it} = (1\{t = 1\}...1\{t = T\})'$ and $Region_{it} = (1\{j(i) = 1\}...1\{j(i) = N_R\})'$. Additionally, let $l_{it} = (Z_{1it}, Z_{2it}, \tilde{Z}'_{3it}, X'_{it}, Time'_{it}, Region'_{it})'$.

$$\textbf{Assumption A}: E[\epsilon_{1it}|L_i] = 0 \ (t = 1, 2, ..., T)$$

$$L'_i = (l_{i1}, l_{i2}, ..., l_{iT})$$

For example, ϵ_{1it} includes unexpected shocks to decrease the labor supply, such as a sudden injury to the respondent. When the assumption is valid, it is easy to show the identifiability of parameters by using the above assumption. T is the total number of periods. We define the following notations $\overline{A}_i \equiv \frac{1}{T}\sum_t A_{it}$ (A is a representative letter).

$$(y_{it} - \overline{y}_i) = \beta_1(IC_{it} - \overline{IC}_i) + (X'_{it} - \overline{X'}_i)\delta_1 + \eta_{jt} - \overline{\eta_j} + (\epsilon_{1it} - \overline{\epsilon}_{1i}) \tag{9}$$

Then, we rewrite equations (7) and (8) in the following way.

$$
\begin{aligned}
(y_{it} - \overline{y}_i) &= \beta_1(IC_{it} - \overline{IC}_i) + (X'_{it} - \overline{X'}_i)\delta_1 + \eta_{jt} - \overline{\eta_j} + (\epsilon_{1it} - \overline{\epsilon}_{1i}) \quad (10) \\
(IC_{it} - \overline{IC}_i) &= \alpha_1(Z_{1it} - \overline{Z}_{1i}) + \alpha_2(Z_{2it} - \overline{Z}_{2i}) + (\tilde{Z}'_{3it} - \overline{\tilde{Z}'}_{3i})\alpha_3 \\
&\quad +(X'_{it} - \overline{X'}_i)\delta_2 + p_{jt} - \overline{p_j} + (\epsilon_{2it} - \overline{\epsilon}_{2i}) \quad (11)
\end{aligned}
$$

Let $\tilde{L}_{it} = [(Z_{1it} - \overline{Z}_{1i}), (Z_{2it} - \overline{Z}_{2i}), (\tilde{Z}_{3it} - \overline{\tilde{Z}}_{3i})', (X_{it} - \overline{X}_i)', (Time_{it} \otimes Region_{it} - \overline{Time_i \otimes Region_i})']'$. Then, $\tilde{L}_{it}$ is a function of L_i, and we can write $\tilde{L}_{it} = A(L_i)$. As a result, $E[\tilde{L}_{it}(\epsilon_{1it} - \overline{\epsilon}_{1i})] = E[A(L_i)(\epsilon_{1it} - \overline{\epsilon}_{1i})] = 0$ by the Assumption A. We can identify the parameter $\eta_{jt} - \overline{\eta_j}$ in equation (10) by using the variables such as $Time_{it} \otimes Region_{it} - \overline{Time_i \otimes Region_i}$.

As explained in the previous section, in Japan, the nursing care level is determined by the local government based on the health condition of an applicant and the situation of household economic and family structure. Let $ParentalHealth_{it}$ be the health condition of an applicant. In other words, it is possible that $NursingCareLevel_{it}$ is a function of variables such as X_{it} and $ParentalHealth_{it}$ in the following way.

$$
NursingCareLevel_{it} = f(X_{it}, ParentalHealth_{it}). \quad (12)
$$

With respect to the unexpected shocks influencing the labor supply, the Assumption A seems to be valid. Here, the validity of Assumption A is checked by an over-identifying restriction test. The variable $1\{NursingCareLevel_{it} \geq n_1\} \cdot PA_{it}$ is a proxy variable of parental health. For example, $ParentalHealth_{it}$ is a function of PA_{it} and M_{it}, which are factors deciding the parental health.

$$ParentalHealth_{it} = g(PA_{it}, M_{it}). \tag{13}$$

On the other hand, $1\{NursingCareLevel_{it} \geq n_2\} \cdot Capacity_{it}$ controls the institutional factor to cause the respondent to provide informal care. When the respondent lives in an area where the capacity of Tokuyo is small, the probability to provide informal care becomes high because it is difficult to get admission to Tokuyo. We discuss this point from the analysis results in section 7.

Finally, we use models (10) and (11) to verify that there is no correlation between $(\epsilon_{1it} - \overline{\epsilon}_{1i})$ and $(IC_{it} - \overline{IC}_i)$. It is possible that $(IC_{it} - \overline{IC}_i)$ is exogenous. In fact, it is reported that providing informal care is exogenous in some studies. (e.g., Ishii (2015)) We check the endogeneity of $(IC_{it} - \overline{IC}_i)$ by using the Durbin-Wu-Hausman (DWH) test. [20] We analyze only samples having a parent who is alive and has a contact with the respondent. The household structure is different between couple and respondent without spouse. In this analysis, it is preferable that the respondent without spouse and couple are separately analyzed because the model differs. However, because the sample size of respondent without spouse is small, we only analyze couple's behavior.

7 Results

7.1 The Validity of Instruments

In this section, we check the validity of using the capacity of Tokuyo as instrument when we estimate the effect of informal care on labor supply. According to our discussion in section 6.1, the capacity of nursing homes (availability) indirectly influences informal care through the

[20] For a terse explanation of the DWH test, see Cameron and Trivedi (2010).

change in formal care utilization and directly influences formal care utilization, as equation (14) shows (informal care is influenced through the change in formal care utilization, which is influenced by the capacity of Tokuyo (Z_{1it})). Here, we estimate the equation (15).

$$IC^j_{it} = F_{IC^j}(y^j, IC^k_{it}, FC_{it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^j_{it}) \quad (14)$$

$$FC_{it} = F_{FC}(IC^A_{it}, IC^B_{it}, Z_{1it}, Z_{2it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}) = f_{FC}(Z_{1it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}^A_{it}, \tilde{X}^B_{it}) \quad (15)$$

Table 5: The Capacity of Tokuyo and Formal Care Utilization

	Age range			
Dependent variable:	**50-60**		**50-70**	
Formal care utilization(facility utilization only)	**Male**	**Female**	**Male**	**Female**
Capacity index				
$Capa \times 1\{NCL \geq C3\}$	0.188***	0.214***	0.149***	0.179***
	(0.046)	(0.047)	(0.032)	(0.034)
Other some controls				
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)	0.003***	0.002***	0.003***	0.003***
	(0.001)	(0.001)	(0.000)	(0.001)
N of certified ($Care$) =1	-0.087	0.043	-0.042	0.018
	(0.085)	(0.075)	(0.050)	(0.059)
N of certified ($Care$) ≥2	-0.240*	0.086	0.037	0.072
	(0.132)	(0.123)	(0.082)	(0.086)
Certified (≥ $C2$) female parent	0.243***	0.105	0.133**	0.104*
	(0.085)	(0.075)	(0.055)	(0.059)
Observations	957	911	2022	1602
Model	FE	FE	FE	FE

1 Standard errors in parentheses.
2 All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
3 * $p < .1$, ** $p < .05$, *** $p < .01$

As per Table 5, there is a positive significant effect of capacity of Tokuyo on formal care utilization. The magnitude is around 0.2 in all categories. With respect to the substitution effect of formal care utilization for the elderly on informal care, please see Nishimura and Oikawa (2017), who show the existence of the substitution effect of formal care utilization for the elderly on informal care, thus explaining why we can use the capacity as an instrumental variable in this study.

Importantly, as discussed in section 5, many people use public home care services when they do not use public nursing homes. As such, while the instrument Z_{1it} influences the allocation of formal care utilization, we do not utilize **the exogenous variation to stop formal care utilization completely** as later discussed in section 8.

7.2 Main Results

The main results are presented in Tables 6, 7, 8, and 9. $C1$ means "Care Level 1," $S1$ indicates "Support Level 1," "N of certified (Care) = 1 ($\geq$ 2)" is the dummy indicating that the number of parents with care level above 1 is 1 ($\geq$ 2), "Certified ($\geq$ C2) female parent" is the dummy variable indicating that the parent with care level above 1 is female. With respect to working hours per week, we use a dummy variable indicating whether working hours per week are more than 5, 10, or 20 hours ($\geq 5, \geq 10, \geq 20$). We test the endogeneity of informal care with the DWH test. When we do not reject the null hypothesis, we support the results of fixed effects (FE) model.

- According to Table 6, there is no effect of informal care on working for pay in male elderly. With respect to working hours per week in male elderly, there is no effect in all categories. On the other hand, the effect of informal care on working for pay is negative in female elderly (0.088). With respect to working hours per week in female elderly, there is no effect in all categories. Whether informal care is exogenous or not depends on gender. Male informal care is endogenous, while female informal care is exogenous. This point can be explained by who decides the allocation of informal care share ratio, α. We also discuss this point in the section 8.

- According to Table 7, the effect of informal care on working for pay is negative. We separate the female sample into two groups: females who are or are not working full time at the first interview or have reached age 54. According to Table 7, in the first group, the effect of informal care on working for pay is negative (0.082). In the second group, the effect of informal care on working for pay is also negative (0.069). The negative effect in females who are not full time workers is stronger than in female people who are full time workers. Informal care is exogenous in both groups.

- We expand the age range in Tables 8 and 9. We check the effect of including more retired elderly. As expected, the effect of informal care on working for pay becomes weaker. The effect is not so much different compared to the age group 50–60. There is no effect of informal care on working for pay in male elderly. The effect of informal care on working for pay is negative in female elderly (0.058). However, the effect is weaker than in the age group 50–60. Additionally, only in full-time working females or those aged 54, the effect of informal care on working for pay is negative (0.079). When we compare Tables 7 and 9, we can discuss the effect of including female retirees more on the "Provide care" coefficient. In the group "Female: Not full time worker at first interview or aged 54," the coefficient is not largely different between Table 7 and Table 9. However, in the group "Female: Full time worker at first interview or aged 54," there is no effect of "Provide care" on labor force participation in Table 9, although there is a negative effect in Table 7 (0.069).

- According to Table 10, we analyze the effect of spending informal care time on labor supply. As per Figure 10, males scarcely spend time on informal care. Thus, we omit the analysis of the effect of male elderlys informal care time spent on labor supply and only analyze female labor supply. As per Table 10, the effect of female elderly's informal care time spending on labor supply is not small. "LTC variables" indicate spending time on informal care more than 0, 5, 10, or 15 hours in each column. Only in the column "$\geq 15h$," informal care is endogenous. However, also in the columns "$\geq 5h$" and "$\geq 10h$," the p-values of informal care are small. According to these results, the effect of spending more than 15 hours per week on informal care on labor supply is not small (0.412).

Table 6: Labor Force Participation and Working Hour (Respondent Age:50-60, Only Couple)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Male								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.130**		-0.136***		-0.136***		-0.136***
		(0.051)		(0.048)		(0.048)		(0.048)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.003**		0.003**		0.003**		0.003**
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.180		0.227*		0.227*		0.227*
		(0.118)		(0.121)		(0.121)		(0.121)
N of certified ($Care$) $\geq$2		0.406**		0.451**		0.451**		0.451**
		(0.192)		(0.189)		(0.189)		(0.189)
Certified ($\geq C2$) female parent		-0.122		-0.190*		-0.190*		-0.190*
		(0.104)		(0.100)		(0.100)		(0.100)
2nd stage								
Provide care	-0.004	-0.090	0.001	0.084	-0.006	0.037	-0.021	0.075
	(0.018)	(0.060)	(0.019)	(0.063)	(0.022)	(0.031)	(0.013)	(0.049)
Observations	983	983	883	883	883	883	883	883
OverID p-value		0.754		0.776		0.573		0.311
DWH p-value		0.066		0.085		0.220		0.042
Female								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.213***		-0.225***		-0.225***		-0.225***
		(0.045)		(0.046)		(0.046)		(0.046)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.006***		0.006***		0.006***		0.006***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.057		0.112		0.112		0.112
		(0.099)		(0.100)		(0.100)		(0.100)
N of certified ($Care$) $\geq$2		0.155		0.165		0.165		0.165
		(0.140)		(0.147)		(0.147)		(0.147)
Certified ($\geq C2$) female parent		0.048		0.009		0.009		0.009
		(0.078)		(0.079)		(0.079)		(0.079)
2nd stage								
Provide care	0.088**	0.150**	-0.046	-0.128*	-0.037	-0.104	-0.036	-0.059
	(0.036)	(0.068)	(0.037)	(0.068)	(0.038)	(0.074)	(0.042)	(0.084)
Observations	921	921	839	839	839	839	839	839
OverID p-value		0.983		0.991		0.936		0.637
DWH p-value		0.220		0.129		0.264		0.755

[1] Standard errors in parentheses.
[2] All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
[3] * $p < .1$, ** $p < .05$, *** $p < .01$

Table 7: Labor Force Participation and Working Hour (Respondent Age:50-60 (Only Female), Only Couple)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Female:Not full time worker at 1st interview or aged 54								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.221***		-0.254***		-0.254***		-0.254***
		(0.058)		(0.057)		(0.057)		(0.057)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.005***		0.004***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.068		0.144		0.144		0.144
		(0.125)		(0.122)		(0.122)		(0.122)
N of certified ($Care$) ≥2		0.271		0.342**		0.342**		0.342**
		(0.167)		(0.164)		(0.164)		(0.164)
Certified ($\geq C2$) female parent		0.054		0.032		0.032		0.032
		(0.089)		(0.089)		(0.089)		(0.089)
2nd stage								
Provide care	0.082*	0.169*	-0.021	-0.131	-0.006	-0.105	-0.028	-0.031
	(0.047)	(0.096)	(0.046)	(0.094)	(0.049)	(0.104)	(0.050)	(0.117)
Observations	680	680	632	632	632	632	632	632
OverID p-value		0.505		0.821		0.541		0.410
DWH p-value		0.238		0.148		0.258		0.977
Female:Full time worker at 1st interview or aged 54								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.225***		-0.203***		-0.203***		-0.203***
		(0.062)		(0.064)		(0.064)		(0.064)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.007***		0.007***		0.007***		0.007***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.106		0.196		0.196		0.196
		(0.168)		(0.171)		(0.171)		(0.171)
N of certified ($Care$) ≥2		-0.000		-0.086		-0.086		-0.086
		(0.259)		(0.325)		(0.325)		(0.325)
Certified ($\geq C2$) female parent		0.068		-0.038		-0.038		-0.038
		(0.151)		(0.145)		(0.145)		(0.145)
2nd stage								
Provide care	0.069*	0.070	-0.084	-0.064	-0.094*	-0.062	-0.023	-0.017
	(0.041)	(0.066)	(0.052)	(0.074)	(0.057)	(0.076)	(0.072)	(0.090)
Observations	233	233	203	203	203	203	203	203
OverID p-value		0.501		0.302		0.315		0.154
DWH p-value		0.973		0.648		0.477		0.908

1 Standard errors in parentheses.

2 All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.

3 * $p < .1$, ** $p < .05$, *** $p < .01$

4 In the estimation for the female full-timer, we replace the year-municipality dummies with the year dummies and year-municipality dummies that have enough non-zero values because we cannot compute over-identifying test statistics due to the dummies without enough non-zero values.

Table 8: Labor Force Participation and Working Hour (Respondent Age:50-70, Only Couple)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Male								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.088***		-0.076**		-0.076**		-0.076**
		(0.031)		(0.031)		(0.031)		(0.031)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.002***		0.002***		0.002***		0.002***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.202***		0.204***		0.204***		0.204***
		(0.063)		(0.064)		(0.064)		(0.064)
N of certified ($Care$) ≥2		0.334***		0.283***		0.283***		0.283***
		(0.099)		(0.102)		(0.102)		(0.102)
Certified ($\geq C2$) female parent		-0.063		-0.088		-0.088		-0.088
		(0.063)		(0.064)		(0.064)		(0.064)
2nd stage								
Provide care	-0.005	-0.073	-0.012	0.089	-0.021	0.079	-0.041	0.144
	(0.025)	(0.072)	(0.026)	(0.083)	(0.028)	(0.088)	(0.028)	(0.101)
Observations	2082	2082	1883	1883	1883	1883	1883	1883
OverID p-value		0.323		0.545		0.289		0.405
DWH p-value		0.318		0.191		0.227		0.058
Female								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.187***		-0.186***		-0.186***		-0.186***
		(0.032)		(0.032)		(0.032)		(0.032)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.005***		0.005***		0.005***		0.005***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.077		0.106		0.106		0.106
		(0.071)		(0.072)		(0.072)		(0.072)
N of certified ($Care$) ≥2		0.183*		0.181*		0.181*		0.181*
		(0.096)		(0.097)		(0.097)		(0.097)
Certified ($\geq C2$) female parent		-0.010		-0.040		-0.040		-0.040
		(0.062)		(0.062)		(0.062)		(0.062)
2nd stage								
Provide care	0.058**	0.036	-0.035	-0.014	-0.024	-0.003	-0.026	-0.039
	(0.027)	(0.059)	(0.027)	(0.061)	(0.029)	(0.068)	(0.030)	(0.071)
Observations	1639	1639	1498	1498	1498	1498	1498	1498
OverID p-value		0.509		0.426		0.554		0.976
DWH p-value		0.664		0.689		0.728		0.837

1 Standard errors in parentheses.
2 All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
3 * $p < .1$, ** $p < .05$, *** $p < .01$

Table 9: Labor Force Participation and Working Hour (Respondent Age:50-70 (Only Female), Only Couple)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Female:Not full time worker at 1st interview or aged 54								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.167***		-0.181***		-0.181***		-0.181***
		(0.038)		(0.038)		(0.038)		(0.038)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.004***		0.004***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.108		0.149*		0.149*		0.149*
		(0.087)		(0.086)		(0.086)		(0.086)
N of certified ($Care$) ≥2		0.289***		0.302***		0.302***		0.302***
		(0.112)		(0.109)		(0.109)		(0.109)
Certified ($\geq C2$) female parent		-0.054		-0.070		-0.070		-0.070
		(0.070)		(0.070)		(0.070)		(0.070)
2nd stage								
Provide care	0.079**	0.096	-0.048	-0.040	-0.023	-0.020	-0.034	-0.069
	(0.032)	(0.081)	(0.030)	(0.085)	(0.034)	(0.094)	(0.034)	(0.096)
Observations	1174	1174	1088	1088	1088	1088	1088	1088
OverID p-value		0.476		0.400		0.286		0.613
DWH p-value		0.813		0.920		0.973		0.704
Female:Full time worker at 1st interview or aged 54								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.217***		-0.179***		-0.179***		-0.179***
		(0.060)		(0.061)		(0.061)		(0.061)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.006***		0.006***		0.006***		0.006***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.064		0.059		0.059		0.059
		(0.123)		(0.130)		(0.130)		(0.130)
N of certified ($Care$) ≥2		-0.005		-0.029		-0.029		-0.029
		(0.171)		(0.185)		(0.185)		(0.185)
Certified ($\geq C2$) female parent		0.060		-0.022		-0.022		-0.022
		(0.138)		(0.130)		(0.130)		(0.130)
2nd stage								
Provide care	0.022	-0.075	-0.008	0.065	-0.027	0.041	-0.002	0.087
	(0.048)	(0.088)	(0.052)	(0.089)	(0.054)	(0.097)	(0.064)	(0.104)
Observations	442	442	401	401	401	401	401	401
OverID p-value		0.306		0.331		0.576		0.081
DWH p-value		0.124		0.209		0.325		0.229

[1] Standard errors in parentheses.
[2] All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
[3] * $p < .1$, ** $p < .05$, *** $p < .01$

Figure 10: The Distribution of Informal Care Time Spending

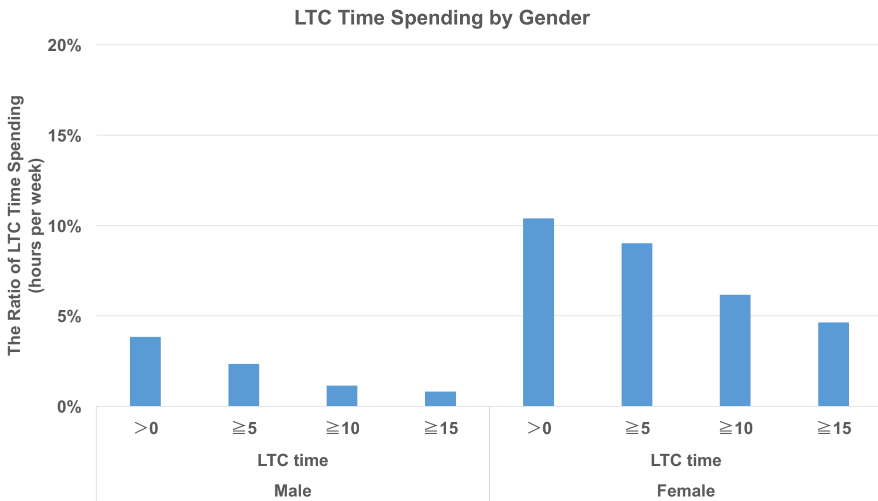


Table 10: Labor Force Participation (Respondent Age:50-60, Only Couple)

Dependent variable: Not working	(1)		(2)		(3)		(4)	
	LTC time (hours per week)							
LTC variables	> 0h		≥ 5h		≥ 10h		≥ 15h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Female								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.213***		-0.150***		-0.096***		-0.064**
		(0.045)		(0.043)		(0.035)		(0.031)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.006***		0.004***		0.003***		0.002***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		0.057		0.039		-0.084		-0.085
		(0.099)		(0.088)		(0.073)		(0.062)
N of certified ($Care$) $\geq$2		0.155		0.108		-0.087		-0.015
		(0.140)		(0.139)		(0.110)		(0.097)
Certified ($\geq C2$) female parent		0.048		0.065		0.119*		0.092
		(0.078)		(0.069)		(0.071)		(0.057)
2nd stage								
LTC variables	0.088**	0.150**	0.080*	0.216**	0.087	0.333**	0.100	0.412**
	(0.036)	(0.068)	(0.047)	(0.101)	(0.058)	(0.168)	(0.071)	(0.201)
Observations	921	921	871	871	871	871	871	871
OverID p-value		0.983		0.979		0.976		0.995
DWH p-value		0.220		0.151		0.122		0.093

[1] Standard errors in parentheses.
[2] All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
[3] * $p < .1$, ** $p < .05$, *** $p < .01$

7.3 The Difference in the Role of Instrumental Variables between Male Elderly and Female Elderly

Next, we discuss the structural difference in the estimated equations between male elderly and female elderly. Table 11 shows the estimated results, adding a spousal informal care dummy, which indicates whether the spouse helps provide informal care in the first stage. According to Table 11, in the first stage of male "Not working," we can find that there is only a significant effect in the coefficient of "Provide care (SP)." On the other hand, in the first stage of female "Not working," we can find that there are also significant effects in the coefficients of "$Capa \times 1\{NCL \geq C3\}$" and "$PA(Parent'sage \times 1\{NCL \geq S1\})$." According to this result, the instruments "$Capa \times 1\{NCL \geq C3\}$" and "$PA(Parent'sage \times 1\{NCL \geq S1\})$" directly influences female informal care. However, these instruments do not directly influence male informal care, but do so indirectly through the influence of female informal care.

According to this discussion, we suggest the following relationship in the male and female informal care functions. We note $A =$ husband and $B =$ wife. In the informal care function of male household members ($A =$ husband), it is possible that formal care is not included.

$$IC_{it}^{A} = F_{IC^{A}}(y_{it}^{A}, IC_{it}^{B}, \tilde{X}_{it}^{A})$$

$$IC_{it}^{B} = F_{IC^{B}}(y_{it}^{B}, IC_{it}^{A}, FC_{it}, Z_{2it}, \tilde{Z}_{3it}, \tilde{X}_{it}^{B})$$

Table 11: Labor Force Participation and Working Hour (Respondent Age:50-60, Only Couple)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Male								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.053		-0.069*		-0.069*		-0.069*
		(0.042)		(0.041)		(0.041)		(0.041)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.000		0.000		0.000		0.000
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		-0.026		0.036		0.036		0.036
		(0.091)		(0.091)		(0.091)		(0.091)
N of certified ($Care$) $\geq$2		-0.083		-0.026		-0.026		-0.026
		(0.145)		(0.146)		(0.146)		(0.146)
Certified ($\geq C2$) female parent		0.057		0.023		0.023		0.023
		(0.077)		(0.073)		(0.073)		(0.073)
Provide care (SP)		0.673***		0.677***		0.677***		0.677***
		(0.056)		(0.059)		(0.059)		(0.059)
2nd stage								
Provide care	-0.005	0.011	0.003	-0.012	-0.004	-0.004	-0.020	-0.028
	(0.018)	(0.030)	(0.019)	(0.033)	(0.022)	(0.035)	(0.013)	(0.029)
Observations	980	980	879	879	879	879	879	879
OverID p-value		0.357		0.369		0.621		0.269
DWH p-value		0.427		0.514		0.990		0.755
Female								
1st stage								
$Capa \times 1\{NCL \geq C3\}$		-0.132***		-0.138***		-0.138***		-0.138***
		(0.037)		(0.038)		(0.038)		(0.038)
PA ($Parent's\ age \times 1\{NCL \geq S1\}$)		0.004***		0.004***		0.004***		0.004***
		(0.001)		(0.001)		(0.001)		(0.001)
N of certified ($Care$) =1		-0.028		-0.012		-0.012		-0.012
		(0.076)		(0.082)		(0.082)		(0.082)
N of certified ($Care$) $\geq$2		-0.027		-0.074		-0.074		-0.074
		(0.106)		(0.108)		(0.108)		(0.108)
Certified ($\geq C2$) female parent		0.099		0.087		0.087		0.087
		(0.061)		(0.063)		(0.063)		(0.063)
Provide care (SP)		0.603***		0.604***		0.604***		0.604***
		(0.059)		(0.063)		(0.063)		(0.063)
2nd stage								
Provide care	0.088**	0.081	-0.046	-0.043	-0.037	-0.026	-0.036	0.006
	(0.036)	(0.057)	(0.037)	(0.054)	(0.038)	(0.056)	(0.042)	(0.061)
Observations	921	921	839	839	839	839	839	839
OverID p-value		0.541		0.371		0.441		0.604
DWH p-value		0.853		0.941		0.807		0.399

[1] Standard errors in parentheses.
[2] All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
[3] * $p < .1$, ** $p < .05$, *** $p < .01$

7.4 Robustness Check: The Instrumental Variables in the Related Literature

Figure 12 shows the estimated results, including in the first stage, when the variables indicate whether both parents and parents-in-law are alive or not. For example, Bolin et al. (2008) and Van Houtven et al. (2013) use whether parents are alive or not as instrumental variables. We also include these variables in the first stage. Figure 12 shows these results. In all age ranges (50–60, 50–64, 50–70), the estimated results are not significantly different compared to the estimated results without the variables of whether both parents and parents-in-law are alive or not. However, the important point is that the p-value of the over-identification test is low compared to the results without these variables. In the age range 50–64, the analysis of "Not working" indicates the rejection of the null hypothesis in the over-identification test.

Table 12: Labor Force Participation and Working Hour (Additional Instruments: Both parents are alive, Both parents-in-law are alive)

	(1)		(2)		(3)		(4)	
			Working hours per week					
Dep.	Not working		$\geq$ 5h		$\geq$ 10h		$\geq$ 20h	
	FE	FE-IV	FE	FE-IV	FE	FE-IV	FE	FE-IV
Age: 50 to 60								
Male								
Provide care	-0.005	-0.080	0.002	0.084	-0.004	0.040	-0.020	0.050
	(0.018)	(0.052)	(0.019)	(0.058)	(0.022)	(0.034)	(0.013)	(0.048)
Observations	976	976	875	875	875	875	875	875
OverID p-value		0.771		0.784		0.650		0.416
DWH p-value		0.053		0.063		0.233		0.137
Female								
Provide care	0.090**	0.123*	-0.049	-0.097	-0.039	-0.073	-0.038	-0.017
	(0.037)	(0.067)	(0.037)	(0.069)	(0.038)	(0.074)	(0.042)	(0.080)
Observations	918	918	836	836	836	836	836	836
OverID p-value		0.293		0.495		0.280		0.182
DWH p-value		0.493		0.389		0.578		0.764
Age: 50 to 64								
Male								
Provide care	-0.014	-0.035	0.019	0.066	0.013	0.052	0.011	0.094
	(0.021)	(0.047)	(0.022)	(0.050)	(0.022)	(0.067)	(0.027)	(0.084)
Observations	1559	1559	1400	1400	1400	1400	1400	1400
OverID p-value		0.380		0.740		0.736		0.289
DWH p-value		0.624		0.312		0.549		0.296
Female								
Provide care	0.095***	0.032	-0.062**	-0.010	-0.057*	0.003	-0.058*	-0.015
	(0.030)	(0.064)	(0.031)	(0.065)	(0.033)	(0.070)	(0.034)	(0.072)
Observations	1334	1334	1219	1219	1219	1219	1219	1219
OverID p-value		0.098		0.124		0.399		0.660
DWH p-value		0.225		0.338		0.312		0.490
Age: 50 to 70								
Male								
Provide care	-0.005	-0.047	-0.012	0.072	-0.020	0.061	-0.041	0.112
	(0.026)	(0.066)	(0.026)	(0.077)	(0.028)	(0.081)	(0.029)	(0.095)
Observations	2072	2072	1873	1873	1873	1873	1873	1873
OverID p-value		0.445		0.700		0.432		0.289
DWH p-value		0.505		0.241		0.289		0.093
Female								
Provide care	0.067**	0.003	-0.039	0.022	-0.028	0.020	-0.030	-0.007
	(0.027)	(0.060)	(0.028)	(0.061)	(0.030)	(0.067)	(0.031)	(0.068)
Observations	1630	1630	1490	1490	1490	1490	1490	1490
OverID p-value		0.176		0.212		0.383		0.814
DWH p-value		0.211		0.239		0.414		0.700

[1] Standard errors in parentheses.
[2] All specification include age, age squared, $Age \geq PEA$(PEA:pension eligibility age), N of children, HH income, house ownership, HH saving(imputed), and year-municipality dummies.
[3] * $p < .1$, ** $p < .05$, *** $p < .01$

8 Discussion

We will shortly discuss our main results as follows.

- **Why is female informal care exogenous? and why is male informal care endogenous?**

  We interpret the results based on the model in section 6.1. Whether male and female informal care is endogenous or not is influenced by who decides the informal care sharing rate, α. According to the discussion of 6.1, if α is decided by a male household member, female informal care is exogenous for female household member.

- **Why is the effect of informal care on labor supply small?**

  As discussed in section 5, in Japan, the public (home) care service is available when a person who requires nursing care lives in a household. This is most important in explaining our results, as we do not separate the samples into a group utilizing home care and a group not utilizing home care. As we discuss in section 7.2, the effect of spending more than 15 hours per week on informal care on labor supply is not small. Overall, spending time on informal care is small both in male and female elderly. This is because home care services are easily available in Japan.

- **The effect of the government intervention on the supply side of elderly the care market**

  For analyzing the effect of government intervention on the supply side of the elderly care market on informal care in Japan, we check the coefficient of "$Capa \times 1\{NCL \geq C3\}$." This coefficient suggests the effect of increasing the capacity of Tokuyo per capita on providing informal care. As per Table 6, the absolute value of the female coefficient is larger than that of the male coefficient. Additionally, as Table 11 shows, male informal care is indirectly influenced by the capacity of Tokuyo per capita through the female informal care. The effect of the capacity of Tokuyo per capita on providing informal care is strong in female elderly. Overall, the effect of informal care on labor supply is small in Japan. With public home care services also available, the government intervention on the supply side of the elderly care market is effective for labor supply in Japan.

9 Conclusion

This study analyzes the effect of informal care for elderly on labor supply, utilizing the exogenous variation of government intervention on the supply side of the elderly care market in Japan to estimate this effect. As a result, the supply of public nursing care is controlled by the government. We utilize this exogenous variation for estimating the effect of informal care for elderly on labor supply. According to our results, the following points are clarified.

- The effect of informal care for elderly on labor supply in both males and females is small. Especially, when compared with literature, the effect is smaller than in extant studies.

- The time spent on informal care in households is the focus on female household members. The government intervention is effective for increasing female labor supply.

In future work, the heterogeneity of utilizing home care services should be considered. Our analysis does not consider separating the group utilizing home care from the group not utilizing home care. As a result, the effect of informal care on labor supply is small. In fact, in the group not utilizing home care service, it is possible that the effect is very strong.

A Appendix

A.1 Comparison with the Japanese Literature

We summarize the results of Japanese studies in Table 13, [21] comparing the results of this study with the results in the listed studies. In Japan, the studies directly analyzing the effect of informal care on labor supply are Yamada and Shimizutani (2015), Ishii (2015), and Moriwaki (2016). Other studies estimate the effect of LTCI or nursing home capacity on labor supply, but do not directly estimate the effect of informal care on labor supply. According to Yamada and Shimizutani (2015) and Moriwaki (2016), there is a negative effect of informal care on male labor supply. Conversely, we find no effect of informal care on male labor supply by using the exogenous variation of the supply side of the elderly care market. Additionally, our estimates with respect to the effect of informal care on female labor supply are small compared to Yamada and Shimizutani (2015) and Ishii (2015). This is because we use the different instruments compared to theses studies.

[21] We omit Wakabayashi and Donato (2005) because it does not consider the endogeneity of informal care.

Table 13: Japanese Literature

	Analysis Method	Instruments	Results	Data
Shimizutani et al (2008)	DID	-	Not Direct Effect	•Survey on Long-term Care Users •Survey on Elderly Medical Care Insurance
Sugawara and Nakamura (2014)	Two part model	-	Not Direct Effect	Comprehensive Survey of Living Conditions
Yamada and Shimizutani (2015)	IV method	Age, Health Status, and Gender of a Parent.	•Negative Effect on Labor Force Participation (-0.202, Male)(-0.581, Female) (The effect of being a maincare-giver on labor supply)	Comprehensive Survey of Living Conditions
Fukahori et al (2015)	DID	-	Not Direct Effect	Comprehensive Survey of Living Conditions
Ishii (2015)	IV method	Age of Eldest Parent	•Exogeneity of Daily Care •Negative Effect on Female Labor Participation (-0.134 (Coresident, OLS))	JSTAR (only female)
Moriwaki (2016)	FEIV method	Care Level	•Negative Effect on Male Labor Force Participation (-0.545)	JSTAR
Kondo (2016)	Probit model	-	Not Direct Effect	•Labour Force Survey •Employment Status Survey

A.2 Asset Level Imputation

Here, we explain saving variable imputation procedures. First, we show the structure of the JSTAR questionnaire with respect to the saving variable and explain reasons why some saving values are missing. Then, we explain the imputation procedures, which are the simplified version of the HRS method. [22] Finally, we compare the imputed saving values with the original saving values and the harmonized JSTAR imputation values.

A.2.1 Questionnaire structure of saving variable

The JSTAR has two types of interviews. One is the leave-behind (LB) questionnaire interview and the other is a computer-assisted personal interviewing (CAPI). Basically, respondents are required to answer the LB questionnaire first and the CAPI afterwards. The questions about saving are asked in both questionnaires. Figure 11 shows the structure of questions with respect to saving values. [23]

First, in the LB questionnaire, respondents are asked to answer questions on the ownership and saving value for a respondent and his/her spouse. The procedure is as follows:

1. A respondent indicates the ownership of their saving. (Q32)

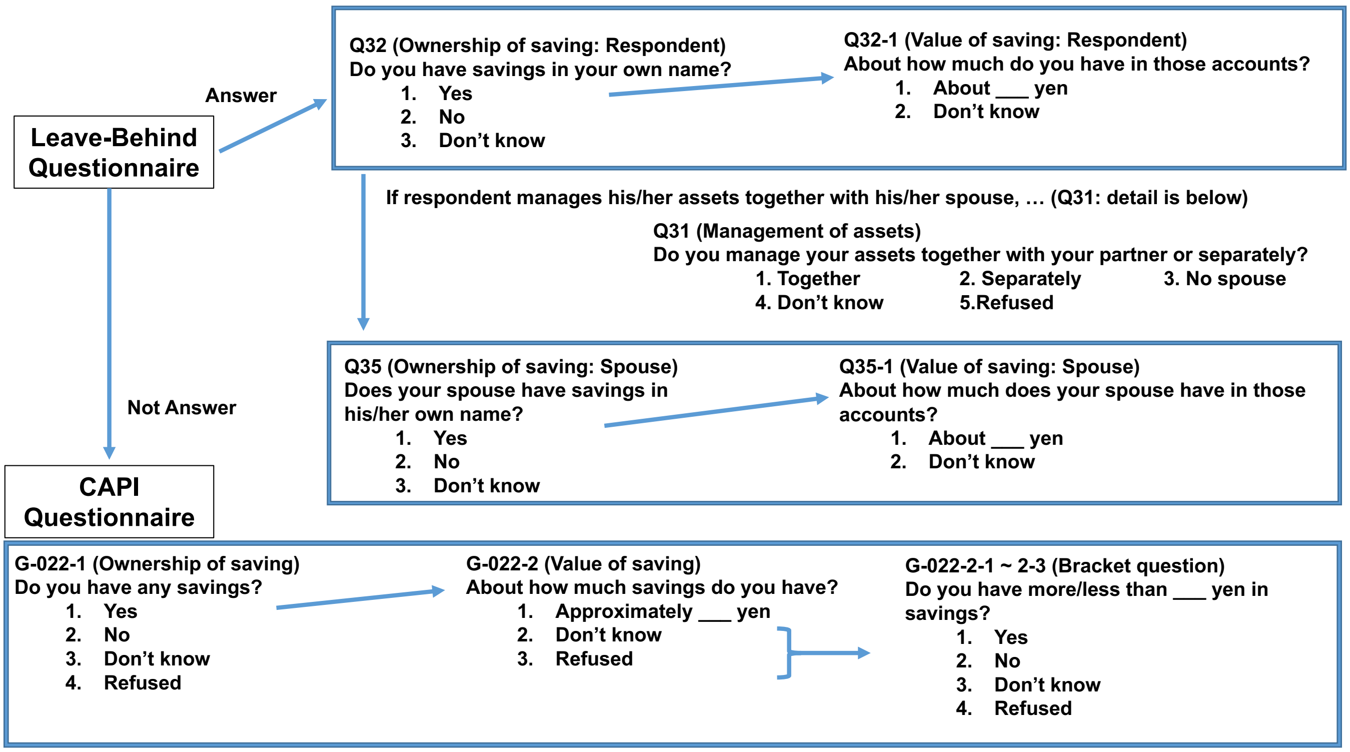
2. If answering "yes" in Q32, respondents indicate the value of their own saving. (Q32-1)

3. If a respondent manages his/her assets together with their spouse (Q31) [24], they move to questions about his/her spouse's saving information.

4. A respondent identifies the ownership of his/her spouse's saving. (Q35)

[22] See Hurd et al. (2016) for details of HRS method.
[23] Figure 11 shows the structure of 2007 JSTAR.
[24] Question Q31 states "Do you manage your assets together with your spouse (or common-law spouse) or separately?" and the answer choices are "1. together"; "2. separately"; "3. no spouse"; "4. don't know"; and "5. refused."

Figure 11: JSTAR's questionnaire structure of saving variable



5. If answering "yes" in Q35, a respondent indicates the value of his/her spouse's saving. (Q35-1)

If not answering the saving information in the LB questionnaire, a respondent is asked to indicate household level saving in the CAPI. The procedure of the CAPI questions is as follows:

1. A respondent indicates the ownership of saving. (G-022-1)

2. If answering "yes" in G-022-1, a respondent identifies the value of saving. (G-022-2)

3. If the saving value is not answered in G-022-2, a respondent is asked to answer the saving value as the brackets three times. (G-022-2-1 ∼ G-022-2-3)

As a result, we can obtain either the individual level (respondent and/or spouse) saving variables (ownership and value) or the household level saving variables (ownership and value (or brackets)). Finally, using this information, we can construct the household-level saving values as follows:

Case 1: continuous values;

Case 2: bracket values;

Case 3: only ownership;

Case 4: no information about ownership.

In Cases 2, 3, and 4, saving values are missing and cannot be used for analysis. We impute the saving values in all these cases, although the Harmonized JSTAR provides the imputed saving values only in Cases 2 and 3. [25]

A.2.2 Imputation Procedures

We use the simplified version of the HRS method for the saving values imputation using cross-sectional variations.[22] The outline of imputation procedure is as follows:

Step 0: Constructing the HH level variables.

- Step 0-1: Construct the HH level variables using LB questionnaire information.
- Step 0-2: If there are missing values in variables constructed above, merge those with the variables surveyed in CAPI.

Step 1: Ownership imputation.

- Step 1-1: Estimate the ownership imputation model using a binary logit model.

[25] See the codebook of the Harmonized JSTAR at `https://g2aging.org/startfile.php?f=codebooks/Harmonized%20JSTAR%20B.pdf` for more details.

Step 1-2: Calculate the predicted probabilities of ownership.

Step 1-3: Take a draw random variables from the uniform distribution.

Step 1-4: Assign ownership using the predicted probabilities and random variables.

Step 2: Bracket imputation.

Step 2-1: Estimate the bracket imputation model using an ordered logit model.

Step 2-2: Calculate the predicted probabilities in the j-th bracket.

Step 2-2: Take a draw random variables from the uniform distribution.

Step 2-2: Assign bracket j using the predicted probabilities and random variables.

Step 3: Value imputation

Step 3-a: Nearest neighbor method for closed brackets

Step 3-a-1: Estimate the linear value imputation model.

Step 3-a-2: Calculate the predicted saving values.

Step 3-a-3: Define donor groups

Step 3-a-4: Assign the imputed values from the donor group.

Step 3-b: Tobit 25 method for upper open brackets

Step 3-b-1: Estimate the tobit value imputation model.

Step 3-b-2: Assign the imputed values from the estimated distribution.

In Step 0, we construct the household level variables such as the ownership, values, and bracket values of savings using both LB questionnaire and CAPI information. First, we construct the household level ownership and values of saving using individual level variables surveyed in LB questionnaire. If there are missing values in the variables, we merge those with the household level variables surveyed in CAPI section. Then, we generate the household level bracket values using CAPI variables.[26] Finally, we obtain three household level variables, the ownership, values, and bracket values of saving and call these as original household level variables.

In Step 1, we impute the ownership of savings using the logit model. First, we regress the original ownership on covariates using logit and obtain the predicted probabilities of saving ownership, p_{it}.[27] Second, we draw a random variable, u_{it}, from the uniform distribution, $U(0,1]$, and assign ownership ($=1$) if $u_{it} < p_{it}$ and non-ownership ($=0$) otherwise.

In Step 2, we impute the bracket value of saving using an ordered logit model. We regress the bracket categories on the covariates using an ordered logit model and obatain the predicted probabilities being in the j-th bracket, p_{ijt}. Then, we calculate the cumulative probabilities for each bracket, $P_{ijt} = \sum_{k=1}^{j} p_{ikt}$. Finally, we draw a random variable, v_{it}, from the uniform distribution, $U(0,1]$, and if $P_{i,j-1,t} < v_{it} \leq P_{ijt}$, we assign bracket j.

In Step 3, we impute the saving values using two imputation methods, depending on the bracket values. There are two types of brackets: closed brackets, which have a closed interval, and upper open brackets, which have an open upper interval. [28] In the case of closed brackets, we use the nearest neighbor (NN) method. First, we regress the saving values which are applied the inverse hyperbolic sine transformation on the covariates using linear regression model for all households and obtain the predicted values of saving. Second, for each bracket, we define a donor group from the households who report a value within the bracket of interest. Finally, from the donor group, the reported value that is closest to the predicted value is assigned to the each household who has missing continuous values and original or imputed bracket.

On the other hand, in the case of upper open brackets, we use the tobit 25 method. First, we regress the logged saving values on covariates using the tobit model with a threshold that is

[26] Here, for simplicity, we reconstruct the brackets as [0,500), [500,1500), [1500,∞). (unit: JPY 10k)

[27] We use female dummy, age, age squared, education dummies, marital status dummies, number of children, and municipality dummies as covariates.

[28] Here, [0,500) and [500,1500) are the closed brackets and [1500,∞) is the upper open bracket.

the 25th percentile of the saving value distribution. Second, from the estimated distribution, we assign the imputed values for households with upper open brackets conditional on the given bracket.

A.2.3 Imputation Results

Table 14 shows the summary statistics of original and imputed saving values for each wave. The column “original” shows the summary of original saving values, column “imputed values: ours” shows the values imputed by our method, column shows “imputed values: H JSTAR,” which is the values imputed by the harmonized JSTAR. The unit of saving values is JPY ten thousand. In all waves, we recover the 1.5 times observations as original values. Figure 12 illustrates the distributions of the values. The blue solid lines indicate the distribution of original values, the red dashed lines that of our imputation values, and the green dashed lines that of the harmonized JSTAR imputation values. The distributions of our imputation variables have roughly similar forms to the distributions of the original values.

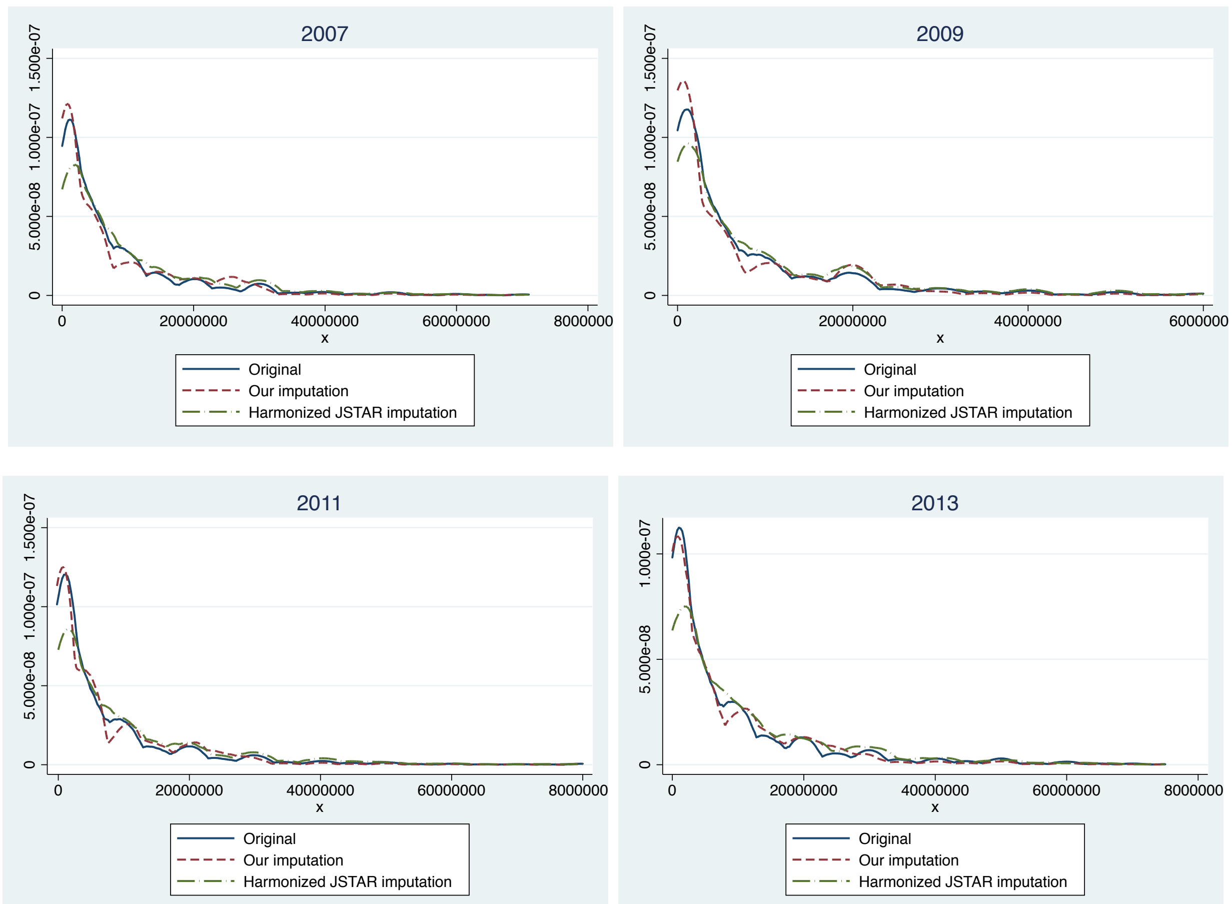
Table 14: Summary statistics of original and imputed saving values

Statistics	Original	Imputed values: Ours	Imputed values: H JSTAR
2007			
Observations	2479	4198	3170
mean	850	783	1060
sd	1460	1260	1550
min	0	0	0
p10	0	0	0
p25	100	40	150
p50	400	303	500
p75	1000	1040	1400
p90	2100	2300	3000
p95	3000	2800	3700
p99	7000	5390	7500
max	30000	30000	30000
2009			
Observations	2574	4555	3369
mean	817	700	994
sd	1580	1300	1670
min	0	0	0
p10	0	0	0
p25	44	10	100
p50	300	200	500
p75	1000	1000	1300
p90	2000	2000	2500
p95	3150	2500	4000
p99	6000	5000	6200
max	40000	40000	40000

Statistics	Original	Imputed values: Ours	Imputed values: H JSTAR
2011			
Observations	2861	5330	4234
mean	1200	915	1420
sd	19100	13700	16000
min	0	0	0
p10	0	0	0
p25	50	30	100
p50	300	350	500
p75	1000	1020	1400
p90	2000	2030	3000
p95	3000	2600	4000
p99	6500	5000	9000
max	1000000	1000000	1000000
2013			
Observations	2495	4370	3143
mean	994	849	1170
sd	2230	1760	1790
min	0	0	0
p10	0	0	0
p25	100	23	100
p50	400	400	600
p75	1010	1100	1500
p90	2500	2200	3000
p95	4000	2840	4000
p99	7500	6000	8000
max	50000	50000	27000

[1] Unit: 10k yen

Figure 12: Distributions of original and imputed saving values



References

Bauer, J. M. and Sousa-Poza, A. (2015). Impacts of Informal Caregiving on Caregiver Employment, Health, and Family. *Journal of Population Ageing*, 8(3):113–145.

Berecki-gisolf, J., Lucke, J., Hockey, R., and Dobson, A. (2008). Transitions into informal caregiving and out of paid employment of women in their 50s. *Social Science & Medicine*, 67:122–127.

Bolin, K., Lindgren, B., and Lundborg, P. (2008). Your next of kin or your own career ? Caring and working among the 50 + of Europe. *Journal of Health Economics*, 27:718–738.

Bonsang, E. (2009). Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, 28(1):143–154.

Carmichael, F. and Charles, S. (1998). The labour market costs of community care. *Journal of Health Economics*, 17(6):747–765.

Carmichael, F. and Charles, S. (2003). The opportunity costs of informal care: Does gender matter? *Journal of Health Economics*, 22(5):781–803.

Carmichael, F., Charles, S., and Hulme, C. (2010). Who will care? Employment participation and willingness to supply informal care. *Journal of Health Economics*, 29(1):182–190.

Ching, A. T., Hayashi, F., and Wang, H. (2015). Quantifying the impacts of limited supply: the case of nursing homes. *International Economic Review*, 56(4):1291–1322.

Chul, J., Done, N., and Anderson, G. F. (2015). Considering long-term care insurance for middle-income countries : comparing South Korea with Japan and Germany. *Health policy*, 119(10):1319–1329.

Ciani, E. (2012). Informal adult care and caregivers ' employment in Europe. *Labour Economics*, 19(2):155–164.

Cohen, J. W. and Spector, W. D. (1996). The effect of Medicaid reimbursement on quality of care in nursing homes. *Journal of Health Economics*, 15(1):23–48.

Connelly, R. (1992). THE EFFECT OF CHILD CARE COSTS ON MARRIED WOMEN ' S LABOR FORCE PARTICIPATION Wl. *Review of Economics and Statistics*, 74(1):83–90.

Ettner, S. L. (1993). Do elderly medicaid patients experience reduced access to nursing home care? *Journal of Health Economics*, 12(3):259–280.

Ettner, S. L. (1996). The opportunity costs of elderly care. *The Journal of Human Resources*, 31(1):189–205.

Fukahori, R., Sakai, T., and Sato, K. (2015). The Effects of Incidence of Care Needs in Households on Employment, Subjective Health, and Life Satisfaction among Middle-aged Family Members. *Scottish Journal of Political Economy*, 62(5):518–545.

Gertler, P. J. (1989). SUBSIDIES, QUALITY, AND THE REGULATION OF NURSING HOMES. *Journal of Public Economics*, 38:33–52.

Gertler, P. J. (1992). Medicaid and the Cost of Improving Access to Nursing Home Care. *The Review of Economic and Statistics*, 74(2):338–345.

Grabowski, D. C. (2001). Medicaid reimbursement and the quality of nursing home care. *Journal of Health Economics*, 20:549–569.

Grabowski, D. C., Gruber, J., and Angelelli, J. J. (2008). Nursing Home Quality as a Common Good. *The Review of Economic and Statistics*, 90(4):754–764.

Hassink, W. H. J. and Berg, B. V. D. (2011). Social Science & Medicine Time-bound opportunity costs of informal care : Consequences for access to professional care , caregiver support , and labour supply estimates. *Social Science & Medicine*, 73(10):1508–1516.

Heitmueller, A. (2007). The chicken or the egg?: Endogeneity in labour market participation of informal carers in England. *Journal of Health Economics*, 26:536–559.

Heitmueller, A. and Inglis, K. (2007). The earnings of informal carers: Wage differentials and opportunity costs. *Journal of Health Economics*, 26(4):821–841.

Hoerger, T. J., Picone, G. A., and Sloan, F. A. (1996). Public Subsidies , Private Provision of Care and Living Arrangements of the Elderly. *The Review of Economics and Statistics*, 78(3):428–440.

Hurd, M. D., Meijer, E., Moldoff, M. B., and Rohwedder, S. (2016). Improved Wealth Measures in the Health and Retirement Study Asset R econciliation and Cross-Wave Imputation.

WR1150, April 2016.

Ichimura, H., Shimizutani, S., and Hashimoto, H. (2009). JSTAR First Results 2009 Report. *RIETI Discussion Papers*, September 2009 09-E-047.

Ishii, K. (2015). Informal care and employment status of Japanese middle aged women : a study using JSTAR. *https://basepub.dauphine.fr/handle/123456789/15047?locale-attribute=en.*

Kondo, A. (2016). Availability of Long-term Care Facilities and Middle-aged people's Labor Supply in Japan. *JCER Working Paper AEPR series*, 2016-1-4.

Kwon, S. (2009). The introduction of long-term care insurance in South Korea. *Eurohealth*, 15(1):15–16.

Leigh, A. (2010). Informal care and labor market participation . *Labour Economics*, 17(1):140–149.

Lilly, M. B., Laporte, A., and Coyte, P. C. (2007). Labor Market Work and Home Care's Unpaid Caregivers: A Systematic Review of Labor Force Participation Rates, Predictors of Labor Market Withdrawal, and Hours of Work. *The Milbank Quarterly*, 85(4):641–690.

Lilly, M. B., Laporte, A., and Coyte, P. C. (2010). Do they care too much to work ? The influence of caregiving intensity on the labour force participation of unpaid caregivers in Canada. *Journal of Health Economics*, 29(6):895–903.

Meng, A. (2013). Informal home care and labor-force participation. *Empirical Economics*, 44:959–979.

Michaud, P.-c., Heitmueller, A., and Nazarov, Z. (2010). A dynamic analysis of informal care and employment in England. *Labour Economics*, 17(3):455–465.

Moriwaki, D. (2016). Informal Care and Labor Supply Among Japanese Middle Aged. *Mimeo.*

Nishimura, Y. and Oikawa, M. (2017). Effects of Informal Elderly Care on Household Time Use and Caregiver's Health: Exploitation of Government Intervention on the Supply Side of Elderly Care Market. *Mimeo.*

Norton, C. (1992). Incentive regulation of nursing homes. *Journal of Health Economics*, 11:105–128.

Nyman, J. A. (1985). PROSPECTIVE AND 'COST-PLUS' MEDICAID REIMBURSEMENT, EXCESS MEDICAID DEMAND, AND THE QUALITY OF NURSING HOME CARE.

Journal of Health Economics, 4:237–259.

Nyman, J. A. (1988). Excess Demand, the Percentage of Medicaid Patients, and the Quality of Nursing Home Care. *Journal of Human Resources*, 23(1):76–92.

Nyman, J. A. (1994). The Effects of Market Concentration and Excess Demand on the Price of Nursing Home Care. *The Journal of Industrial Economics*, 42(2):76–92.

Schaffer, M. (2010). xtivreg2: Stata module to perform extended iv/2sls, gmm and ac/hac, liml and k-class regression for panel data models. Available online via `http://repec.org/bocode/x/xtivreg2.html` (last accessed 19.9.2016).

Schulz, E. (2010). The Long-Term Care System for the Elderly in Germany. *ENEPRI Research Report*, 78.

Shimizutani, S., Suzuki, W., and Noguchi, H. (2008). The socialization of at home elderly care and female labor market participation: Micro level evidence from Japan. *Japan and the World Economy*, 20(1):82–96.

Sugawara, S. and Nakamura, J. (2014). Can formal elderly care stimulate female labor supply? The Japanese experience. *Journal of the Japanese and International Economies*, 34:98–115.

Tamiya, N., Noguchi, H., Nishi, A., Reich, M. R., Ikegami, N., Hashimoto, H., Shibuya, K., Kawachi, I., and Campbell, J. C. (2011). Japan : Universal Health Care at 50 years 4 Population ageing and wellbeing : lessons from Japan ' s long-term care insurance policy. *The Lancet*, 378(9797):1183–1192.

Tan, F. E. S. (2000). The Competing Demands of Paid Work and Parent Care. *Research on Aging*, 22(2):165–187.

Trong, H. and Brian, L. (2014). Social Science & Medicine The effect of unpaid caregiving intensity on labour force participation : Results from a multinomial endogenous treatment model. *Social Science & Medicine*, 100:115–122.

Van Houtven, C. H., Coe, N. B., and Skira, M. M. (2013). The effect of informal care on work and wages. *Journal of Health Economics*, 32(1):240–252.

Van Houtven, C. H. and Norton, E. C. (2004). Informal care and health care use of older adults.

Journal of Health Economics, 23(6):1159–1180.

Vermeulen, F. (2002). COLLECTIVE HOUSEHOLD MODELS: PRINCIPLES AND MAIN RESULTS. *JOURNAL OF ECONOMIC SURVEYS*, 16(2):533–564.

Wakabayashi, C. and Donato, K. M. (2005). The consequences of caregiving: Effects on women's employment and earnings. *Population Research and Policy Review*, 24:467–488.

Wolf, D. A. and Soldo, B. J. (1994). Married Women's Allocation of Time to Employment and Care of Elderly Parents. *The Journal of Human Resources*, 29(4):1259–1276.

Won, C. W. (2013). Elderly long-term care in Korea. *Journal of Clinical Gerontology & Geriatrics*, 4:4–6.

Yamada, H. and Shimizutani, S. (2015). Labor market outcomes of informal care provision in Japan. *The Journal of the Economics of Ageing*, 6:79–88.

Chapter 4: Effects of the Japanese Social Security System on Retirement Behavior

1 Introduction

Since the Japanese social security system has started to put pressure on the Japanese budget, there has been many discussions about how the Japanese government should reform the social security system. However, these discussions proceed without analyzing how elderly people will behave after a reform on the social security system. This paper provides an evidence to answer this question and the goal of this study is to predict what will happen after a reform on the Japanese social security system. This study is the first analysis to simulate the Japanese male elderly's labor supply under some counterfactual reforms by using a household panel data of the Japanese elderly people. After 2000, the Japanese economics researchers have started to answer this question by using a dynamic framework. In the preceding studies, there are two types of studies analyzing retirement by using a dynamic model. One is the analysis to use the option value model. The other is the analysis which does not use the option value model.

Iwamoto (2000) is the first study to answer this question by using a model, which is not the option value model. Iwamoto (2000) derives the reduced form from a dynamic model without uncertainty. In this research, to estimate the model parameters, this study sets some assumptions because a panel dataset with enough individual information was not available. For example, it is assumed that there is no cohort effect in wages, the time preference is equal to the interest rate. Since a panel data was not available, the lagrange multiplier identifying a fixed effect is assumed to be a function having only a respondent birth year. Higuchi and Yamamoto (2002) is the second work to answer this question. Higuchi and Yamamoto (2002) use a model which has three available leisure options such as the leisure when people work full-time as a full-time worker, the leisure when people work as a part-time worker and the leisure when people do not work. In Higuchi and Yamamoto (2002), the discount factor is zero because a panel dataset was not available.[1] Oshio and Oishi (2000), Oshio and Sato Oishi (2003) and Oshio and Sato Oishi (2004) are the studies which answer this question by using the option value model. In these studies, the model without saving is analyzed. An individual's indirect utility function over work and leisure is defined. The option value is defined as the difference between indirect utility from retirement at optimal age and indirect utility from retiring today. The effect of the incentive measure to work such as the option value is estimated. Oshio et al. (2011) is the most recent work answering the question by using the option value model. Oshio et al. (2011) shows the simulation results of some counterfactual social security reforms based on the estimated results. In Oshio and Oishi (2000), Oshio and Sato Oishi (2003), Oshio and Sato Oishi (2004) and Oshio et al. (2011), there is no representation of direct connection between labor supply and the option value.

French (2005) has introduced asset accumulation in dynamic life cycle model which Rust and Phelan (1997) does not considered. French (2005) includes three types uncertainty in the model, which are uncertainty of death, health shock and wage uncertainty. French (2005) has clarified that the effect of social security benefit, pension and wages is more important factor for retirement behavior than the health shock effect. French and Jones (2011) has also analyzed the relationship between the health insurance and retirement in a similar framework.

[1] It might be appropriate to make the assumption that the discount factor is zero under some conditions. However, in an empirical research analyzing households which consider the utility in the future, it is difficult to justify this argument because it doesn't seem meaningful to use a dynamic model under zero discount factor. In addition, it is not justified the assumption that the discount factor is zero according to the preceding studies. For example, see Karuma and Ikeda (2006).

In this study, the goal is to do a Japanese empirical research of the Japanese male elderly retirement behavior. This study has advanced the Japanese previous studies. Iwamoto (2000) constructs a model without uncertainty. Health uncertainty and death uncertainty are introduced in this study. Oshio and Oishi (2000), Oshio and Sato Oishi (2003) and Oshio and Sato Oishi (2004) do not include saving. In this study, saving is included in the model. However, wage uncertainty is not considered because the JSTAR (the Japanese Study of Aging and Retirement), which is used for estimating wage equation, has a few waves. It is difficult to reveal the structure of wage uncertainty. If the JSTAR project continues in the future, it is possible that the wage process of the Japanese elderly will be clarified. At this moment, there is a limit of estimating wage uncertainty because the information of the wage transition is not enough. We also cannot get the information to calculate exact amount of pension which an individual will receive.

The remainder of this paper is organized as follows: section 2 reviews analysis method; section 3 discusses the model and data uses; section 4 explains the estimation procedure; section 5 concludes this paper and identifies the scope for future research.

2 Analysis method

There are two analysis steps in this study. First, we construct a model which describes a dynamic utility maximization problem of the Japanese male elderly. Second, we estimate the parameter of the utility function. We estimate the parameters of the utility function by using the MSM (the method of simulated moments).

In this study, there are four steps in the estimation procedure. First, the exogenous data processes are estimated. Second, the initial distribution is produced. Third, we produce the sample moments available from the data which we will match with the simulated moments. Finally, based on the estimated exogenous data processes and the initial distribution, the iteration of simulation is carried out to minimize the distance between the simulated moments available from the simulated data and the sample moments available from the data until the iteration finishes. In the estimation, the most important point is the characteristic of the data which we have. We use the labor participation information from the JSTAR which has only two waves. It is desirable to estimate a dynamic model that both the cohort group used for producing the initial distribution and the cohort group used for estimating the sample moments should be the same. However, we have only two waves. We will use the cohort group used for estimating the sample moments, which is different from the cohort group used for the initial distribution. To cope with this problem, we use the estimation strategy of French (2005). However, the estimation method is different from French (2005) because the method by French (2005) causes a problem when we estimate the structural parameters. We discuss this point in section A.

There are many important points in French (2005) for this study. First, the cohort group used for the initial distribution and the cohort group used for estimating the sample moment are different. Second, the exogenous data process and the life cycle profiles for asset, labor hour and participation rate are all adjusted to 1940 birth year cohort effect to consider the difference in the cohort gourd between the sample moments and the initial distribution. In French (2005), the initial distribution is produced based on the individuals aged 29-31. Thus, there must be the assumption that the difference of the initial distribution is not so different among the individuals with age 29-31 although French (2005) clearly does not describe. If there is no difference in the initial distribution of the individuals with age 29-31, only problem is the difference between cohort group used for estimating the sample moments of the life cycle profiles (e.g., labor participation rate) and the cohort group used for estimating the initial distribution.[2] Thus, French (2005) cope with this problem by adjusting life cycle profile data and exogenous data process to 1940 birth year cohort effect. We use this estimation

[2]There is also a cohort effect in an exogenous probability process.

strategy of French (2005). In other words, the initial distribution is estimated by using the individuals with aged 55-60 in the JSTAR 2007. Next, we adjust the individual profiles (e.g., labor participation rate) to age 60 cohort effect in the JSTAR 2007. Then, we also adjust the exogenous data processes to age 60 cohort effect in the JSTAR 2007. We assume that the joint initial distribution of the cohorts aged 55-60 in the JSATR 2007 are not different among the cohorts aged 55-60.[3] We also assume that the exogenous data processes are not so different among the cohorts aged 55-60 in the JSATR 2007. In Japan, there are some social security programs which influence the incentive to supply labor among the elderly. The basic pension and the earning-related pension are included in this study. We omit the Japanese tax system.

[3]To increase the number of observations, we have used the individuals aged 55-60 in the JSTAR 2007 when estimating the initial distribution. It is desirable that only the cohort aged 60 in the JSATR 2007 is used for the estimation.

3 Analysis framework

3.1 Model

The model is simple. French (2005) and French and Jones (2011) have a similar framework. The households with a male respondent are analyzed. A male respondent in the household maximizes a household expected utility at age t ($t = 61, \ldots, 94$). Let $T = 94$. The utility function after age 86 is the function (1). The household member cannot supply labor after age 86. C_t is a consumption at period t. $M_t \in \{good, bad\}$ describes a health state.

$$U_t(C_t, H_t) = \frac{1}{1-\eta} C_t^{1-\eta} \quad \text{if } 86 \leq t \leq 94 \tag{1}$$

The utility function before age 85 is the function (2).

$$U_t(C_t, H_t) = \frac{1}{1-\eta}\{C_t^{\gamma}(L - H_t - \phi_1 P_t - \phi_2 I\{M_t = bad\})^{1-\gamma}\}^{1-\eta} \quad \text{if } 61 \leq t \leq 85 \tag{2}$$

$b(A_t)$ is a bequest function. A_t is an asset at period t.

$$b(A_t) = \theta \frac{1}{1-\eta}(A_t + K_b)^{(1-\eta)\gamma} \tag{3}$$

H_t is a **weekly** labor hour. H_t is in the set (4).

$$H_t \in \{a|0,5 \leq a \leq L - \phi_1 \cdot 1 - \phi_2 I\{M_t = bad\}\} \tag{4}$$
$$\text{if} \quad 61 \leq t \leq 85$$
$$H_t \in \{0\}$$
$$\text{if} \quad 86 \leq t \leq 94$$

P_t describes labor participation in the following way.

$$P_t = \begin{cases} 1 & \text{if} \quad H_t > 0 \\ 0 & \text{if} \quad H_t = 0 \end{cases} \tag{5}$$

Male mortality rate s_{t+1}^{man} is a mortality rate at age $t+1$ given living at age t. The mortality rate depends on only respondent's age.

$$s_{t+1}^{man} = s^{man}(t) \tag{6}$$
$$s_{T+1}^{man} = 1$$

Health status follows a two-state transition matrix at each age in the following way. (7) is an example.

$$\pi_{good,bad,t} = prod(M_{t+1} = good|M_t = bad, t) \tag{7}$$

The logarithm of wage per hour at age t, $\ln W_t$, is a **deterministic** function of the logarithm of working hours. After age 75, the wage process is the same with the wage process at $t = 75$.

$$\text{If} \quad 61 \leq t \leq 75, \tag{8}$$

$$\begin{cases} \ln W_t = \alpha_1 \ln H_t + \alpha_2 I\{t \geq 65\} + \sum_{k=1}^{K} \delta_k t^k \\ \text{if} \quad H_t > 0 \\ W_t = 0 \\ \text{if} \quad H_t = 0 \end{cases}$$

$$\text{If} \quad 76 \leq t \leq 94, \tag{9}$$

$$\begin{cases} \ln W_t = \alpha_1 \ln H_t + \alpha_2 I\{t \geq 65\} + \sum_{k=1}^{K} \delta_k 75^k \\ \text{if} \quad H_t > 0 \\ W_t = 0 \\ \text{if} \quad H_t = 0 \end{cases}$$

The asset accumulation follows the equation (10). The tax system is omitted. $weeks$ is an adjustment factor of labor hours, which is the number of annual weeks. $pb1st_t$ is a basic pension covering every people in Japan.[4] $pb2nd_t$ is a full amount of earning-related pension covering only company employees or public servants.[5] $\widetilde{pb2nd_t}$ is an actual earning-related pension. The full amount of $pb2nd_t$ can be reduced depending on the recipient's working hours. $\widetilde{pb2nd_t}$ is determined by the government based on the amount of $pb2nd_t$.[6] $pb2nd_t$ is exogenously determined after age 61. $pb1st_t$ is the quantity of the full basic pension amount, which does not change after age 61. The eligibility age of the basic pension is age 65;

[4] The pension is called *Kiso-Nennkin* in Japan.

[5] The pension is called *Kousei-Nennkin* or *Kyosai-Nennkin* in Japan.

[6] We explain how to calculate the full amount of $\widetilde{pb2nd_t}$ in section A.

$pb1st_t = 0$ if $61 \leq t \leq 64$.

$$\begin{aligned} A_{t+1} = (1+r)A_t + W_t H_t weeks - C_t + \\ I\{spousestatus_t = 1\} spouseincome_t \\ + pb1st_t + \widetilde{pb2nd_t} \\ A_{t+1} \geq 0 \end{aligned} \tag{10}$$

$spousestatus_t = spousestatus_{61}$ $(62 \leq t \leq 94)$ and $spousestatus_{61} = 1$ if a male respondent has a spouse at $t = 61$. The variable $spouseincome_t$ is an annual income of a spouse. It is assumed that a spouse does not die before a male respondent. And a male respondent does not marry and divorce after age 61.[7] A spouse income is deterministic process depending on only respondent's age.

Finally, the recursive formulation is the equation (11). This is a finite horizon simple problem with a Markov process. The recursive formulation is the problem (11) such that (6)-(10). Let $X_t = \{A_t, M_t, spousestatus_t, spouseincome_t, pb1st_t, pb2nd_t, t\}$. In addition, $V_{T+1}(X_{T+1}) = b(A_{T+1})$.

$$V_t(X_t) = \max_{C_t, H_t}\{U_t(C_t, H_t) + \beta[(1 - s^{man}_{t+1}) \sum_{M_{t+1} \in \{good, bad\}} V_{t+1}(X_{t+1}) prob(M_{t+1}|M_t, t) + s^{man}_{t+1} b(A_{t+1})]\} \tag{11}$$

3.2 Data

In this section, we explain three datasets used in this paper. We briefly explain our datasets in Table 1. We explain how to use datasets in each section. In Japan, there are

[7] In fact, people scarcely divorce after age 60 in Japan. The website: http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001082332 (in Japanese) (Table 10-7) (Ministry of Health, Labour and Welfare, Monthly Labour Survey).

three panel datasets targeting the elderly. They are the JSTAR (Japanese Study of Aging and Retirement), the NUJLSOA (Nihon University Japanese Longitudinal Study of Aging), the NSJE (National Survey of the Japanese Elderly). We use these three datasets. These datasets have different characteristics. The JSTAR has enough information of each household such as asset, labor, wage and social security. However, the JSTAR has a few waves. The NUJLSOA has enough health information of each individual in detail. However, the NUJLSOA has also only two waves and include the small number of economic variables. The NSJE has four waves although the NSJE has many missing values in economic variables. The JSTAR and the NUJLSOA has a two-year interval although the NSJE has a three-year interval. This characteristic is important for estimating the health transition probability.

We use three datasets to utilize the characteristics in each dataset. First, when we estimate the health transition probability, we use the JSTAR and the NUJLSOA because these datasets have the same interval (two-year). Second, when we estimate a male mortality rate, we use all three datasets because the mortality information is not related to the size of each interval. Third, when we estimate a wage process, we use the JSTAR because this is a only panel data which has the details of both wage and labor supply. When we estimate the parameters of the utility function, we use the method of simulated moments by adjusting the fixed effect of the moments and the exogenous data processes to the aged 60 cohort in the JSTAR 2007. We estimate the initial joint distribution from the cohort aged 55 to 60 in the JSTAR 2007.

3.3 Discussion of the model assumptions

3.3.1 Labor hour

There are many assumptions in this study. We will discuss these points. We use weekly hours worked in this study. We use weekly hours worked in the JSTAR when we estimate the

Table 1: Panel Data Sets Used in This Analysis

	JSTAR	NSJE	NUJLSOA	
Response Rate (Wave 1)	59.4	67.2	74.6	
Age	50-75	60-	65-	
Number of Waves	2	2	4	
Intervals (years)	2	3	2	
Characteristics -Health -Mortality -Social Security -Economic Activity	 Rich Rich Rich Rich	 Rich Rich Not Rich Not Rich	 Rich Rich Not Rich Not Rich	
Where is the data used?	Mortality rate Health Transition Wage Process Moment Initial Distribution	Mortality rate	Mortality rate Health Transition	

wage process. Usually, in the preceding studies, an annual hour worked of a respondent has been used. There are two reasons why we use weekly hours worked. Only the JSTAR provides the detail information of both hours worked and wage. First, the questionnaire to ask how many days an individual works is different between 2007 and 2009. In 2007, only weekly hours worked are available. Then, the calculation method has to be changed when an annual hour worked is calculated. As a result, to keep consistency of the calculation, we use only weekly hours worked. Second, the value of annual hours worked is large. Please see section A. Finally, we assume that the adjustment of annual weeks in the model is equal to 42.95 in the budget constraint and we use only hours worked of an respondent who works throughout the year. 42.95 is the average weeks worked of male respondents who work throughout the year in the JSTAR 2009. We use only weekly hours worked of male respondents who work throughout the year.

3.3.2 Starting age and terminal age

We consider the utility maximization problem between age 61 and age 94. In the model, there is a health uncertainty between age 61 and age 85. There are some reasons why we consider the problem between age 61 and age 94. The most important reason why the starting age is age 61 is a wage uncertainty. As we mentioned, the structure of the wage uncertainty is not clarified. We choose the age when the wage uncertainty doesn't matter in a wage process. In Japan, it seems that most of the male elderly aged 50s working for a institution has a wage uncertainty. In addition, the sample size of health state before age 61 is very small. The small sample size becomes an important problem when we estimate the health transition probability. As a result, we set that the starting age is 61 because of the two reasons. With respect to the terminal age, the terminal age 94 is decided based on a death transition probability prediction. When the age is 94, the prediction of the male death transition probability is less than zero. Finally, the health uncertainty exists only between age 61 and age 85. After age 85, we assume that an individual only consumes. The health transition information after age 85 is not available and we omit the uncertainty of health transition after age 85. After age 85, most Japanese elderly don't work (Figure 1). However, we cannot clearly confirm whether this fact is caused by an exogenous factor (he /she is not able to work because of his age even if he/she wants to work) or an endogenous factor (he/she does not work because he/she does not want to work) because, in the NUJLSOA and the NSJE, which covers male respondents aged greater than age 85, there is no question about whether he is looking for a job or not. In Figure 1, we must show an evidence that the Japanese elderly usually do not work after age 85.

3.3.3 Wage process

There are two assumptions in the wage process. First, the wage is assumed to be determined by age and labor hours at t. This wage process should be modified in the future.

Second, it is assumed that after age 75, the wage process is the same as the process at $t = 75$. This assumption is based on the availability of the data. In the JSTAR, there is only information about wage between age 55 and age 75. We assume that the wage process is the same as the process at $t = 75$. In public aggregated data, there is little information about wage of respondents aged more than age 75. We cannot verify whether this assumption is valid or not. Reconsidering this assumption can be also important future work.

at_work_respondent_w1	age_respondent_w1 65	66	67	68	69	70	71	72	73	74	Total
1	61 57.01	51 48.11	70 51.47	37 36.27	49 44.14	40 47.62	29 39.19	42 43.75	30 33.71	23 33.33	672 33.30
2	45 42.06	55 51.89	66 48.53	62 60.78	61 54.95	44 52.38	45 60.81	54 56.25	58 65.17	44 63.77	1,319 65.36
4	1 0.93	0 0.00	0 0.00	3 2.94	1 0.90	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	14 0.69
5	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 1.12	2 2.90	13 0.64
Total	107 100.00	106 100.00	136 100.00	102 100.00	111 100.00	84 100.00	74 100.00	96 100.00	89 100.00	69 100.00	2,018 100.00

at_work_respondent_w1	age_respondent_w1 75	76	77	78	79	80	81	82	83	84	Total
1	38 31.40	52 37.14	25 24.75	27 24.32	27 25.47	8 10.67	16 21.62	7 10.00	13 21.31	4 9.52	672 33.30
2	82 67.77	84 60.00	75 74.26	82 73.87	78 73.58	66 88.00	56 75.68	61 87.14	47 77.05	37 88.10	1,319 65.36
4	0 0.00	3 2.14	0 0.00	1 0.90	0 0.00	0 0.00	2 2.70	2 2.86	1 1.64	0 0.00	14 0.69
5	1 0.83	1 0.71	1 0.99	1 0.90	1 0.94	1 1.33	0 0.00	0 0.00	0 0.00	1 2.38	13 0.64
Total	121 100.00	140 100.00	101 100.00	111 100.00	106 100.00	75 100.00	74 100.00	70 100.00	61 100.00	42 100.00	2,018 100.00

at_work_respondent_w1	age_respondent_w1 85	86	87	88	89	90	91	92	93	94	Total
1	11 25.00	4 10.81	1 11.11	3 21.43	2 18.18	1 12.50	0 0.00	0 0.00	0 0.00	1 25.00	672 33.30
2	33 75.00	32 86.49	8 88.89	11 78.57	9 81.82	6 75.00	4 80.00	6 100.00	3 100.00	3 75.00	1,319 65.36
4	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	14 0.69
5	0 0.00	1 2.70	0 0.00	0 0.00	0 0.00	1 12.50	1 20.00	0 0.00	0 0.00	0 0.00	13 0.64
Total	44 100.00	37 100.00	9 100.00	14 100.00	11 100.00	8 100.00	5 100.00	6 100.00	3 100.00	4 100.00	2,018 100.00

at_work_respondent_w1	age_respondent_w1 96	97	Total
1	0 0.00	0 0.00	672 33.30
2	1 100.00	1 100.00	1,319 65.36
4	0 0.00	0 0.00	14 0.69
5	0 0.00	0 0.00	13 0.64
Total	1 100.00	1 100.00	2,018 100.00

Figure 1: Working Situation of Male Respondents (NUJLSOA, Only Respondents)
1=at work, 2=retirement, 3=student, 4=housework, 5=retirement from housework

3.3.4 Spouse income

In French and Jones (2011), a spouse income is assumed to be deterministic function of the household head's age and health status. However, we do not assume that a spouse income does not depend on the health state of a male respondent because, in Japan, it seems that there is little relationship between a spouse income and a health status of a male household head. Also in French (2005), French and Jones (2011), whether a male respondent has a spouse or not is not considered. We consider whether a male respondent has a spouse or not. It seems that whether a male respondent has as spouse or not influences a male worker's labor supply. The difference in whether a male respondent has a spouse or not causes the difference in labor supply. Anyway, we must show the evidence about the above two points. With respect to a female spouse income, there is no systematic difference depending on the male respondent health state. The average is a little higher because the maximum of income at good health is higher than the income at bad health (Table 2). The frequency around 40-50 hours worked is different because a single male respondent works more (Table 3). It is strongly possible that a male respondent with a spouse works less because he has a spouse income. This tendency is also observed at 2009.

Table 2: Spouse Income (JSTAR, Only Male Respondents, 2007 and 2009)

male age	mean (yen)	s.d	health state of male respondents
61-65	2272225	1630911	good
61-65	2122800	1290308	bad
66-70	2282894	1555424	good
66-70	2178824	1115343	bad
71-75	2111462	1416986	good
71-75	2213929	1220932	bad

Table 3: Male Hours Worked at 2007 (JSTAR, Only Male Respondents)

hour	10	20	30	40	50	60	70	sum
single	2	7	11	36	61	24	12	153
(freq)	1.31	4.58	7.19	23.53	39.87	15.69	7.84	100
with spouse	77	70	81	339	357	192	114	1230
(freq)	6.26	5.69	6.59	27.56	29.02	15.61	9.27	100

4 Estimation

4.1 Estimation of the exogenous data processes

4.1.1 Mortality rate

The mortality transition probability is estimated by using all three datasets. The information about when a respondent died is available in the NSJE, the NUJLSOA and the JSTAR. The equation (12) is estimated to control the cohort effect.

$$I\{\, living\, at\, t_i + 1\} = f_i + \sum_{k=1}^{K} \gamma_k t_i^k + u_{it_i+1} \tag{12}$$

The estimation result is shown (Table 4, Table 5). The value of mortality transition probability is shown (Table 6). The fixed effect f_i is adjusted. We replace the fixed effect with the cohort effect of a respondent aged 60 at the JSTAR 2007.

Table 4: Probability of Living at t_i+1 Given Living at t_i (Fixed effect estimator of men)

variable	coefficients	S.E.
t_i	8.796	2.227
t_i^2	-.254	.062
t_i^3	.004	.001
t_i^4	-.254e-04	.602e-05
t_i^5	.702e-07	.166e-07
Number of obs	17367	

Table 5: Probability of Living at t_i+1 Given Living at t_i (Fixed effect estimator of women)

variable	coefficients	S.E.
t_i	-6.142	3.548
t_i^2	.170	.095
t_i^3	-.002	.001
t_i^4	.165e-4	.831e-05
t_i^5	-.460e-07	.219e-07
Number of obs	17045	

Table 6: Mortality Rate at $t_i + 1$ Given Living at t_i(Fixed effect adjusted value of men)

age	value	age	value	age	value	age	value
61	0.0011	71	0.0597	81	0.2551	91	0.806
62	0.0086	72	0.0677	82	0.2941	92	0.872
63	0.0154	73	0.0772	83	0.3371	93	0.938
64	0.0215	74	0.0889	84	0.3843		
65	0.0271	75	0.1029	85	0.4355		
66	0.0323	76	0.1197	86	0.4904		
67	0.0372	77	0.1395	87	0.5488		
68	0.0421	78	0.1628	88	0.6102		
69	0.0473	79	0.1896	89	0.6739		
70	0.0531	80	0.2204	90	0.7395		

4.1.2 Health transition probability

We will explain how to make a health measure. We use a subjective health state question in each dataset. In the JSTAR, there is a question, "Choose one option which describes your current health state.". A respondent chooses one option from five options, "1. Fairly good 2. Good 3. Middling 4. Not good 5. Bad ". The value of a respondent health state at age t is equal to zero if the option number is 1 or 2 or 3 (Health state is good.). The value of a respondent health state at age t is equal to one if the option number is 4 or 5 (Health state is bad.). Also in the NSJE and the NUJLSOA, there is a question which is similar to this question in the JSTAR. Also in the NSJE and the NUJLSOA, we produce the same health measure in the same way.

When estimating the health transition probability, we estimate aggregately the following value, $Pr(h_{it+2} = bad|h_{it}, age_{it})$ from the JSTAR and the NUJLSOA. In fact, we aggregate the data among an appropriate age range to estimate a health transition probability. However, there are two problems in this method[8]. We show the calculated health transition probability in Table 7.

Table 7: Health Transition Probability at $t_i + 1$ Given t_i and Health State of t_i

age	from good to bad	age	from bad to bad
61-64	0.0832	61-64	0.4056
65-69	0.0912	65-69	0.5101
70-74	0.1633	70-74	0.5254
75-79	0.1782	75-79	0.6247
80-	0.2926	80-	0.6543

[8]First, this method cannot control fixed effect of each individual. We assume that there is no fixed effect in health transition probability. Second, this method also has a discrepancy with the value $Pr(h_{it+1} = bad|h_{it}, age_{it})$ which we want to use.

4.1.3 Wage process

We estimate the equation (13) and (14) as the wage process. There are three components. $\phi log(h_{it})$ is a component to represent a premium of wage per hours worked. $I\{t \geq 65\}$ is a component to control a Japanese employment practice. Age polynomials control an individual's productivity (Aaronson and French (2004)).

$$
\begin{aligned}
log(h_{it}) &= f_i + \sum_{k=1}^{K} \gamma_k age_{it}^k + \beta_1 I\{M_t = bad\} \qquad (13)\\
&+\beta_2 I\{age_{it} \geq 60 \ and \ Earning-related \ pension_{it} = 1\}\\
&+\beta_3 I\{age_{it} \geq 65\} + \beta_4 year2009_t + e_{it}
\end{aligned}
$$

$$
log(w_{it}) = z_i + \sum_{k=1}^{K} \delta_k age_{it}^k + \phi log(h_{it}) + \alpha_1 I\{age_{it} \geq 65\} + \alpha_2 year2009_t + u_{it} \qquad (14)
$$

Aaronson and French (2004) estimate a labor supply premium per hour on wage. Aaronson and French (2004) uses an indicator showing that a respondent's age is over age 65. People start to receive social security benefit from age 65. This indicator is an instrumental variable for labor hour. However, in Japan, this is not appropriate because a male worker's wage is decreased because of age itself between 60 to 65.[9] In Japan, other instrumental variables are needed.

There is a clear decrease in log wage per hour after age 65 (Figure 2). This is a statistic from only working respondents. We want to discuss the existence of a premium per hours worked. There is a clear distribution difference between workers who work more than 40 hours and workers who do not work more than 40 hours (Figure 3). In Japan, 40 is an

[9]The age indicator is included in the wage process directly. The indicator is not an instrumental variable in Japan. In Japan, there is the *Teinenn Taisyoku Seido* which means that an institution forces a male worker to retire or change his position with lower wage between age 60 to age 65. The point of this practice is that the retirement age is determined by the custom which is not necessarily related to a worker's productivity. In 2007, about 70 percent of male correspondents retire because of the *Teinenn Taisyoku Seido*. About 86 percent respondents who experienced the *Teinenn Taisyoku Seido* retire between age 60 to age 65.

log_wageperhour07	respondent_age_man 58	59	60	61	62	63	64	65	66	67	68	Total
5	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 2.94	4 9.52	1 3.23	1 3.33	7 1.48
7	6 10.34	12 15.79	8 18.18	9 24.32	11 25.58	15 35.71	10 27.03	9 26.47	19 45.24	13 41.94	12 40.00	124 26.16
9	46 79.31	58 76.32	35 79.55	25 67.57	30 69.77	24 57.14	24 64.86	21 61.76	15 35.71	14 45.16	14 46.67	306 64.56
11	6 10.34	5 6.58	0 0.00	3 8.11	1 2.33	2 4.76	3 8.11	3 8.82	3 7.14	3 9.68	2 6.67	31 6.54
13	0 0.00	1 1.32	0 0.00	0 0.00	1 2.33	1 2.38	0 0.00	0 0.00	1 2.38	0 0.00	1 3.33	5 1.05
15	0 0.00	0 0.00	1 2.27	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	0 0.00	1 0.21
Total	58 100.00	76 100.00	44 100.00	37 100.00	43 100.00	42 100.00	37 100.00	34 100.00	42 100.00	31 100.00	30 100.00	474 100.00

Figure 2: Male Workers' Log Wage per Hour at 2007 (JSTAR, Only Respondents)

log_wageperhour07	hourless40_07 0	1	Total
5	5 1.47	2 1.48	7 1.48
7	66 19.47	58 42.96	124 26.16
9	246 72.57	60 44.44	306 64.56
11	18 5.31	13 9.63	31 6.54
13	3 0.88	2 1.48	5 1.05
15	1 0.29	0 0.00	1 0.21
Total	339 100.00	135 100.00	474 100.00

Figure 3: Male Workers' Log Wage per Hour and 40 Hours Weekly Labor Hour Dummy at 2007 (JSTAR, Only Respondents, the Same Respondent in Figure 3)
An individual works less than 40 if $hourless40$=1

parttime_07	hourless40_07 0	1	Total
0	463 92.42	122 66.30	585 85.40
1	38 7.58	62 33.70	100 14.60
Total	501 100.00	184 100.00	685 100.00

Figure 4: 40 Hours Weekly Hours Worked Dummy and Part Time Dummy at 2007 (JSTAR, Only Respondent, Age 58-68 Male Respondents)
A respondent works as an employee and a part time position if $parttime07$=1

important point which distinguishes a full time worker from a part time worker (Figure 4).[10]

We explain the estimation process. We explain how to make the data of wage per hour. As we mentioned, we should not use annual working hours which are used as in the preceding studies. In the preceding studies, an annual wage is divided by annual working hours to make wage per working hours. However, in this study, a wage per working hours except a bonus is produced by directly using questions in the JSTAR. For example, with respect to monthly salary workers, to produce monthly payment per working hours, a monthly payment is divided by the amount of four times weekly working hours. To produce a bonus per working hours, a bonus is divided by a product between the adjustment of annual weeks[11] and weekly working

[10] In Japan, the weekly hours worked of full time workers are concentrated in the range between 40 hours and 60 hours.

[11] 42.95

hours. We use only wage per working hours of a male respondent who works throughout the year to estimate the wage process. This keeps consistency.

$I\{age_{it} \geq 60\ and\ Earning - related\ pension_{it} = 1\}$ is an indicator showing that a respondent is over age 60 with an earning-related pension eligibility. Aaronson and French (2004) use an indicator showing that a respondent's age is over age 65. This is an instrumental variable of working hours. However, in Japan, an indicator showing that a respondent's age is over age 65 should be in the wage equation directly. We use both a health state and an indicator showing that a respondent is over age 60 with an earning-related pension eligibility[12] as other instrumental variables while Aaronson and French (2004) does not use a health state as an instrumental variable. Table 8 and Table 9 are the estimation results. We use this result as the wage equation in the model. In Aaronson and French (2004), the value of ϕ is from 0.371 to 0.949 (the PSID, the HRS, the CPS March, the CPS ORG) with respect to male workers. In French (2005), ϕ is set to 0.415 which is not estimated. In this study, we estimate the value of ϕ. We replace the fixed effect in the wage equation with the cohort effect of individuals aged 60 at the JSTAR 2007.

Table 8: Fixed Effect Estimator of Male Respondents Aged More Than 60 at 2009 (Second stage)

variable	coefficients	S.E.
$log(h_{it})$	.282	.660
$year2009_t$	.086	.160
age_{it}	4.22	9.39
age_{it}^2	-.063	.140
age_{it}^3	.3051e-3	.698e-3
$I\{age_{it} \geq 65\}$	-.222	.188
Number of obs	725	

The sample of log wage per hour is omitted if it is over 9.

[12] In Japan, after age 60, the earning-related pension is reduced if an individual earns more than the base amount.

Table 9: Fixed Effect Estimator of Male Respondents Aged More Than 60 at 2009 (First stage)

variable	coefficients	S.E.
$year2009_t$	-.036	.1184
age_{it}	-.407	7.07
age_{it}^2	.011	.105
age_{it}^3	-.753e-4	.522e-3
$I\{age_{it} \geq 65\}$	-.076	.137
$I\{M_t = bad\}$	-.229	.110
$I\{age_{it} \geq 60 \ and \ Earning - related \ pension_{it} = 1\}$	-.012	.061

4.2 Initial distribution

We explain how to produce the asset profile. There are two types of answers in the question asking how much asset a respondent has. They are an exact asset price and approximate asset price in each asset component. We use the price, housing, housing loan, loan except housing, saving, stock and bond, and the price of their firms. With respect to an approximate price, we use the medium between the upper value and the lower value in each asset component. A spouse status is available in the dataset. A spouse income is estimated by using the average income given a male respondent's age. We use the data in both 2007 and 2009.

A full pension amount is assumed to be deterministic after age 61. We explain how to produce a full pension amount. First, with respect to basic pension (the *Kiso-Nennkin*), all respondents are assumed to have a full pension amount.[13] Second, with respect to earning-

[13] 792100 yen.

related pension (the *Kosei-Nennkin* or the *Kyosai-Nennkin*), the pension amount is imputed. The method of imputation is in section A. Table 10 is the initial distribution.

Table 10: Basic Statistics of Initial Distribution at 2007 (JSTAR, Only Male Respondents Aged 55-60)

variable	mean	s.d.
age	57.54	1.67
asset(yen)	38102681.21	63153833.23
health(bad=1)	0.12	0.33
spouse status(with spouse = 1)	0.86	0.35
earning-related pension(monthly payment, yen)	144376.07	108949.49
Number of obs	149	

4.3 The estimation of parameters of utility function

Let $K_b = 50000000$, $R = 1.04$. We estimate the parameters of utility function by using the method of simulated moments (Table 11). We use the joint estimation which both adjusts a participation rate and estimate the parameters of the utility function.[14] Table 11 is the estimation result.[15] Under the model and the moment conditions, identification is not verified. We want to compare the result in French (2005). In French (2005), the annual labor hour is used in the model. There is no meaning to compare the result of ϕ_1, ϕ_2 and L. It is interesting that the estimated value of the consumption weight γ is close to the value of French (2005). The specification test is significantly rejected as in the preceding studies.[16] It is interesting that the estimated consumption weight $\hat{\gamma}$ in Japan is close to the estimated value in the U.S.

[14] The details are in section A.

[15] The calculation method is similar to the NFXP (Rust (1987)).

[16] For example, French (2005), French and Jones (2011)

Table 11: Parameter Estimates of Utility Function

Parameter	the estimated value (S.E.)
ϕ_1	12.751(0.0692)
ϕ_2	10.693(0.0237)
γ	0.623(0.0631)
η	3.864(0.005)
θ	0.783(0.051)
β	0.920(0.189)
L	71.854(0.114)
χ^2 Statistic (degrees of freedom)	196(8)

4.4 Simulation

Figure 5 shows both the without-case simulation result and the adjusted participation rate. The gap between those two cases are large around age 65. In the model, age 65 is an important point. A respondent starts to receive the basic pension and there is a large decrease in the wage process. There is a very strong incentive to retire at age 65 in the model. This point is discussed later.

The simulation results are shown in Figure 6. Case 1 is the without-case. Case 2 is the with-case. The with-case is that the eligibility age of receiving basic pension is assumed to be changed to age 70. Case 3 is the second with-case. The second with-case is that the full amount of basic pension is assumed to be cut in half. There is a difference between Case 1 and Case 2 after age 65. The reform of Case 2 increases the participation rate by about 4.7 percent in average between age 63 and age 69 compared to Case 1. The reform of Case 3 increases the participation rate by about 1.4 percent in average between age 63 and age 75 compared to Case 1.

Finally, the reasons why the participation rate decreases suddenly at age 65 in Figure 6 are considered. Both the wage sharp decrease at age 65 and receiving a pension are the most important factors which explain the participation decrease at age 65.

In Figure , three cases are simulated. Case 1 is the same as Case 1 of Figure 6. Case 5 is

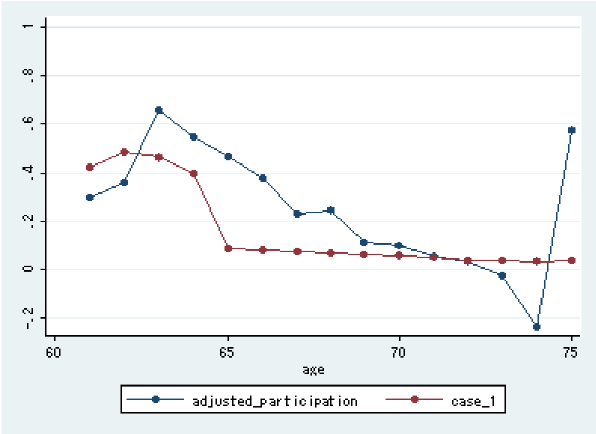


Figure 5: Simulation 1

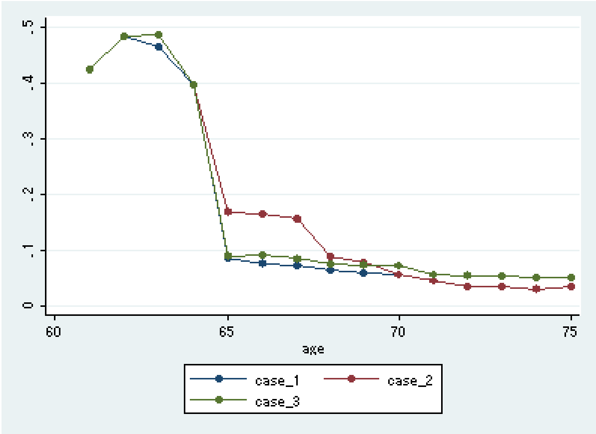


Figure 6: Simulation 2

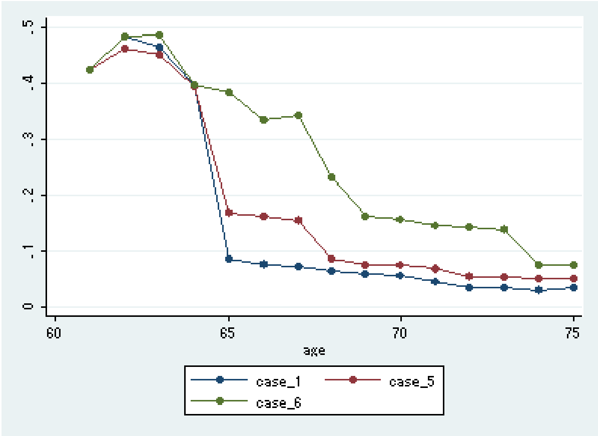


Figure 7: Simulation 3

that α_2 in the wage equation (8) and (9) is assumed to be zero in addition to the condition of Case 1. Case 6 is that the full amount of basic pension is assumed to be zero in addition to the conditions of Case 5. As Figure shows, there are three reasons why the participation rate considerably decreases at age 65. First, a wage significantly decreases at age 65. The wage decrease is one of the reasons why the participation rate considerably decreases at age 65. The second reason is receiving basic pension at age 65. The participation decrease at age 65 is explained by the two factors as Figure shows.

5 Conclusion

We show the exogenous data process, the estimated parameters of utility function and some simulations under some counterfactual social security reforms. We estimated the dynamic model which includes labor supply and asset accumulation decision under a health uncertainty and a mortality uncertainty. We show the some counterfactual simulations. For example, the reform that the eligibility age of basic pension is changed to age 70 influences the labor participation rate under the estimated parameters. However, there are many problems in this study. These problems have been discussed in this study.

The first problem is that there are some problematic assumptions. As discussed in section 3.3, some problematic assumptions are used. These assumptions influence the final specification test in Table 11. Whether the model framework is appropriate or not should also be considered. In this framework, respondents who have retired can participate labor force again. However, in Japan, this situation is unrealistic. With respect to the model assumptions, reconsidering these assumptions can be a future work. For example, the wage uncertainty can be estimated from the wage transition information of individuals if the number of waves becomes large. The second problem is pension information. As mentioned in section 4.2, there is a problem in the way to impute pension information. This point will be also be a future work. The final problem is the method of estimation. In this study, identification of the model parameters is not verified.

Finally, we refer to the method of estimation. Of course, the estimation of a dynamic model should be done based on a multi-period panel dataset. If there is a panel data with enough periods to estimate all parameters, the estimation method which both adjust a cohort effect and estimate parameters of the utility function is not necessary. However, at this moment, this analysis has to be done by using two period panel datasets. This is the reason to use this complex estimation method. Estimating a dynamic model by using a panel dataset with enough periods can be a future work.

The result shows the large gap between the simulation result and the adjusted participation rates. In the model, age 65 is important point. A respondent starts to receive basic pension and there is a sharp decrease in the wage process. These two factors are the cause of a large decrease in labor participation at age 65. However, it is possible that this specification has a problem. In fact, in the dataset, it is not observed that elderly people sharply retire at age 65. The wage equation is a key factor of this model to explain when people retire. Estimating a better wage equation to explain the retirement behavior can be also a future work.

A Appendix

A.1 Validity of using annual hours worked

We compare annual working hours between 2007 and 2009. In 2007, 1989.6 is the average working hours per a male worker in Japan.[17] Next, we explain how to produce annual labor hours in each year. As we mentioned, the way to ask how many days a respondent works

[17]Ministry of Health, Labour and Welfare, Monthly Labour Survey.
The website: http://www.e-stat.go.jp/SG1/estat/List.do?lid=000001038554 (in Japanese), Table 19.

changes. In 2007, we explain how to produce the number of working weeks per year. With respect to the workers working throughout the year, the number of working days is calculated by subtracting the number of national holidays in 2006 plus absent days except holidays from the number of days per year (365). Finally, the number of the annual working days is divided by seven. With respect to the workers who do not work throughout the year, we use the results of the respondent answer. In 2009, with respect to the workers working throughout the year, the number of working days is calculated by subtracting the number of paid holidays and national holidays from seven times working weeks including paid holidays. In 2009, with respect to the workers who do not work throughout the year, the number of working days is calculated by subtracting the number of paid holidays from seven times weeks worked including paid holidays.

In particular, there is a problem in annual working hours in 2007 because the average annual working hours of elderly people is more than the average annual working hours of all Japanese male workers. In 2009, we calculated the annual hours worked by subtracting the number of paid holidays, which means that a worker used all paid holidays. We underestimated the number of annual working days in 2009. Since it seems that there is a problem when calculating annual working hours, we don't use the annual working hours. We set the average annual number of working weeks in 2009 of male respondents who work throughout the year as the annual number of working weeks in the model. The value is 42.95.

annuallabor 07_category	Freq.	Percent	Cum.
500	75	6.17	6.17
1000	64	5.27	11.44
1500	79	6.50	17.94
2000	345	28.40	46.34
2500	359	29.55	75.88
3000	293	24.12	100.00
Total	1,215	100.00	

Figure 8: Distribution of Male Labor Hour in 2007 (JSTAR, Only Respondent)
500=less than 500, 1000=larger than 500 and less than1000, 3000=larger than 2500

annuallabor 09_category	Freq.	Percent	Cum.
500	48	6.95	6.95
1000	48	6.95	13.89
1500	96	13.89	27.79
2000	232	33.57	61.36
2500	163	23.59	84.95
3000	104	15.05	100.00
Total	691	100.00	

Figure 9: Distribution of Male Working Hours in 2009 (JSTAR, Only Respondents)

Table 12: Basic Statistics of Male Working Hours in 2007 and 2009 (JSTAR, Only Respondents)

year	number of obs	mean	s.d.
2007	1209	2090.469	806.7646
2009	690	1881.866	797.9687

min = 0 and max =5000 are set.

A.2 Reduced payment system of earning-related pension

In this section, we will explain the reduced payment system of an earning-related pension. (15) is based on the descriptions of the Japan Pension Service.[18] (15) is one part of the whole reduced payment system. The whole reduced payment system is included.

If $61 \leq t \leq 64$,

[18]The Japan Pension Service
The website: http://www.nenkin.go.jp/new/topics/zaisyoku_23_0318.html (in Japanese)

$$\text{if } (pb2nd_t + W_tH_tweeks) \leq 28 \cdot 12 \cdot 10000$$

$$\widetilde{pb2nd_t} = pb2nd_t \tag{15}$$

$$\text{if } \quad (pb2nd_t + W_tH_tweeks > 28 \cdot 12 \cdot 10000) \wedge (pb2nd_t \leq 28 \cdot 12 \cdot 10000)$$

$$\wedge (W_tH_tweeks \leq 46 \cdot 12 \cdot 10000)$$

$$\begin{cases} \widetilde{pb2nd_t} = \frac{pb2nd_t}{2} - \frac{W_tH_tweeks}{2} + 14 \cdot 12 \cdot 10000 \\ \text{if } \frac{pb2nd_t}{2} - \frac{W_tH_tweeks}{2} + 14 \cdot 12 \cdot 10000 > 0 \\ \widetilde{pb2nd_t} = 0 \\ \text{if } \frac{pb2nd_t}{2} - \frac{W_tH_tweeks}{2} + 14 \cdot 12 \cdot 10000 \leq 0 \end{cases}$$

A.3 The imputation earning-related pension

It is important that an earning-related pension is decided by a salary during working periods as en employee. There are many missing values in the question of pension. We impute the value of pension.

We produce an indicator of income. We use a wage per hour in 2007 and a weekly working hours which is used in the estimation of the wage process. In addition, the number of years during which an individual works as a company employee or a public servant is used. We multiple the working years as a company employee or a public servant and a wage per week. We call this value "income indicator". We make an adjustment factor which is set so that the average "income indicator" of individuals aged less than 60 is equal to the average of monthly *Kousei-Nennkin* in 2007.[19] Next, we produce the value multiplying "income indicator" of individuals aged 55-60 with an adjustment factor. We call this value an imputed pension. Finally, we use both pension data in the JSTAR and the imputed pension.

[19] The Statistics of Older People Society (*Kourei Syakai Toukei Yourann*)(2011). The website: http://www.jeed.or.jp/data/elderly/statistics/statistics01.html#sec11 (in Japanese), Table 11-4. The calculation means that (adjustment factor) × (the average "income indicator") = (the average monthly *Kousei-Nennkin* in 2007)

First if there is an imputed pension, we use the imputed pension. Second, if there exists no imputed pension, we use the pension data in 2007 if there exists the pension data in JSTAR 2007. Then, we subtract the monthly amount of basic pension from the pension data in JSTAR 2007. In addition, if the pension value in the JSATR 2007 is less than 792100, then we set that an earning-related pension is equal to zero. If there is no pension information, we set that an earning-related pension is equal to zero if there exists a response in the question whether a respondent does not have an earning-related pension eligibility. Finally, we show the basic statistics about the initial distribution. The number of samples is small (Table 10).[20] We assume that every individual in the initial distribution follows the estimated wage process although the wage process is estimated by only workers who work throughout the year. Even if this assumption exists, the pension imputation will produce some biases. About 85 percent of individuals in the initial distribution have an earning-related pension which is more than zero.

A.4 The method of estimation

There are other two candidates which should be used as the moment conditions including labor participation rate. The asset moment condition is the first candidate. We must adjust cohort effects in sample moments. We use only asset including all kinds of asset. Then, the asset sample size is very small. As a result, we omit asset moment conditions. Second, the moment conditions of working hours is the second candidate. We use weekly working hours in the model. Then, the weekly working hours means the annual average weekly working hours. However, in the JSTAR, the data of individuals who do not work annually is included. We omit the moment conditions of working hours.

We explain how to adjust the fixed effect of a labor participation rate. When we adjust the fixed effect of labor participation rate, we do the joint estimation which both adjusts the

[20]We do not use male samples with a female respondent because we select a male sample with a wife.

cohort effect and estimates the parameters of the utility function. The method of adjustment in French (2005) has a problem in this study. When we used the estimation method in French (2005), we estimated very large consumption weight. As a result, it is showed that some reforms of social security system in Japan do not so much influence the participation rate of the Japanese male elderly. For example, the reform that the eligibility age of basic pension is changed to age 70 increases the labor participation rate of male elderly by about 0.4 percent in average between age 63 and age 66. In Japan, the estimated consumption weight is very large. This estimated consumption weight largely influences the effect of the reforms of social security system. The gap between the simulated participation rate and the adjusted participation rate was large. This was caused by that the adjustment of the fixed effect and the parameters estimation were carried out separately.

We took another strategy to estimate the parameters of utility function while adjusting the cohort effect of labor participation rate.

Cohort effect adjustment

We explain the equation of labor participation rate. Let t' be the period number(t'=1 (in 2007), 2 (in 2009)). Let $age_{it'}$ be the age of a respondent at period t'. Let $P_{it'}$ be the participation indicator of a respondent i at period t'.

$$P_{it'} = z_i + \sum_{k=60}^{K} \delta_{1k} I\{age_{it'} = k\} \cdot I\{M_{it'} = good\} + \sum_{k=60}^{K} \delta_{2k} I\{age_{it'} = k\} \cdot I\{M_{it'} = bad\} \quad (16)$$
$$+\alpha_1 year2009_{t'} + u_{it'}$$

The equation (17) is derived from (16) given an arbitrary set of parameters, $(\theta_1 = (\{(\delta_{1k}, \delta_{2k})\}_{k=60}^{K}, \alpha_1)$,

θ_2). Let θ_2 be a vector of the parameters of the utility function.

$$
\begin{aligned}
\tilde{y_{it'}} = (P_{it'} - P_{it'-1}) = & \quad (17)\\
(\sum_{k=60}^{K} \delta_{1k} I\{age_{it'} = k\} \cdot I\{M_{it'} = good\} + \sum_{k=60}^{K} \delta_{2k} I\{age_{it'} = k\} \cdot I\{M_{it'} = bad\} \\
+\alpha_1 year2009_{t'}) \\
-(\sum_{k=60}^{K} \delta_{1k} I\{age_{it'-1} = k\} \cdot I\{M_{it'-1} = good\} + \sum_{k=60}^{K} \delta_{2k} I\{age_{it'-1} = k\} \cdot I\{M_{it'-1} = bad\} \\
+\alpha_1 year2009_{t'-1}) \\
+u_{it'} - u_{it'-1} = \tilde{x_{it'}}\delta + \tilde{u_{it'}}
\end{aligned}
$$

The moment conditions (18), (19) are assumed in the equation (17). These moment conditions are used to estimate the parameters ($\theta_1 = (\{(\delta_{1k}, \delta_{2k})\}_{k=60}^{K}, \alpha_1)$, θ_2)

$$E[(u_{it'} - u_{it'-1})(I\{age_{it'} = k\} \cdot I\{M_{it'} = good(bad)\} - I\{age_{it'-1} = k\} \cdot I\{M_{it'-1} = good(bad)\})] = 0 \quad (18)$$

$$E[(u_{it'} - u_{it'-1})(year2009_{t'} - year2009_{t'-1})] = 0 \quad (19)$$

We remember that a set of parameters is given. With respect to this fixed parameter values, we can produce the estimator of the cohort effect. The variables of k_i and $cohort_{2007}$ are defined. If an individual i is k_i years old in the JSTAR 2007, let $cohort_{2007} = k_i$. For convenience, consider that t' is fixed. For example, $t' = 1$. Then, we can produce the estimator of the cohort effect given a set of an arbitrary parameters. Let $n_{k_i 1}$ be the number of respondents aged k_i at period 1.

$$\frac{1}{n_{k_i1}}\Sigma_{l=1}^{n_{k_i1}}(z_l + u_{l1}) =$$

$$\frac{1}{n_{k_i1}}\Sigma_{l=1}^{n_{k_i1}}[P_{l1} - (\sum_{k=60}^{K}\delta_{1k}I\{age_{l1} = k\} \cdot I\{M_{l1} = good\}) - (\sum_{k=60}^{K}\delta_{2k}I\{age_{l1} = k\} \cdot I\{M_{l1} = bad\}$$

$$+\alpha_1 year2009_1)]$$

$$= \hat{E}[z_i|cohort_{2007} = k_i] \quad (20)$$

Then, the variable of participation given a set of parameters is adjusted as follows.

$$\hat{P}_{it'} = P_{it'} + \hat{E}[z_i|cohort_{2007} = 60] - \hat{E}[z_i|cohort_{2007} = k_i] \;\; (t' = 1, 2) \tag{21}$$

$\hat{P}_{it'}$ is an adjusted participation indicator. Let $\hat{P}_{it'} = \hat{P}^t_{it'}$ if $age_{it'} = t$. The moment conditions (23) are assumed. These moment conditions are also used to estimate the parameters ($\theta_1 = (\{(\delta_{1k}, \delta_{2k})\}_{k=60}^{K}, \alpha_1)$, θ_2). Let $\tilde{P}_t$ be a simulated moment at age t. t is from age 61 to age 75.

$$E[\hat{P}^t_{it'} - \tilde{P}_t|t] = 0 \;\; (t = 61, 62, ..., 75) \tag{22}$$

Finally, we produce a distance vector of the equation (23). n is the number of separate individuals in this dataset. n_t is the number of individuals who are observed to participate or not participate at age t. It is assumed that the participation data does not contribute to moment conditions if it is missing.

$$\frac{1}{n}\sum_{i=1}^{n_t}\{\hat{P}^t_{it'} - \tilde{P}_t\} \tag{23}$$

GMM objective function

We correct the MSM (Pakes and Pollard (1989), Duffie and Singleton (1993)) used in French (2005) to use our estimation method. There are two corrections in this study. The objective function and the estimator of the asymptotic variance. The GMM criterion function is the equation (24). Let $\hat{W}$ be an optimal weighting matrix. q is the ratio of the number of observations to the number of simulated observations. Let $\tilde{d}_1(\theta_1)$ be the distance vector of the moment conditions of a participation equation (the moment conditions (18), (19)). Let $\tilde{d}_2(\theta_2 : \chi)$ be the distance vector between an actual participation and a simulated participation moment.

$$\min_{\theta \in \Theta} n(\tilde{d}_1(\theta_1)' \tilde{d}_2(\theta_2 : \chi)')\hat{W} \begin{pmatrix} \tilde{d}_1(\theta_1) \\ \tilde{d}_2(\theta_2 : \chi) \end{pmatrix} \tag{24}$$

$$\hat{W}^{-1} = \begin{pmatrix} \hat{\Sigma_{11}} & \hat{\Sigma_{12}} \\ \hat{\Sigma_{21}} & (1+q)\hat{\Sigma_{22}} \end{pmatrix} \tag{25}$$

$\hat{\Sigma_{11}}, \hat{\Sigma_{12}}, \hat{\Sigma_{21}}, \hat{\Sigma_{22}}$ have typical elements of each optimal weighting matrix which is associated with $\tilde{d}_1(\theta_1)$ and $\tilde{d}_2(\theta_2 : \chi)$.

Let $\theta_0 = (\theta_1^0, \theta_2^0)$ be the true parameter. Under some conditions which include the regularity conditions in Pakes and Pollard (1989) and Duffie and Singleton (1993), the GMM estimator $\hat{\theta} = (\hat{\theta_1}, \hat{\theta_2})$ is consistent and the estimator $\hat{\theta}$ is asymptotically normally distributed. We correct the estimator of the asymptotic variance.

$$\sqrt{n} \begin{pmatrix} \hat{\theta_1} - \theta_1^0 \\ \hat{\theta_2} - \theta_2^0 \end{pmatrix} \to_d N(0, \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix})$$

Then, the elements of the asymptotic variance is estimated by $\begin{pmatrix} \hat{A_{11}} & \hat{A_{12}} \\ \hat{A_{21}} & \hat{A_{22}} \end{pmatrix} = (\hat{G}'\hat{W}\hat{G})^{-1}$

and $\hat{G} = \begin{pmatrix} \frac{\partial \tilde{d_1}}{\partial \theta}\big|_{\theta=\hat{\theta}} \\ \frac{\partial \tilde{d_2}}{\partial \theta}\big|_{\theta=\hat{\theta}} \end{pmatrix}$.

References

Aaronson, D. and French, E. (2004). The Effect of Part-Time Work on Wages: Evidence From the Social Security Rules. *Journal of Labor Economics*, 22(2):329–352.

Duffie, D. and Singleton, K. J. (1993). Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica*, 61(4):929–952.

French, E. (2005). The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior. *The Review of Economic Studies*, 72:395–427.

French, E. and Jones, J. B. (2011). The Effects of Health Insurance and Self-Insurance on Retirement Behavior. *Econometrica*, 79(3):693–732.

Hayashi, F. (2000). Econometrics. *Princeton university Press.*

Higuchi, Y. and Yamamoto, I. (2002). Older People Labor Supply in Japan (Wagakoku Dannseikoureisya no Rodo Kyokyu Kodo Mekanizumu)(Japanese Article). *Research of Finance (Kinyu Kenkyu)*, 10:31–78.

Iwamoto, Y. (2000). Reduced-payment of Earning-related Pension and Older People Labor Supply (Zaisyoku Rourei Nennkin Seido to Koureisya no Syugyo Kodo)(Japanese Article). *Quarterly of Social Security (Kikan Syakai Hosyo Kenkyu)*, 35:364–376.

Karuma, F. and Ikeda, S. (2006). The Estimation of Time Preference Rate by Experiments and Questionnaires (Keizai Zikkken to Anketo Cyosa ni Motozuku Zikan Waribikiritu no Kenkyu)(Japanese Article). *PRI Discussion Paper Series*, No.06A-26.

Oshio, T. and Oishi, A. (2000). Retirement Behavior of Older People and Social Security Asset (Koureisya no Intai Kodo to Syakaihosyoshisann)(Japanese Article). *The Quarterly of Social Security Research (Kikan Syakai Hosyo Kenkyu)*, 35:405–419.

Oshio, T. and Sato Oishi, A. (2003). Consequence of Japanese Budget After a Pension Reform (Nenkin Kaikaku no Zaiseiteki Kiketu)(Japanese Article). *The Quarterly of Social Security Research (Kikan Syakai Hosyo Kenkyu)*, 39:216–233.

Oshio, T. and Sato Oishi, A. (2004). Social Security and Retirement in Japan: An Evaluation

Using Micro-Data. In Gruber, J. and Wise, D. A., editors, *Social Security Programs and Retirement around the World: Micro-Estimation*. University of Chicago Press.

Oshio, T., Sato Oishi, A., and Shimizutani, S. (2011). Social Security Reforms And Labour Force Participation of The Elderly in Japan. *The Japanese Economic Review*, 62(2):248–271.

Pakes, A. and Pollard, D. (1989). Simulation and the Asymptotics of Optimization Estimators. *Econometirca*, 57(5):1027–1057.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometirca*, 55(5):999–1033.

Rust, J. and Phelan, C. (1997). How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets. *Econometirca*, 65(4):781–831.

Chapter 5: What Explains the Difference in the Effect of Retirement on Health?: Evidence from Global Aging Data

1 Introduction

Retirement related policies, such as pension system reform, have become important for developed countries to sustain their social security systems. Numerous developed countries have faced the same problems of a decreasing birthrate and an ageing population. As population ages, the cost of social security and social welfare increases, eroding the country's budget. As such, developed countries have reformed their pension systems to reduce the cost of social security and social welfare. Moreover, many developed countries, such as the United States, the United Kingdom, and Korea have already decided to increase pension eligibility age for the next decades. Japan has already increased the pension eligibility age. These pension reforms in developed countries are expected to delay retirement. As Gruber and Wise (1998) discuss, the relationship between the social security system and retirement in developed countries generated a lot of attention in economics. When policy makers evaluate the effect of these reforms, health is a key factor. If working is beneficial for the health of the elderly, it would lead to reduced medical expenses and vice-versa.

Along with a growing interest in the effect of these retirement delaying policies, a number of studies have investigated the relation between retirement and health over the last two decades. [1)] Using various health indexes, numerous researchers have examined the relationship between health and retirement. To the best of our knowledge, Kerkhofs and Lindeboom (1997) is one of the first papers suggesting endogenous decisions between retirement and health, and identifying the effect of retirement on health. They find that the Hopkins Symptom Checklist (HSCL) health index can be improved after early retirement in the Netherlands by applying FE methods. Lindeboom et al. (2002) extend Kerkhofs and Lindeboom (1997) study to other indices such as the mini-mental state examination (MMSE) test on cognitive ability, the Center for Epidemiological Studies-Depression (CES-D) test of depressing feelings, and others, and apply FE methods to Dutch data different from that of Kerkhofs and Lindeboom (1997).[2)] Charles (2004) is also one of the first investigations that analyze the causal effect of retirement on health focusing on subjective well-being (SWB) in the economic literature by using instrumental variables (IVs).

Additionally, there are numerous other papers that study the effect of retirement on various health indexes (Bound and Waidmann, 2007; Coe and Lindeboom, 2008; Dave et al., 2008; Neuman, 2008; Johnston and Lee, 2009; Latif, 2011; Coe and Zamarro, 2011; Kajitani, 2011; Behncke, 2012; Bonsang et al., 2012; Mazzonna and Peracchi, 2012; Hernaes et al., 2013; Bingley and Martinello, 2013; Hashimoto, 2013; Insler, 2014; Kajitani et al., 2014; Hashimoto, 2015; Kajitani et al., 2016). There are, however, no unified views on the impact of retirement on various health indexes. While some studies conclude that retirement has a positive impact on health defined as mental or physical health, other studies conclude that retirement has no or negative effect. Additionally, these results depend on characteristics such as gender and education.

The goal of this paper is to explain why the estimated effect of retirement on health in the previous studies differ. One of the keys to understanding these differences is a better understanding of the path through which retirement influences health. If there is an important link between retirement and health (i.e., a mechanism through which retirement influences health outcomes), the effect of retirement on health could be heterogeneous. In fact, some researchers focus on the change in the health investment behaviors after retirement to explain why the estimated effect of retirement on health in the previous studies differ (Zhao et al., 2013; Ayyagari, 2014; Insler, 2014; Eibich, 2015; Motegi et al., 2016). Eibich (2015) is the first study to clearly point out the importance of the mechanism to explain the difference in the effect of retirement on health. On the other hand, we investigate the differences by focusing on the analysis methods. There is no study to focus on the analysis methods to explain why the effect of retirement on health estimated results in the previous studies differ.[3)] The contribution of this paper is to provide two verification frameworks to examine which factor causes these differences. We will discuss which factor causes the difference in the estimated results by the previous studies.

According to our analysis, the analysis method is one of the determinants of these differences. By choosing an analysis methodology, we also comprehensively reexamine the effect of retirement on health in eight countries. We analyze five health indexes, such as self-reported health, depression, cognitive function, body mass index (BMI), and activities of daily living (ADL). We examine the five health indexes by using the same analysis method. By doing so, we show the comprehensive results of the effect of retirement on health.

The rest of this paper is organized as follows: Section 2 reviews preceding studies; Section 3 discusses the data; Section 4 examines why the estimated results of the effect of retirement on health in previous studies differ from each other; Section 5 performs harmonized analysis on the effect of retirement on health; and Section 6 concludes this paper and discusses future research scope.

2 Literature Review

This section summarizes related studies, focusing on economic literature. As such, we introduce studies that examine the effect of retirement on health. The study by Kerkhofs and Lindeboom (1997) is one of the first to suggest an endogenous decision linking retirement and health regarding the effects of retirement on health. Using a fixed effects (FE) method, they find that, in the Netherlands, the HSCL health index can be improved after early retirement. Lindeboom et al. (2002) examined other measurement scales, such as MMSE and CES-D, with FE methods, using Dutch data. Charles (2004) also conducted an early investigation analyzing the causal effects of retirement on health by focusing on SWB and through IV. Psychological and psychiatric literature boasts a large body of research on the correlation of retirement and SWB, but has paid scant attention to causal effects. [4)]

Furthermore, Rohwedder and Willis (2010), who investigated the effects of retirement on cognitive abilities and compared micro data across the USA, the UK, and 11 European countries, found a negative influence of retirement on cognitive abilities. They suggest that institutional differences across countries, such as pensions, taxes, and disability policies, are also important in explaining the differences in health outcomes across countries. As such, Rohwedder and Willis (2010) gave an impetus to research on the effect of retirement on cognitive abilities, making possible studies such as those by Bonsang et al. (2012), Mazzonna and Peracchi (2012), Coe et al. (2012), and Bingley and Martinello (2013). Additionally, numerous other studies assessed the effects of retirement on other aspects of health. [5)] Finally, Tables A.2.1, A.2.2, A.2.3, and A.2.4 in section A.2 show a summary of relevant studies, chosen based on the following criteria:

- We chose all papers analyzing the effect of retirement on health that have been published by November 2015. We used Google Scholar to identify these research papers.
- We choose all working papers that have more than 50 citations on Google Scholar by November 2015.

We restrict our analysis to only papers in economics, thus excluding literature on public health. In Tables A.2.1, A.2.2, A.2.3, and A.2.4, we show the category of health outcome, method, the definition of retirement, control variables information, dataset, the method of sample selection, and the surveyed country. Here, "positive" means the positive impact on a health status (better after retirement), "negative" means a negative impact worse after retirement, and "no" means no impact. According to Tables A.2.1, A.2.2, A.2.3, and A.2.4, there is no unifying result in all health indexes except the health index, which only a few studies analyze. Numerous studies analyze CES-D, self-report of health, ADL, and cognitive functioning. According to Tables A.2.1, A.2.2, A.2.3, and A.2.4, the datasets such as the HRS, the SHARE and the ELSA have been frequently used. The fixed effects method or the IV method have been typically used as the analysis method. According to Tables A.2.1, A.2.2, A.2.3, and A.2.4, many studies suggest that cognitive function decreases after retirement. In addition, many studies suggest that self-report of health improves after retirement. However, there is no agreement in other health indexes.

We consider why they obtain different results. We also add BMI to the analyzed indexes, although only two studies in our list use it. This is because we comprehensively analyze the effect of retirement on health indexes. In the Appendix A.2, we show the other indexes on illness. However, this paper does not focus on the health indexes of illness.

3 Data

This paper uses the Health and Retirement Study (HRS) [6)] and other related datasets, such as the English Longitudinal Study of Ageing (ELSA), the Health Survey for England (HSE), the Survey of Health, Ageing, and Retirement in Europe (SHARE), and the Japanese Study of Ageing and Retirement (JSTAR). These are panel surveys of individuals 50 or older. These family datasets are constructed so that the questions in the HRS family studies are as similar to the original questions in the HRS as possible. They include a rich variety of variables to capture living aspects in terms of economic status, health status, family background, as well as social and work status. We subsequently explain all health indexes used.

Cognitive score: We use the cognitive function score in the HRS and other related datasets. In the HRS, we use the immediate word recall scores (first half of the word recall test), delayed word recall (second half of the word recall test), [7)] and word recall summary score (immediate word recall plus delayed word recall). The word recall summary score is between 0 and 20. The immediate word recall and delayed word recall tests ask the respondent to recall as many words as possible from a list of 10 words. The score of immediate word recall and delayed word recall is the number of words from the 10-word list that were recalled correctly.

Self-report of health: In the HRS, there is a variable that indicates self-reported health conditions. The variable measures the categories of health self-reports as excellent, very good, good, fair, poor. The health categories are numbered from 1 (excellent) to 5 (poor). In all related datasets, the same variable is present. We convert the five values into two health statuses, poor health or not poor health. Additionally, in the ELSA and the SHARE, we can use another scale of self-assessed health: very good, good, fair, bad, and very bad. We also define the health self-report index of "bad health." [8)]

ADL: This variable measures the change in the index for ADL. In the HRS and other related datasets, all respondents are asked to answer questions such as "Because of a health or memory problem do you have any difficulty with bathing or showering?" We use this information when calculating the ADL score.

Depression: In the HRS, there is a question targeting whether a respondent has symptoms of depression. For example, one of the statements is “Much of the time during the past week, you felt depressed.” We use these questions when we calculate the CES-D score. In the HRS and other related datasets, there are similar questions. Additionally, we use another depression scale, EURO-D, which is available in all version of the SHARE. We mainly use the EURO-D scale in the SHARE because the CES-D scale is only available in waves 1 and 2 of the SHARE.

BMI: In the HRS and other related datasets, all respondents are asked to provide their weight and height, and BMI is calculated using this information. We use the value of BMI and create a dummy variable that takes the value 1 if the respondent's BMI value is greater than or equal to 30.

We summarize all scores and values of these health indexes in Tables 1 and 2. In Table 1, we show the descriptive statistics of the age group above 50 in all countries and the descriptive statistics for the USA in Table 2. According to Table 1, the scores and values are not at the same level in all countries, BMI in the US being higher than in other countries. In Table 2, we can observe characteristics of the cognitive function. Females have a higher score than males in the word recall summary score. Highly educated individuals have higher overall cognitive scores.

In Section 5, we perform a dynamic analysis for selected countries. We utilize both the pension eligibility age and the long-term variation of retirement behavior. Moreover, we choose the analyzed countries based on the availability of information regarding pension eligibility age. We mainly use the harmonized datasets. 9) However, when our preferred variables are not available in the harmonized datasets, we use the variables of the original datasets. In Table 3, we show a summary explaining which dataset we use in Section 5 of this paper.

More importantly, we use the pensionable age when we calculate our IVs. We explain this point in Appendix (A.1), while in section 5, we use only the pensionable age confirmed to be correct.

4 Critical Literature Assessment

4.1 Targeted Literature

Our goal is to explain why the estimated results of the effect of retirement on health in previous studies differ. We investigate the difference by focusing on the research framework. First, we create pairs of related studies for each health index, based on the following criteria:

- Step 1: We choose papers from Tables A.2.1, A.2.2, A.2.3, and A.2.4.
- Step 2: We can replicate them by using the HRS, related studies (the Global Aging Data), and the HSE.
- Step 3: We choose only published papers in Health Economics or Labor Economics.
- Step 4: We choose only published papers that estimate a linear model to analyze the effect of retirement on health.
- Step 5: We choose published papers in journals with higher impact factor as much as possible.

Based on these criteria, we choose the studies in Table 4, which we use in the next sections. We show how these criteria determine which paper we analyze in Appendix (A.3). In the subsequent section, we explain how we analyze why the effect of retirement on health differs.

4.2 Verification Framework 1

Having chosen the targeted studies, we first analyze the effect of the difference in each factor on the final results. Each study consists of certain factors, such as surveyed country, analysis method, retirement definition, etc. (see Table 4). These studies use various identification strategies, analysis methods, and definitions of retirement. As such, we analyze why the estimated results of the effect of retirement on heath in previous studies differ by focusing on the differences in these factors. In each pair of studies, we first replace only one factor (e.g., the estimation method), as shown in Figure 1. In section 4.3, we replace all the factors, one by one in the paired studies. In section 4.2, by replacing only one factor, we analyze the effect of each factor on the difference in the estimated results. There are five characteristics in each study: “index,” “def. of retire,” “controls,” “method,” “sample,” and “survey country.” The differences in these characteristics explain the different results on the effect of retirement on health. The details of these characteristics are as follows.

- Index: characteristics of the index used (e.g., CES-D versus EURO-D);
- Def. of retire: definition of retirement (e.g., retired for at least one year versus not working for pay);
- Controls: What the researchers include as control variables (e.g., only family structure variables versus family structure variables + economic variables);
- Method: analysis method (e.g., FE methods versus IV methods);
- Sample: sample selection method (e.g. only male versus full sample);
- Survey country: surveyed country (e.g., the USA versus France).

Here, we summarize our results.

- The results are not sensitive to replacing the definition of retirement.
- The results are sensitive to replacing the analysis method. In almost all indexes, the estimated results change when replacing the analysis method.
- The results are also sensitive to replacing the surveyed country.
- The difference in the estimated results cannot be explained by only one-factor replacement.

In this section, by replacing only one factor, we have checked the sensitivity of each factor on the estimated results. We explain the details of this procedure by using an example in the results section (**Cognitive score**). We show detailed results, except for the “**Cognitive score**” of the verification framework 1 in section A.4. According to our results, it is difficult to explain why the estimated results are different by replacing only one factor. In the next section, we provide another framework to explain why the estimated results in the previous studies differ.

In the Appendix (A.5), we summarize the replication and replacement notes in this section. When we replicate and replace the analysis of related literature, we make some adjustments if needed (see section A.5 for details).

Cognitive score (Bonsang et al. (2012) versus Coe and Zamarro (2011)):

- Table 5 shows the result replacing each factor from Bonsang et al. (2012) to those in Coe and Zamarro (2011) and from Coe and Zamarro (2011) to those in Bonsang et al. (2012). The upper panel implies how the estimated results will change if we replace either the definition (Def. of retire), the set of control variables (Controls), the analysis method (Method), sample selection method (Sample), or the surveyed country (Survey country) in Bonsang et al. (2012) with the one in Coe and Zamarro (2011). When we replace all factors at the same time, the result in Bonsang et al. (2012) (-1.036) is replaced with the one in Coe and Zamarro (2011) (-0.120). In the method replacement from Bonsang et al. (2012) to Coe and Zamarro (2011), we only replace FE-IV (the method in Bonsang et al. (2012)) with IV (the method in Coe and Zamarro (2011)). In all health indexes, we perform the same analysis.

- According to Table 5, when transplanting one factor from Bonsang et al. (2012) to Coe and Zamarro (2011), the replacement of the surveyed country yields the opposite results (negative-positive) and vice-versa. However, the sensitivity of replacing the control variables and the surveyed country are important.

4.3 Verification Framework 2

In the previous section, we have discussed the sensitivity of each factor on the estimated results. We have also found that there are multiple factors that explain why the estimated results are different. In this section, we propose another framework to explain why the estimated results are different. As such, we start from one study and arrive at another study, replacing factors one by one (see Figure 2). If the source of the difference in the effect of retirement on health exists, the result will change after we change this source as per Figure 2. We discuss the results in the following. As in verification framework 1, we explain the details of this procedure by using an example in the results section (**Cognitive score**).

Cognitive score (Bonsang et al. (2012) versus Coe and Zamarro (2011)):

- Table 6 shows the result for replacing factors one by one from Bonsang et al. (2012) to those in Coe and Zamarro (2011) and from Coe and Zamarro (2011) to those in Bonsang et al. (2012). For example, in Pattern A of Table 6, we first replace "Method," "Controls',' and "Country" from Bonsang et al. (2012) to those in Coe and Zamarro (2011). In the second replacement, we further replace "Def. of Retirement." Finally, we replace "Sample." We perform the same analysis in all health indexes.

- In Table 6, we combine method, controls, and country, as these are the factors producing the change in the results in Review 1. We consider that these factors are important for explaining the difference in the effect of retirement on health between two different studies. The figure on the left shows the change in the estimation results when we change the order of replacing the block (method + controls + country). On the other hand, the right-hand figure shows the change in the estimation results when we change the order of replacing the retirement definition. We compare these cases as follows.

- In all patterns (A, B, C), we observed that the estimated results change after replacing the block (method + controls + country) (Negative $\rightarrow$ No)(left-hand figure). On the other hand, we do not observe any change just after replacing the definition of retirement (right-hand figure).

Self-report of health (Dave et al. (2008) versus Coe and Zamarro (2011)):

- In Table 7, we show the same procedure as in Table 6. The left-hand figure shows the change in the estimation results when we change the order of replacing the block (method + controls + country + index), as these factors (method + controls + country + index) produce the change in the results in Review 1. On the other hand, the right-hand figure shows the change in the estimation results when we change the order of replacing the retirement definition. We compare these cases as follows.

- In all patterns (A, B, C), we observed that the estimated results change after replacing the block (method + controls + country + index) (Negative → Positive)(left-hand figure). On the other hand, we do not observe any change just after replacing the definition of retirement except in pattern B (right-hand figure).

ADL (Dave et al. (2008) versus Neuman (2008)):

- In Table 8, we show the same procedure as in Table 6. The left-hand figure shows the change in the estimation results when we change the order of replacing the block (method + controls), as these factors (method + controls) produce the change in the results in Review 1. On the other hand, the right-hand figure shows the change in the estimation results when we change the order of replacing the retirement definition. We compare these cases as follows.

- In all patterns, changing both the estimation method and the difference in what the researcher uses as control variables produce a change in the results. In particular, in pattern C (left-hand figure), the change in method + controls produces the opposite impact for female samples. In patterns A and B, "sample" is also significant. The estimated results changes just after replacing "sample" (No → No (male) and Positive (female))(left-hand figure). As such, the definition of retirement seems to have no impact on the results (right-hand figure).

Depression (Dave et al. (2008) versus Coe and Zamarro (2011)):

- In Table 9, we show the same procedure as in Table 6. The left-hand figure in Figure 9 shows the change in the estimation results when we change the order of replacing the block (method + controls), as these factors (method + controls) produce the change in the results in Review 1. On the other hand, the right-hand figure shows the change in the estimation results when we change the order of replacing the retirement definition. We compare these cases as follows.

- In all patterns (A, B, C), we observe that the estimated results change after replacing the block (method + controls) (Negative → No). In pattern D, "country + index" is also significant. The estimated results changes just after replacing "country + index" (Negative → No)(left-hand figure). On the other hand, we do not observe any change just after replacing the retirement definition (right-hand figure).

BMI (Johnston and Lee (2009) versus Godard (2016)):

- In Table 10, we show the same procedure as in Table 6. The left-hand figure in Table 10 shows the change in the estimation results when we change the order of replacing the block (method + controls + sample), as these factors (method + sample) produce the change in the results in Review 1. On the other hand, the right-hand figure shows the change in the estimation results when we change the order of replacing the index. There is no difference in the definition of retirement between Johnston and Lee (2009) and Godard (2016). Here, we replace the index, and compare these cases as follows.

- In all patterns (A, B), we observe that the estimated results change after replacing the block (method + controls + sample) (Negative → No). In patterns C and D, "country" is also significant. The estimated results changes just after replacing "country" (Negative → No)(left-hand figure). On the other hand, we do not observe any change just after replacing the index except for pattern A (right-hand figure).

Finally, we summarize our results.

- The choice of the estimation method seems to be the key factor for explaining the difference in the estimation results in all indexes. Additionally, the use of control variables is also important. What the researcher uses as control variables is also included in all health indexes. In all health indexes, the estimation method plus other factors (e.g., method + controls) changes in the estimation result.

- The influence of the difference in the surveyed country is also important for explaining the difference in the effect of retirement on health.

- Changes in the definition of retirement have a lower impact.

According to our results, the difference in the estimation method is a key factor in explaining why the estimated effects of retirement on health in preceding studies differ. It is intuitive that the sensitivity of the surveyed country chosen is strong. However, we do not consider this as problematic. On the other hand, a strong sensitivity of the analysis method choice is problematic because it is possible that we do not appropriately estimate the effect of retirement on health, depending on the choice of the analysis method. In some studies, it is possible that there remains room for further improvement. For example, Coe and Zamarro (2011) estimate the effect of retirement on cognitive function by using cross-sectional data. They use the exogenous variation of the pensionable age as an IV, the SHARE being their data source. As such, we can use a dynamic variation of the retirement behavior in the SHARE. Dave et al. (2008) only use FE and do not use an IV. Consequently, we can use the FE-IV method, often used in recent studies to estimate the effect of retirement on health indexes. For example, Bonsang et al. (2012), Insler (2014) and Godard (2016) use the FE-IV method to estimate the effect of retirement on health.

5 Harmonized Analysis of the Effect of Retirement on Health

5.1 Analysis Framework

Here, we use the FE-IV estimation method and estimate the impact of retirement on certain health indexes for eight countries. Coe and Zamarro (2011) estimate the effect of retirement on cognitive function by using cross-sectional data, and use the cross-country variation of pensionable age to control for retirement endogeneity, using SHARE. However, we use a dynamic variation of the retirement behavior, and control for retirement endogeneity by using the pensionable age in the surveyed countries. We also estimate the effect of retirement on health indexes for each country. While Dave et al. (2008) only use FE, we use the FE-IV method to estimate the effect of retirement on health indexes as follows:[10)]

$$health\_index_{it} = \beta_0 + \beta_1 retire_{it} + \gamma' x_{it} + a_{1i} + \lambda_{1t} + \epsilon_{1it} \quad (1)$$

$$retire_{it} = \alpha_0 + \alpha_1 1\{age_{it} \geq A_i^{eb}\} + \alpha_2 1\{age_{it} \geq A_i^{fb}\}$$
$$+\alpha_1 1\{age_{it} \geq A_i^{eb}\} \cdot age_{it} + \alpha_2 1\{age_{it} \geq A_i^{fb}\} \cdot age_{it} + \eta' x_{it} + a_{2i} + \lambda_{2t} + \epsilon_{2i} \quad (2)$$

A_i^{eb}: the early retirement benefit eligibility age

A_i^{fb}: the full retirement benefit eligibility age where $retire_{it}$ is an indicator which is equal to 1 when a respondent retires at period t. We use two retirement definitions. The first is "not work for pay," which means that a respondent is retired if he/she is not working for payment. The second definition is "complete retire," which is the same retirement definition of Dave et al. (2008). λ_{1t} and λ_{2t} are time FE; a_{1i} and a_{2i} are individual FE; x_{it} are control variables at period t. We restrict the sample to those aged above 50.

Our identification strategy utilizes the fact that the proportion of retired elderly in many developed countries starts to increase dramatically after the pensionable age. Pension eligibility age is exogenous. The incentive to retire from the labor market increases after the exogenous pensionable age. However, the pension eligibility threshold does not directly influence health status, but while it increases the incentive to retire form the labor market. We exploit this fact to identify the effect of retirement on health. [11)] As such, we use dummy variables (e.g., $\{age_{it} \geq A_i^{eb}\}$) and the cross terms between the dummy variable and age (e.g., $\{age_{it} \geq A_i^{eb}\} \cdot age_{it}$) to identify changes in retirement after the pensionable age. We analyze only countries where pensionable age is confirmed to be correct (the USA, the UK, Germany, France, Denmark, Switzerland, Czech, Estonia, Japan, China, and Korea), and discuss how to confirm each pensionable age in Appendix (A.1). For the countries included in the SHARE, we analyze only the surveyed countries included in the first wave (Austria, Belgium, Switzerland, Germany, Denmark, Spain, France, Greece, Italy, Netherlands, Sweden, and Israel) because we utilize enough dynamic variation of retirement and health indexes. We use the UK, Denmark, France, Germany, Switzerland as the European analyzed countries.

Figure 3 shows the proportions of retired elderly by age by pooling all samples. In Figure 3, the pensionable age is represented by the red line. In the US, the UK, Denmark, France, Germany, Switzerland, Japan (male), and South Korea, there is a sharp increase in the proportion of retired around the pensionable age. In the US, Denmark, France, Germany, and Switzerland, around the early retirement age, there is also a sharp increase in the proportion of retired elderly. In the UK, Japan (male) and South Korea, there is a sharp increase in the proportion of retired elderly around the normal retirement age. Additionally, after the early retirement age, the slope of the proportion of the retired elderly changes in many countries. As a result, we use the cross term (e.g., $\{age_{it} \geq A_i^{eb}\} \cdot age_{it}$) to identify this movement. In the next section, the first stage results are presented as to check the validity of our strategy. Eibich (2015) uses a similar strategy to exploit the discontinuity of retirement status around the pensionable age. Furthermore, we control individual demographics (x_{it}), including variables to control the age effect. Around the pensionable age, it is possible that there is a change in individual demographics. As such, we control for these effects.

5.2 The Results

We discuss the estimated results only when the coefficients of IV in the first stage are significant. We also test the endogeneity of retirement with the Durbin-Wu-Hausman test. When we do not reject the null hypothesis, we support the results of FE model. We use the retirement definition of “not work for pay” in all countries except Korea and Japan. On the other hand, in Korea and Japan, we use the retirement definition of “complete retire.” This is because, in Korea and Japan, we do not obtain a significant result in the first stage regression by using the retirement definition of not work for pay. We perform a robustness check with respect to the retirement definition in the next section. With respect to Depression, we use both CES-D and Euro-D, and identify which scale we use in the analysis (e.g., US (CES-D(0-8), Denmark (Euro-D)). The total score of CES-D is seven or eight. On the other hand, the total score of Euro-D is 20. We use Euro-D in the European countries because the sample size is larger when we use Euro-D. The first stage results are shown in Table 11.[12)]

- As per Table 12, in each health index, only Korea has an opposite effect compared to the US. [13)] With respect to self-reported health and CES-D, in half of the surveyed countries, we observe a positive effect of retirement on health. However, only in Korea and the US there is a significant effect on cognitive function. Nonetheless, there is an opposite effect (positive or negative) between these countries.

- As per Table 13, there is a negative effect or no effect of retirement on BMI (BMI: negative = increase and positive = decrease). However, in half of the surveyed countries, there is a positive effect of retirement on ADL.

- Summarizing the estimated results (Table 12 and 13), in the US, we observe a change in health outcomes after retirement for almost all health outcomes. BMI increases after retirement in the US. With respect to poor health, CES-D and ADL summaries, health outcomes improve after retirement, as do in the UK. On the other hand, in Denmark, France, Germany, and Japan, almost all health outcomes remain constant after retirement. In Switzerland, no health outcome changes after retirement.

- Summarizing the results by gender (Table 12 and 13), with respect to poor health, CES-D and ADL summaries, in the US and UK, the coefficients are similar for both elderly males and females. In these countries, health outcomes improve after retirement for both elderly males and females. Regarding the CES-D summary, the magnitude of the coefficient is large (-2.435) for elderly Japanese males, and their CES-D summary improves after retirement. Additionally, BMI increases after retirement, and the magnitude of the coefficient is large (2.796) in Japan.

- In subjective indexes, such as the self-report of health and depression, the index improves in many countries, while the health self-report index worsens only in Korea. With respect to objective indexes, such as BMI and ADL, BMI increases or remains constant and ADL improves or remains constant.

Subsequently, we check the sensitivity of the retirement definition and the pattern of control variables on the effect of retirement on health. We prepare two retirement definitions ("not work for pay" and "complete retire") and four control patterns ("Pattern 1," etc.). According to Table 14, in most analyzed countries and patterns, the estimates are robust, although we change the retirement definition and control variable patterns for each country regardless of health outcomes. The results are sensitive depending on the definition of retirement in Denmark (Pattern 4, ADL) and Germany (Pattern 4, health self-report). The results are not significant for some countries, but there is no opposite effect. In Table 14, we show only the final results after performing the DWH test, by choosing FE or FE-IV.

5.3 Discussion

We summarize our main results in Table 15. Our analysis method (FE-IV) is established in this section. According to Table 15, when we fix our analysis method, we find a few of opposite results (positive or negative effects) (health self-report, cognition). For each health index, we obtain positive (negative) or no effects of retirement on health in all surveyed countries. The important point is that there is heterogeneity of the effect of retirement on health, even if we fix our method and control for retirement endogeneity. Heterogeneities depending on the surveyed countries cannot be explained by the differences in the analysis method. It is possible that these differences can be explained by the heterogeneity of the health investment behavior change after retirement. Consequently, we should investigate the relationship between the heterogeneity of the effect of retirement on health observed in many countries and the heterogeneity of the change in health investment behaviors after retirement. Eibich (2015) discusses this point solely for Germany.

6 Conclusion

We summarize the results of this study as follows.

- Review 1:
  - The results are not sensitive to replacing the definition of retirement.
  - The results are sensitive to replacing the analysis method. In almost all indexes, the estimated results change when replacing the analysis method.
  - The results are also sensitive to replacing the surveyed country.
- Review 2:
  - The choice of the estimation method seems to be the key factor for explaining the difference in the estimation results in all indexes. Additionally, what the researcher uses as control variables is also important. In all health indexes, the estimation method plus other factors (e.g., method + controls) changes the estimation result. What the researcher uses as control variables is also included in all health indexes.

- The influence of the difference in the surveyed country is also important for explaining the difference in the effect of retirement on health.
- Changes in the definition of retirement have a lower impact.

We summarize our main results in Table 15, and fix our analysis method (FE-IV) in Section 5. According to Table 15, when we fix our analysis method, we obtain comparatively stable results. However, there is heterogeneity of the effect of retirement on health even if we fix our methods and control for the endogeneity of retirement. As such, future work could answer on why is there heterogeneity of the effect of retirement on health among different countries. It is possible that the change in health investment behaviors after retirement is an important factor for explaining these heterogeneities. Future work can investigate the relationship between the heterogeneity of the effect of retirement on health observed in many countries and the one of the change in health investment behaviors after retirement.

References

Ayyagari, P. (2014). The Impact of Retirement on Smoking Behavior. *Eastern Economic Journal*, pages 1–18.

Behncke, S. (2012). Does Retirement Trigger Ill Health? *Health economics*, 21:282–300.

Bingley, P. and Martinello, A. (2013). Mental retirement and schooling. *European Economic Review*, 63:292–298.

Bonsang, E., Adam, S., and Perelman, S. (2012). Does retirement affect cognitive functioning? *Journal of Health Economics*, 31(3):490–501.

Bound, J. and Waidmann, T. (2007). Estimating the Health Effects of Retirement. *University of Michigan Retirement Research Center Working Papers*.

Charles, K. K. (2004). Is Retirement Depressing?: Labor Force Inactivity and Psychological Well being in Later Life. *Research in Labor Economics*, 23:269–299.

Coe, N. B., Gaudecker, H.-M. V., Lindeboom, M., and Maurer, J. (2012). The Effect Of Retirement On Cognitive Functioning. *Health Economics*, 21:913–927.

Coe, N. B. and Lindeboom, M. (2008). Does Retirement Kill You? Evidence from Early Retirement Windows. *IZA Discussion Papers*, 93.

Coe, N. B. and Zamarro, G. (2011). Retirement effects on health in Europe. *Journal of Health Economics*, 30(1):77–86.

Dave, D., Rashad, I., and Spasojevic, J. (2008). The Effects of Retirement on Physical and Mental Health Outcomes. *Southern Economic Journal*, 75(2):497–523.

Devore, E. E., Stampfer, M. J., Breteler, M. M. B., Rosner, B., Kang, J. H., Okereke, O., Hu, F. B., and Grodstein, F. (2009). Dietary fat intake and cognitive decline in women with type 2 diabetes. *Diabetes Care*, 32(4):635–640.

Eibich, P. (2015). Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity. *Journal of Health Economics*, 43:1–12.

Fonseca, R., Kapteyn, A., Lee, J., Zamarro, G., and Feeney, K. (2014). A Longitudinal Study of Well-being of Older Europeans: Does Retirement Matter? *Journal of population ageing*, 7(1):21–41.

Godard, M. (2016). Gaining weight through retirement ? Results from the SHARE survey. *Journal of Health Economics*, 45:27–46.

Gruber, J. and Wise, D. A. (1998). Social Security and Retirement: An International Comparison. *American Economic Review, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association*, 88(2):158–163.

Hashimoto, H. (2013). Health Consequences of Transitioning to Retirement and Social Participation: Results based on JSTAR panel data. *RIETI Discussion Paper Series 13-E-078*.

Hashimoto, H. (2015). Impacts of Leaving Paid Work on Health, Functions, and Lifestyle Behavior: Evidence from JSTAR panel data. *RIETI Discussion Paper Series 15-E-114*.

Hernaes, E., Markussen, S., Piggott, J., and Vestad, O. L. (2013). Does retirement age impact mortality? *Journal of Health Economics*, 32(3):586–598.

Insler, M. (2014). The Health Consequences of Retirement. *Journal of Human Resources*, 49(1):195–233.

Iparraguirre, J. (2014). Physical Functioning in work and retirement : commentary on age-related trajectories of physical functioning in work and retirement - the role of sociodemographic factors , lifestyle and disease by Stenholm et al. *Journal of Epidemiology and Community Health*, 68:493–499.

Johnston, D. W. and Lee, W.-S. (2009). Retiring to the good life? The short-term effects of retirement on health. *Economics Letters*, 103(1):8–11.

Kajitani, S. (2011). Working in old age and health outcomes in Japan. *Japan and The World Economy*, 23:153–162.

Kajitani, S., Sakata, K., and McKenzie, C. (2014). Occupation, Retirement and Cognitive Functioning. *Discussion Paper Series Graduate School and School of Economics Meisei University*, Discussion Paper Series, No.27.

Kajitani, S., Sakata, K., and McKenzie, C. (2016). Occupation, retirement and cognitive functioning. *Ageing and Society*, May 2016:1–29.

Kerkhofs, M. and Lindeboom, M. (1997). Age related health dynamics and changes in labour market status. *Health economics*, 6(4):407–423.

Latif, E. (2011). The impact of retirement on psychological well-being in Canada. *Journal of Socio-Economics*, 40(4):373–380.

Latif, E. (2013). The impact of retirement on mental health in Canada. *Journal of Mental Health Policy and Economics*, 16(1):35–46.

Lee, J. and Smith, J. P. (2009). Work, Retirement, and Depression. *Journal of Population Ageing*, 2(1-2):57–71.

Lindeboom, M. and Kerkhofs, M. (2009). Health and Work of The Elderly: Subjective Health Measures, Reporting Errors and Endogeneity in The Relationship Between Health and Work. *Journal of Applied Econometrics*, 24:1024–1046.

Lindeboom, M., Portrait, F., and Van Den Berg, G. J. (2002). An econometric analysis of the mental-health effects of major events in the life of older individuals. *Health Economics*, 11(6):505–520.

Mazzonna, F. and Peracchi, F. (2012). Ageing, cognitive abilities and retirement. *European Economic Review*, 56(4):691–710.

McEwen, B. S. and Sapolsky, R. M. (1995). Stress and Cognitive Function. *Current Opinion in Neurobiology*, 5:205–216.

McGarry, K. (2004). Health and Retirement: Do Changes in Health Affect Retirement Expectations? *Journal of Human Resources*, 39(3):624–648.

Motegi, H., Nishimura, Y., and Terada, K. (2016). Does Retirement Change Lifestyle Habits? *Japanese Economic Review*, 67(2):169–191.

Neuman, K. (2008). Quit your job and get healthier? The effect of retirement on health. *Journal of Labor Research*, 29(2):177–201.

Rohwedder, S. and Willis, R. J. (2010). Mental Retirement. *Journal of Economic Perspectives*, 24(1):119–138.

Schaffer, M. (2010). xtivreg2: Stata module to perform extended iv/2sls, gmm and ac/hac, liml and k-class regression for panel data models. Available online via http://repec.org/bocode/x/xtivreg2.html (last accessed 19.9.2016).

van der Heide, I., van Rijn, R. M., Robroek, S. J., Burdorf, A., and Proper, K. I. (2013). Is retirement good for your health? A systematic review of longitudinal studies. *BMC public health*, 13(1):1180.

Wang, M. and Shi, J. (2014). Psychological research on retirement. *Annual review of psychology*, 65:209–33.

Zhao, M., Konishi, Y., and Noguchi, H. (2013). Retiring for Better Health Evidence from Health Investment Behaviors in Japan. Mimeo.

Zhu, R. (2016). Retirement and its consequences for women's health in Australia. *Social Science & Medicine*, 163(2016):117–125.

Zhu, R. and He, X. (2015). How does women's life satisfaction respond to retirement? A two-stage analysis. *Economics Letters*, 137:118–122.

Figure 1: Replacement 1

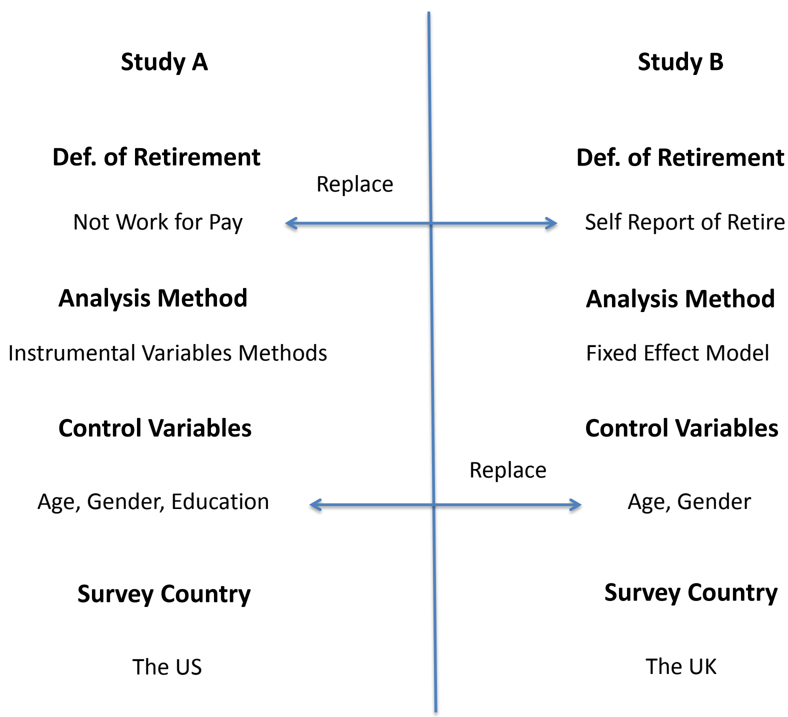


Table 1: Summary Statistics of Health Outcomes (Age 50 or older) around 2010

	Obs.	Mean	S.D.	Min	Max
HRS					
Word Recall Summary Score	19681	9.61	3.41	0	20
Serial 7's Score	19681	3.41	1.68	0	5
Poor health	21029	0.09	0.28	0	1
ADL summary score (0-3)	20892	0.25	0.66	0	3
CESD summary score (0-8)	19480	1.51	2.03	0	8
BMI	20645	28.46	6.16	7	79
ELSA*1					
Word Recall Summary Score	9536	10.40	3.73	0	20
Poor health	9570	0.08	0.27	0	1
ADL summary score (0-3)	10087	0.26	0.63	0	3
CESD summary score (0-8)	9435	1.51	1.96	0	8
BMI*2	8230	28.26	5.30	15	71
SHARE*3					
Word Recall Summary Score	55472	8.91	3.76	0	20
Serial 7's Score	53332	3.78	1.75	0	5
Poor health	56790	0.13	0.33	0	1
ADL summary score (0-3)	56770	0.17	0.53	0	3
EURO-D summary score (0-12)	55229	2.58	2.31	0	12
BMI	54110	26.92	4.93	6	222
JSTAR					
Word Recall Summary Score	1690	9.56	3.04	0	20
Serial 7's Score	1740	4.16	1.18	0	5
Poor health	2263	0.03	0.17	0	1
ADL summary score (0-3)	2265	0.05	0.33	0	3
CESD summary score (0-8)	1865	2.11	1.75	0	8
BMI	2222	23.52	2.96	13	41
KLoSA					
Word Recall Summary Score*4	7231	4.48	1.57	0	6
Serial 7's Score	7231	3.57	1.76	0	5
Poor health	7649	0.24	0.43	0	1
ADL summary score (0-3)	7649	0.10	0.49	0	3
CESD summary score (0-7)	7596	2.64	1.95	0	7
BMI	7458	23.20	2.81	12	37

*1: No Serial 7's Score in ELSA.

*2: We use BMI in Wave 4 ELSA because Wave 5 ELSA does not include BMI.

*3: Calculated using weight.

*4: KLoSA's Word Recall Scores are not comparable with other dataset.

Table 2: Summary Statistics: The US (Age 50 or older) at 2010

	Obs.	**Mean**	**S.D.**	**Min**	**Max**	**Obs.**	**Mean**	**S.D.**	**Min**	**Max**
			Male					**Female**		
Word Recall Summary Score	8291	9.07	3.31	0	20	11390	10.01	3.42	0	20
Serial 7's Score	8291	3.66	1.57	0	5	11390	3.22	1.74	0	5
Poor health	8993	0.08	0.28	0	1	12036	0.09	0.29	0	1
ADL summary score (0-3)	8929	0.22	0.61	0	3	11963	0.27	0.70	0	3
CESD summary score (0-8)	8202	1.30	1.88	0	8	11278	1.67	2.12	0	8
BMI	8904	28.42	5.27	7	61	11741	28.49	6.75	9	79
			Not Univ. Graduate					**Univ. Graduate**		
Word Recall Summary Score	15286	9.18	3.32	0	20	4391	11.12	3.29	0	20
Serial 7's Score	15286	3.17	1.73	0	5	4391	4.21	1.18	0	5
Poor health	16441	0.10	0.30	0	1	4584	0.03	0.18	0	1
ADL summary score (0-3)	16332	0.29	0.70	0	3	4556	0.13	0.49	0	3
CESD summary score (0-8)	15116	1.67	2.10	0	8	4360	0.96	1.63	0	8
BMI	16103	28.69	6.30	7	79	4538	27.65	5.53	12	61
			White-collar workers					**Blue-collar workers**		
Word Recall Summary Score	8634	10.16	3.43	0	20	3187	8.52	3.27	0	20
Serial 7's Score	8634	3.65	1.59	0	5	3187	3.14	1.74	0	5
Poor health	9095	0.06	0.24	0	1	3528	0.10	0.30	0	1
ADL summary score (0-3)	9082	0.20	0.61	0	3	3528	0.27	0.68	0	3
CESD summary score (0-8)	8560	1.26	1.87	0	8	3147	1.49	1.98	0	8
BMI	8993	28.12	5.92	7	72	3491	28.57	5.68	11	59

Table 3: The datasets which we use in each section

	Wave	**Year**
Section 5 (The Harmonized Analysis)		
The HRS	3-11	1996-2011
The SHARE*[1]	1,2,4,5	2004-2006, 2010-2012
The ELSA	1-6	2002-2014
The JSTAR	1-4	2007-2013
The KLoSA	1-4	2006-2012

*[1]: We analyze only Denmark, France, Germany, and Switzerland.

Table 4: The Targeted Literature

	(1)Cognition	
	Bonsang et al. (2012)	**Coe and Zamarro (2011)**
Impact	Negative	No
Survey countries	United States	European countries
Dataset	HRS	SHARE
Index	Word recall	Word recall, Verbal fluency
Method	FE-IV	IV
Definition of Retirement	Retired for at least one year	Not working for pay*2
Control variables*1	Only Age variables	B, E, L, H

	(2)Self-report of health	
	Dave et al. (2008)	**Coe and Zamarro (2011)**
Impact	Negative	Positive
Survey countries	United States	European countries
Dataset	HRS	SHARE
Method	FE	IV
Definition of Retirement	Reporting retired and not working	Not working for pay*2
Control variables*1	B, E	B, E, L

	(3)Depression	
	Dave et al. (2008)	**Coe and Zamarro (2011)**
Impact	Negative	No
Survey countries	United States	European countries
Dataset	HRS	SHARE
Index	CESD	EUROD
Method	FE	IV
Definition of Retirement	Reporting retired and not working	Not working for pay*2
Control variables*1	B, E	B, E, L

	(4)ADL	
	Dave et al. (2008)	**Neuman (2008)**
Impact	Negative	No (Male)/Positive(Female)
Survey countries	United States	United States
Dataset	HRS	HRS
Method	FE	IV
Definition of Retirement	Reporting retired and not working	Work less than 1200 h per year
Control variables*1	B, E	B, E, H

	(5)Obesity	
	Johnston and Lee (2009)	**Godard (2016)**
Impact	No	Negative
Survey countries	England	European countries
Dataset	HSE	SHARE
Index	BMI	BMI≥30
Method	RDD	FEIV
Definition of Retirement	Reporting retired	Reporting retired
Control variables*1	No	B

*1 B:Basic variables(Ex:Age, education), E:Economic variables(Ex:Income), L:Labor force status(Ex:Self-employed), H:Health variables(Ex:Number of ADLs).

*2 Strictly speaking, the retirement variable definition used in Coe and Zamarro (2011) is reporting retired, unemployed, a homemaker, sick and disabled.

Table 5: Cognitive score

	Bonsang et al. (2012)		Coe and Zamarro (2011)
Estimated result in the original paper	-0.942***	→	-0.0390
Our replication	-1.036*** (0.005*2)		
Repracing ...	Def. of retire	→	-1.244***
	Controls	→	-1.189***
	Method	→	-1.444***
	Sample	→	-1.266*
	Survey country	→	15.570**

	Coe and Zamarro (2011)		Bonsang et al. (2012)
Estimated result in the original paper	-0.0390	→	-0.942***
Our replication	-0.120 (0.004*2)		
Repracing ...	Def. of retire	→	-0.035
	Controls	→	-4.647***
	Method	→	1.778
	Sample	→	-0.321
	Survey country	→	-2.649**

*1 The red (blue) character indicates the positive (negative) impact.

*2 $\frac{\text{(Coeff.: original paper)} - \text{(Coeff. our replication)}}{\text{(maximum value of index)}}$

Figure 2: Replacement 2

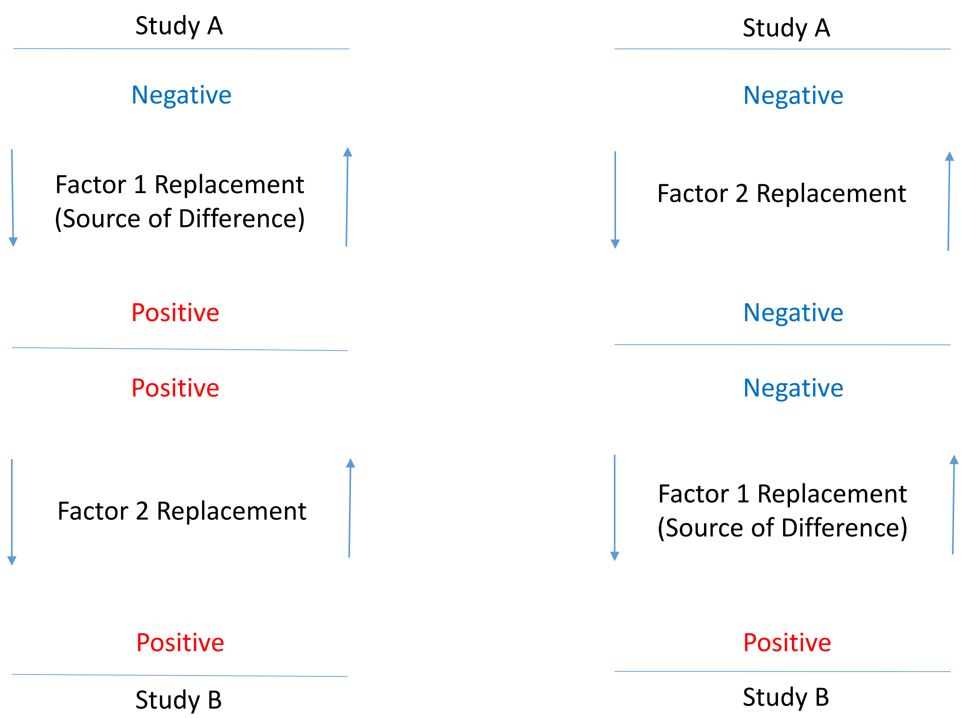


Table 6: Cognitive score

Panel A: change the order of replacing "methods+controls"

Pattern A	Pattern B
Bonsang et al. (2012)	**Bonsang et al. (2012)**
-0.942***(Original paper)	-0.942***(Original paper)
-1.036***(Replication)	-1.036***(Replication)
↓ Method + Controls + Country ↑	↓ Def. of Retirement ↑
-0.348	-1.244***
↓ Def. of Retirement ↑	↓ Method + Controls + Country ↑
-0.321	-0.321
↓ Sample ↑	↓ Sample ↑
-0.120(Replication)	-0.120(Replication)
-0.0390(Original paper)	-0.0390(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**

Pattern C
Bonsang et al. (2012)
-0.942***(Original paper)
-1.036***(Replication)
↓ Def. of Retirement ↑
-1.244***
↓ Sample ↑
-1.825*
↓ Method + Controls + Country ↑
-0.120(Replication)
-0.0390(Original paper)
Coe and Zamarro (2011)

Panel B: change the order of replacing the retirement definition

Pattern A	Pattern B
Bonsang et al. (2012)	**Bonsang et al. (2012)**
-0.942***(Original paper)	-0.942***(Original paper)
-1.036***(Replication)	-1.036***(Replication)
↓ Def. of Retirement ↑	↓ Method + Controls + Country ↑
-1.244***	-0.348
↓ Method + Controls + Country ↑	↓ Def. of Retirement ↑
-0.321	-0.321
↓ Sample ↑	↓ Sample ↑
-0.120(Replication)	-0.120(Replication)
-0.0390(Original paper)	-0.0390(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**

Pattern C
Bonsang et al. (2012)
-0.942***(Original paper)
-1.036***(Replication)
↓ Method + Controls + Country ↑
-0.348
↓ Sample ↑
-0.035
↓ Def. of Retirement ↑
-0.120(Replication)
-0.0390(Original paper)
Coe and Zamarro (2011)

Table 7: Self-report of health

Panel A: change the order of replacing "methods+controls"

Pattern A	Pattern B
Dave et al. (2008)	**Dave et al. (2008)**
0.0268***(Original paper)	0.0268***(Original paper)
0.025***(Replication)	0.025***(Replication)
↓ Method + Controls + Country + Index ↑	↓ Sample ↑
-3.570	0.027***
↓ Sample ↑	↓ Method + Controls + Country + Index ↑
-0.314**	-0.314**
↓ Def. of Retirement ↑	↓ Def. of Retirement ↑
-0.368**(Replication)	-0.368**(Replication)
-0.3545**(Original paper)	-0.3545**(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**

Pattern C
Dave et al. (2008)
0.0268***(Original paper)
0.025***(Replication)
↓ Sample ↑
0.027***
↓ Def. of Retirement ↑
0.051***
↓ Method + Controls + Country + Index ↑
-0.368**(Replication)
-0.3545**(Original paper)
Coe and Zamarro (2011)

Panel B: change the order of replacing the retirement definition

Pattern A	Pattern B
Dave et al. (2008)	**Dave et al. (2008)**
0.0268***(Original paper)	0.0268***(Original paper)
0.025*** (Replication)	0.025*** (Replication)
↓ Def. of Retirement ↑	↓ Method + Controls + Country + Index ↑
0.023***	-3.570
↓ Method + Controls + Country + Index ↑	↓ Def. of Retirement ↑
-0.545***	-0.545***
↓ Sample ↑	↓ Sample ↑
-0.368**(Replication)	-0.368**(Replication)
-0.3545**(Original paper)	-0.3545**(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**

Pattern C
Dave et al. (2008)
0.0268***(Original paper)
0.025***(Replication)
↓ Method + Controls + Country + Index ↑
-3.570
↓ Sample ↑
-0.314**
↓ Def. of Retirement ↑
-0.368**(Replication)
-0.3545**(Original paper)
Coe and Zamarro (2011)

Table 8: ADL

Panel A: change the order of replacing "methods+controls"		Panel B: change the order of replacing the retirement definition	
Pattern A	**Pattern B**	**Pattern A**	**Pattern B**
Dave et al. (2008)	**Dave et al. (2008)**	**Dave et al. (2008)**	**Dave et al. (2008)**
0.0267***(Original paper)	0.0267***(Original paper)	0.0267***(Original paper)	0.0267***(Original paper)
0.043***(Replication)	0.043***(Replication)	0.043***(Replication)	0.043***(Replication)
↓ Method + Controls ↑	↓ Def. of Retirement ↑	↓ Def. of Retirement ↑	↓ Method + Controls ↑
-0.01	0.021***	0.021***	-0.01
↓ Def. of Retirement ↑	↓ Method + Controls ↑	↓ Method + Controls ↑	↓ Def. of Retirement ↑
0.01	0.01	0.01	0.01
↓ Sample ↑	↓ Sample ↑	↓ Sample ↑	↓ Sample ↑
-0.013(M)/0.211***(F)(Replication)	-0.013(M)/0.211***(F)(Replication)	-0.013(M)/0.211***(F)(Replication)	-0.013(M)/0.211***(F)(Replication)
-0.025(M)/0.101**(F)(Original paper)	-0.025(M)/0.101**(F)(Original paper)	-0.025(M)/0.101**(F)(Original paper)	-0.025(M)/0.101**(F)(Original paper)
Neuman (2008)	**Neuman (2008)**	**Neuman (2008)**	**Neuman (2008)**

Panel A	Panel B
Pattern C	**Pattern C**
Dave et al. (2008)	**Dave et al. (2008)**
0.0267***(Original paper)	0.0267***(Original paper)
0.043***(Replication)	0.043*** (Replication)
↓ Def. of Retirement ↑	↓ Method + Controls ↑
0.021***	-0.01
↓ Sample ↑	↓ Sample ↑
0.062***(M)/0.084***(F)	-0.03(M)/0.219***(F)
↓ Method + Controls ↑	↓ Def. of Retirement ↑
-0.013(M)/0.211***(F)(Replication)	-0.013(M)/0.211***(F)(Replication)
-0.025(M)/0.101**(F)(Original paper)	-0.025(M)/0.101**(F)(Original paper)
Neuman (2008)	**Neuman (2008)**

Table 9: Depression

Panel A: change the order of replacing "methods+controls"		Panel B: change the order of replacing the retirement definition	
Pattern A	**Pattern B**	**Pattern A**	**Pattern B**
Dave et al. (2008)	**Dave et al. (2008)**	**Dave et al. (2008)**	**Dave et al. (2008)**
0.1157***(Original paper)	0.1157***(Original paper)	0.1157***(Original paper)	0.1157***(Original paper)
0.116***(Replication)	0.116***(Replication)	0.116***(Replication)	0.116***(Replication)
↓ Method + Controls ↑	↓ Def. of Retirement ↑	↓ Def. of Retirement ↑	↓ Method + Controls ↑
0.274	0.165***	0.165***	0.274
↓ Def. of Retirement ↑	↓ Method + Controls ↑	↓ Method + Controls ↑	↓ Def. of Retirement ↑
0.282	0.282	0.282	0.282
↓ Sample ↑	↓ Sample ↑	↓ Sample ↑	↓ Sample ↑
-0.227	-0.227	-0.227	-0.227
↓ Country + Index ↑	↓ Country + Index ↑	↓ Country + Index ↑	↓ Country + Index ↑
0.303(Replication)	0.303(Replication)	0.303(Replication)	0.303(Replication)
-0.0691(Original paper)	-0.0691(Original paper)	-0.0691(Original paper)	-0.0691(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**	**Coe and Zamarro (2011)**	**Coe and Zamarro (2011)**

Panel A		Panel B	
Pattern C	**Pattern D**	**Pattern C**	**Pattern D**
Dave et al. (2008)	**Dave et al. (2008)**	**Dave et al. (2008)**	**Dave et al. (2008)**
0.1157***(Original paper)	0.1157***(Original paper)	0.1157***(Original paper)	0.1157***(Original paper)
0.116***(Replication)	0.116***(Replication)	0.116***(Replication)	0.116***(Replication)
↓ Def. of Retirement ↑	↓ Def. of Retirement ↑	↓ Method + Controls ↑	↓ Method + Controls ↑
0.165***	0.165***	0.274	0.274
↓ Sample ↑	↓ Sample ↑	↓ Sample ↑	↓ Sample ↑
0.259***	0.259***	-0.285	-0.285
↓ Method + Controls ↑	↓ Country + Index ↑	↓ Def. of Retirement ↑	↓ Country + Index ↑
-0.227	-0.038	-0.227	0.267
↓ Country + Index ↑	↓ Method + Controls ↑	↓ Country + Index ↑	↓ Def. of Retirement ↑
0.303(Replication)	0.303(Replication)	0.303(Replication)	0.303(Replication)
-0.0691(Original paper)	-0.0691(Original paper)	-0.0691(Original paper)	-0.0691(Original paper)
Coe and Zamarro (2011)	**Coe and Zamarro (2011)**	**Coe and Zamarro (2011)**	**Coe and Zamarro (2011)**

Table 10: BMI

Panel A: change the order of replacing "methods+controls"		Panel B: change the order of replacing the index	
Pattern A	**Pattern B**	**Pattern A**	**Pattern B**
Godard (2016)	**Godard (2016)**	**Godard (2016)**	**Godard (2016)**
0.115**(Original paper)	0.115**(Original paper)	0.115**(Original paper)	0.115**(Original paper)
0.122**(Replication)	0.122**(Replication)	0.122**(Replication)	0.122**(Replication)
↓ Method + Controls + Sample ↑	↓ Def. of Retirement ↑	↓ Index ↑	↓ Method + Controls + Sample ↑
0.002	0.122**	0.371	0.002
↓ Def. of Retirement ↑	↓ Method + Controls + Sample ↑	↓ Method + Controls + Sample ↑	↓ Index ↑
0.002	0.002	0.291	0.291
↓ Country ↑	↓ Country ↑	↓ Def. of Retirement ↑	↓ Def. of Retirement ↑
-0.018	-0.018	0.291	0.291
↓ Index ↑	↓ Index ↑	↓ Country ↑	↓ Country ↑
0.118(Replication)	0.118(Replication)	0.118(Replication)	0.118(Replication)
0.092(Original paper)	0.092(Original paper)	0.092(Original paper)	0.092(Original paper)
Johnston and Lee (2009)	**Johnston and Lee (2009)**	**Johnston and Lee (2009)**	**Johnston and Lee (2009)**

Pattern C	**Pattern D**	**Pattern C**	**Pattern D**
Godard (2016)	**Godard (2016)**	**Godard (2016)**	**Godard (2016)**
0.115**(Original paper)	0.115**(Original paper)	0.115**(Original paper)	0.115**(Original paper)
0.122**(Replication)	0.122**(Replication)	0.122**(Replication)	0.122**(Replication)
↓ Def. of Retirement ↑	↓ Def. of Retirement ↑	↓ Method + Controls + Sample ↑	↓ Method + Controls + Sample ↑
0.122**	0.122**	0.002	0.002
↓ Country ↑	↓ Country ↑	↓ Def. of Retirement ↑	↓ Def. of Retirement ↑
-0.386	-0.386	0.002	0.002
↓ Method + Controls + Sample ↑	↓ Index ↑	↓ Index ↑	↓ Country ↑
-0.018	-2.057	0.291	-0.018
↓ Index ↑	↓ Method + Controls + Sample ↑	↓ Country ↑	↓ Index ↑
0.118(Replication)	0.118(Replication)	0.118(Replication)	0.118(Replication)
0.092(Original paper)	0.092(Original paper)	0.092(Original paper)	0.092(Original paper)
Johnston and Lee (2009)	**Johnston and Lee (2009)**	**Johnston and Lee (2009)**	**Johnston and Lee (2009)**

Table 11: The Results of 1st Stage Regression (only Poor health)

	US			UK			Denmark			France		
	Full	Male	Female	Full	Male	Female	Full	Male	Female	Full	Male	Female
$1\{Age_{it} \geq A_i^{eb}\}$	0.105***	0.120***	0.089***				0.115***	0.076***	0.124***	0.168***	0.157***	0.176***
	(0.005)	(0.007)	(0.006)				(0.017)	(0.023)	(0.027)	(0.020)	(0.030)	(0.026)
$1\{Age_{it} \geq A_i^{fb}\}$	-0.457***	-0.472***	-0.323***	0.153***	0.176***	0.131***	0.165***	0.150***	1.464***	1.775***	2.143***	1.481***
	(0.067)	(0.114)	(0.083)	(0.008)	(0.012)	(0.012)	(0.019)	(0.027)	(0.507)	(0.270)	(0.446)	(0.337)
$1\{Age_{it} \geq A_i^{fb}\} \times age_{it}$	0.008***	0.008***	0.006***	0.001***	0.000**	0.001***			-0.020**	-0.027***	-0.033***	-0.022***
	(0.001)	(0.002)	(0.001)	(0.000)	(0.000)	(0.000)			(0.008)	(0.004)	(0.007)	(0.005)
Observations	162130	68199	93931	45070	20062	25008	6672	3120	3552	11214	4894	6320

	Germany			Switzerland			Japan	Korea	
	Full	Male	Female	Full	Male	Female	Male	Full	Male
$1\{Age_{it} \geq A_i^{eb}\}$	0.142***	0.075**	0.180***	0.090***	0.061*	0.114***		-1.257***	-2.161***
	(0.024)	(0.037)	(0.034)	(0.026)	(0.037)	(0.038)		(0.253)	(0.326)
$1\{Age_{it} \geq A_i^{fb}\}$	0.107***	0.153***	0.092***	-2.062***	-1.519*	-2.578***	-1.409**	-0.043***	-0.067***
	(0.021)	(0.034)	(0.028)	(0.551)	(0.857)	(0.699)	(0.682)	(0.013)	(0.017)
$1\{Age_{it} \geq A_i^{fb}\} \times age_{it}$				0.034***	0.026*	0.042***	0.024**		
				(0.009)	(0.013)	(0.011)	(0.011)		
$1\{Age_{it} \geq A_i^{eb}\} \times age_{it}$								0.022***	0.038***
								(0.005)	(0.006)
Observations	5380	2512	2868	5358	1977	3381	3721	24353	10898

[1] Standard errors in parentheses and * ($p < .1$), **($p < .05$), ***($p < .01$).
[2] All specifications include age, age squared, married, number of children, HH income, housing, HH total wealth, region and wave dummy.

Figure 3: The Proportion of Retired Elderly By Age and Country (US, UK, Denmark and France, Germany, Switzerland, Japan(Male) and South Korea)

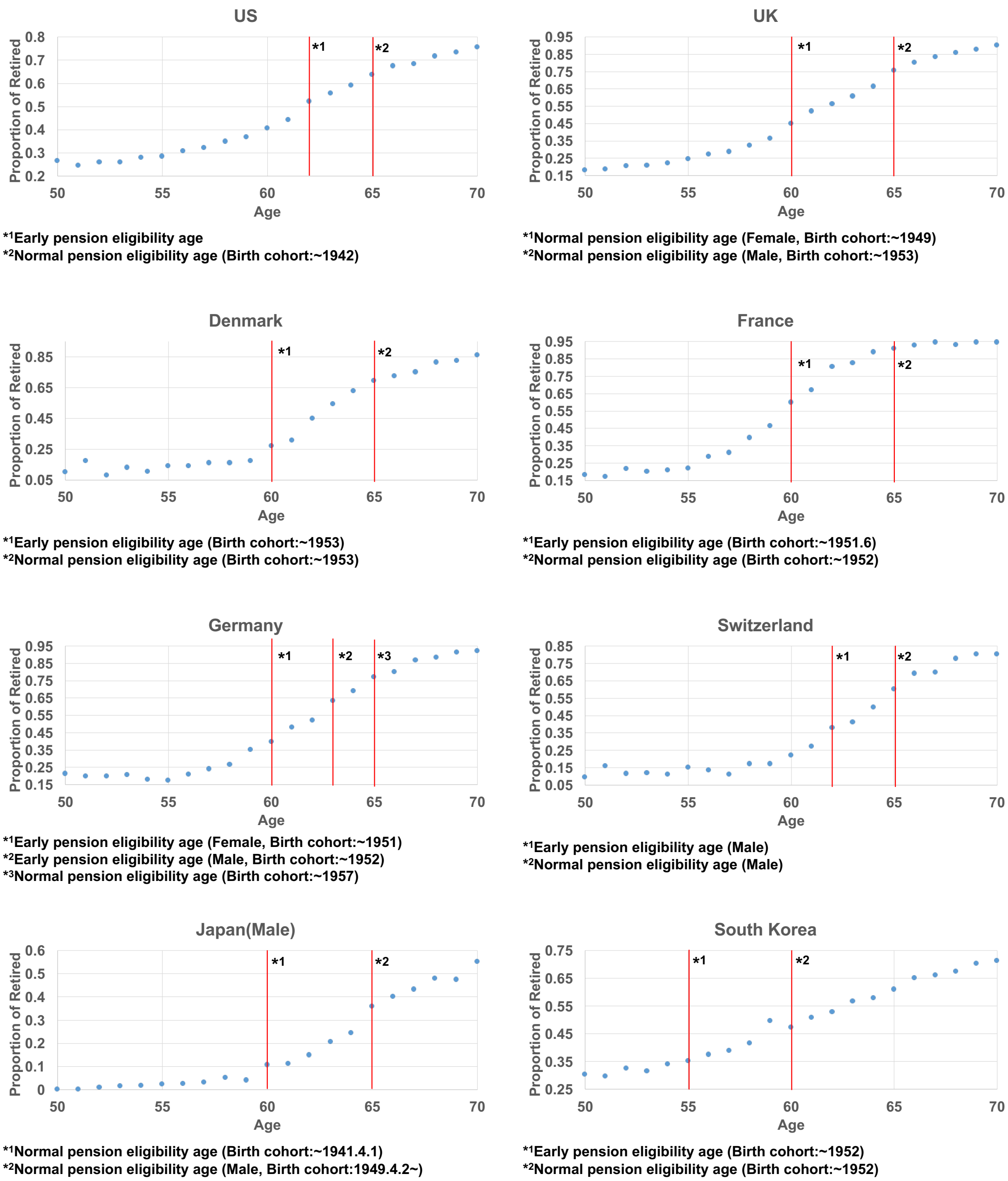


Table 12: FEIV estimation 1

Poor health	US		England		Denmark		France		Germany		Switzerland		Japan		Korea	
Full sample																
FE	0.034***	DWH p-val	0.020***	DWH p-val	**0.003**	DWH p-val	-0.007	DWH p-val	-0.013	DWH p-val	**0.004**	DWH p-val			0.063***	DWH p-val
FE-IV	-0.138***	0.000	-0.097***	0.000	-0.053	0.297	-0.158***	0.016	-0.143*	0.079	0.003	0.997			0.071	0.996
Male																
FE	0.041***	DWH p-val	0.026***	DWH p-val	**-0.009**	DWH p-val	**-0.001**	DWH p-val	**0.003**	DWH p-val	**-0.005**	DWH p-val	-0.035***	DWH p-val	0.072***	DWH p-val
FE-IV	-0.119***	0.000	-0.061*	0.010	-0.012	0.967	-0.070	0.674	-0.059	0.586	-0.073	0.352	**0.146**	0.070	-0.060	0.362
Female																
FE	0.029***	DWH p-val	0.014***	DWH p-val	**0.017**	DWH p-val	-0.011	DWH p-val	-0.025	DWH p-val	**0.010**	DWH p-val				
FE-IV	-0.159***	0.000	-0.154***	0.000	-0.080	0.144	-0.206***	0.015	-0.201*	0.078	0.011	0.939				

Depression	US(CESD(0-8))		England(CESD(0-8))		Denmark(EuroD)		France(EuroD)		Germany(EuroD)		Switzerland(EuroD)		Japan(CESD(0-8))		Korea(CESD(0-7))	
Full sample																
FE	0.193***	DWH p-val	0.084***	DWH p-val	-0.121	DWH p-val	**0.020**	DWH p-val	**-0.035**	DWH p-val	**-0.064**	DWH p-val			**0.045**	DWH p-val
FE-IV	-1.153***	0.000	-0.501***	0.000	-1.336***	0.005	-0.040	0.887	-0.194	0.736	-0.099	0.954			1.155	0.252
Male																
FE	0.194***	DWH p-val	0.043	DWH p-val	-0.158	DWH p-val	**0.137**	DWH p-val	**0.047**	DWH p-val	**-0.004**	DWH p-val	-0.001	DWH p-val	0.105*	DWH p-val
FE-IV	-0.865***	0.000	-0.586***	0.001	-1.358**	0.073	0.881	0.224	-0.192	0.725	0.021	0.910	-2.234*	0.064	1.185	0.190
Female																
FE	0.189***	DWH p-val	0.116***	DWH p-val	-0.067	DWH p-val	**-0.083**	DWH p-val	**-0.083**	DWH p-val	**-0.103**	DWH p-val				
FE-IV	-1.308***	0.000	-0.406*	0.027	-1.265**	0.027	-0.793	0.245	-0.736	0.390	-0.008	0.961				

Word Recall	US		England		Denmark		France		Germany		Switzerland		Japan		Korea	
Full sample																
FE	-0.102***	DWH p-val	**0.039**	DWH p-val	**-0.014**	DWH p-val	**0.140**	DWH p-val	**-0.176**	DWH p-val	**0.091**	DWH p-val			0.037	DWH p-val
FE-IV	-0.091	0.936	0.358	0.236	0.585	0.397	-0.356	0.445	0.949	0.202	1.359	0.169			1.895**	0.007
Male																
FE	-0.092**	DWH p-val	**-0.010**	DWH p-val	**-0.115**	DWH p-val	**-0.056**	DWH p-val	**-0.354**	DWH p-val	**-0.176**	DWH p-val			0.046	DWH p-val
FE-IV	-0.781**	0.072	0.102	0.817	0.576	0.571	-0.478	0.647	1.002	0.221	1.166	0.403			1.316**	0.030
Female																
FE	-0.122***	DWH p-val	**0.081**	DWH p-val	**0.060**	DWH p-val	**0.310**	DWH p-val	**-0.043**	DWH p-val	**0.269**	DWH p-val				
FE-IV	0.354	0.164	0.255	0.644	0.688	0.464	-0.284	0.494	1.707	0.149	1.248	0.388				

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

All specifications include age, age squared, married, number of children, HH income, housing, HH total wealth, region and wave dummy.

The red (blue, **bold**) character indicates the positive and significant (negative and significant, **insignificant**) impact.

Table 13: FEIV estimation 2

BMI	US		England		Denmark		France		Germany		Switzerland		Japan		Korea	
Full sample																
FE	0.115***	DWH p-val	0.124**	DWH p-val	**-0.035**	DWH p-val	**0.136**	DWH p-val	**-0.048**	DWH p-val	**0.072**	DWH p-val			**0.016**	DWH p-val
FE-IV	1.406***	0.000	0.179	0.840	0.121	0.730	-0.056	0.645	-0.331	0.708	0.776	0.245			0.532	0.612
Male																
FE	0.079**	DWH p-val	0.176**	DWH p-val	**0.037**	DWH p-val	**0.151**	DWH p-val	**-0.035**	DWH p-val	**0.080**	DWH p-val	0.185*	DWH p-val	**-0.092**	DWH p-val
FE-IV	1.419***	0.000	0.880**	0.065	1.003	0.169	-0.027	0.941	-0.073	0.901	0.785	0.375	2.796***	0.003	-0.504	0.535
Female																
FE	0.153***	DWH p-val	0.069	DWH p-val	**-0.108**	DWH p-val	**0.124**	DWH p-val	**-0.090**	DWH p-val	**0.076**	DWH p-val				
FE-IV	1.524***	0.000	**-0.833**	0.081	-0.686	0.406	-0.180	0.622	-0.939	0.340	0.796	0.438				

BMI≥30	US		England		Denmark		France		Germany		Switzerland		Japan		Korea	
Full sample																
FE	0.006**	DWH p-val	**-0.011**	DWH p-val	**0.007**	DWH p-val	**-0.007**	DWH p-val	**-0.013**	DWH p-val	**0.013**	DWH p-val			0.006**	DWH p-val
FE-IV	0.069***	0.013	-0.004	0.881	0.064	0.382	0.005	0.652	0.025	0.667	0.121*	0.106			-0.074	0.257
Male																
FE	**0.003**	DWH p-val	**-0.008**	DWH p-val	**0.022**	DWH p-val	**-0.015**	DWH p-val	**-0.015**	DWH p-val	**0.001**	DWH p-val	**0.002**	DWH p-val	**-0.001**	DWH p-val
FE-IV	0.039	0.339	0.062	0.291	0.098	0.477	0.062	0.365	-0.070	0.574	0.143	0.343	0.081	0.400	-0.036	0.302
Female																
FE	0.009**	DWH p-val	**-0.014**	DWH p-val	**-0.006**	DWH p-val	**0.000**	DWH p-val	**-0.012**	DWH p-val	0.021*	DWH p-val				
FE-IV	0.094***	0.016	-0.067	0.431	-0.004	0.969	-0.040	0.934	0.035	0.643	0.088	0.365				

ADL summary (0-3)	US		England		Denmark		France		Germany		Switzerland		Japan		Korea	
Full sample																
FE	0.041***	DWH p-val	0.044***	DWH p-val	0.030*	DWH p-val	**-0.018**	DWH p-val	0.003	DWH p-val	**-0.018**	DWH p-val			0.021**	DWH p-val
FE-IV	-0.484***	0.000	-0.146***	0.000		0.060	0.049	0.781	-0.328***	0.011	-0.173*	0.255			-0.471**	0.002
Male																
FE	0.056***	DWH p-val	0.040***	DWH p-val	**0.045**	DWH p-val	**-0.022**	DWH p-val	-0.014	DWH p-val	**-0.026**	DWH p-val	**-0.030**	DWH p-val	0.065***	DWH p-val
FE-IV	-0.376***	0.000	-0.168***	0.000	-0.108	0.308	0.155	0.437	-0.316*	0.051	-0.266	0.187	-0.134	0.665	-0.248*	0.011
Female																
FE	0.031***	DWH p-val	0.048***	DWH p-val	**0.017**	DWH p-val	**-0.015**	DWH p-val	0.020	DWH p-val	**-0.011**	DWH p-val				
FE-IV	-0.523***	0.000	-0.123**	0.003	-0.046	0.267	-0.009	0.367	-0.393**	0.011	-0.153	0.497				

Standard errors in parentheses

* $p < .1$, ** $p < .05$, *** $p < .01$

All specifications include age, age squared, married, number of children, HH income, housing, HH total wealth, region and wave dummy.

The red (blue, **bold**) character indicates the positive and significant (negative and significant, **insignificant**) impact.

Table 14: Robustness check: ADL (0-3)(Left) and Poor health (Self-report of health)(Right)

ADL(0-3)	Def. of retire	Controls Pattern 1	Pattern 2	Pattern 3	Pattern 4
United States	Not work	-0.493***	-0.484***	-0.473***	-0.484***
	Complete retire	-0.323***	-0.318***	-0.310***	-0.284***
England	Not work	-0.173***	-0.166***	-0.149***	-0.146***
	Complete retire	-0.102***	-0.098***	-0.090***	-0.088***
Denmark	Not work	-0.114	-0.112	-0.119	-0.104
	Complete retire	-0.008	-0.007	-0.127**	-0.120**
France	Not work	-0.009	-0.009	-0.017	-0.018
	Complete retire	-0.009	-0.01	-0.018	-0.019
Germany	Not work	-0.326***	-0.252**	-0.328***	-0.328***
	Complete retire	-0.206***	--0.170**	-0.227***	-0.225***
Switzerland	Not work	-0.004	-0.005	-0.017	-0.018
	Complete retire	0.019	-0.097*	0.018	0.017

Pattern 1 includes age and age squared.
Pattern 2 includes age, age squared, married and number of children(basic variables).
Pattern 3 includes basic variables and, HH income, housing and HH total wealth(economic variables).
Pattern 4 includes basic variables, economic variables and, region dummy and wave dummy.
* $p < .1$, ** $p < .05$, *** $p < .01$
The red (blue) character indicates the positive (negative) impact.

Poor health	Def. of retire	Controls Pattern 1	Pattern 2	Pattern 3	Pattern 4
United States	Not work	-0.107***	-0.110***	-0.107***	-0.138***
	Complete retire	-0.082***	-0.084***	-0.082***	-0.105***
England	Not work	-0.094***	-0.095***	-0.098***	-0.097***
	Complete retire	-0.059***	-0.060***	-0.062***	-0.061***
Denmark	Not work	0.010	0.008	0.002	0.003
	Complete retire	0.009	0.007	0.006	0.008
France	Not work	-0.135**	-0.130**	-0.149***	-0.158***
	Complete retire	-0.013	-0.011	-0.076**	-0.082**
Germany	Not work	--0.132*	-0.129*	-0.140*	-0.143*
	Complete retire	-0.025*	-0.023*	-0.023*	-0.023
Switzerland	Not work	0.002	0.002	0.002	0.004
	Complete retire	0.006	0.006	0.005	0.007

Pattern 1 includes age and age squared.
Pattern 2 includes age, age squared, married and number of children(basic variables).
Pattern 3 includes basic variables and, HH income, housing and HH total wealth(economic variables).
Pattern 4 includes basic variables, economic variables and, region dummy and wave dummy.
* $p < .1$, ** $p < .05$, *** $p < .01$
The red (blue) character indicates the positive (negative) impact.

Table 15: International comparison of the effect of retirement on health

	US	England	Denmark	France	Germany	Switzerland	Japan	South Korea
Self-report of health	+	+		+	+			-
Depression	+	+	+				+	
Cognition	-							+
BMI	-	-					-	
ADL	+	+			+			+

A Appendix

A.1 Pension Eligibility Age

To obtain pensionable age, we use the information from the Bureau of Labor Statistics in each country. However, this information is not available for some countries. Subsequently, we contact the Bureau of Labor Statistics or Bureau of Statistics directly, and obtain the information if possible. If we cannot find any information in the previous step, we use the OECD Pensions at a Glance, social security programs throughout the world (Europe, Asia and the Pacific, and the Americas) and The EUs Mutual Information System in Social Protection (MISSOC) as data sources. However, we cannot obtain the detailed pension eligibility age for many countries. Finally, the correct pension eligibility ages are obtained for the USA, the UK, Germany, France, Denmark, Switzerland, Czech, Estonia, Japan, China, and Korea. We do not consider countries where this information is missing. In this paper, we analyze the USA, the UK, Denmark, France, Germany, Switzerland, Japan and Korea. We show the pension eligibility ages used in this paper, as per the following tables.

Table A.1.1: Pension eligibility age in Section 5

Country: Birth cohort	**the U.S. PEA**
Early PEA	
	62y0m
Normal PEA	
~ 1937.12	65y0m
1938.1 ~ 1938.12	65y2m
1939.1 ~ 1939.12	65y4m
1940.1 ~ 1940.12	65y6m
1941.1 ~ 1941.12	65y8m
1942.1 ~ 1942.12	65y10m
1943.1 ~ 1943.12	66y0m
1944.1 ~ 1944.12	66y0m
1945.1 ~ 1945.12	66y0m
1946.1 ~ 1946.12	66y0m
1947.1 ~ 1947.12	66y0m
1948.1 ~ 1948.12	66y0m
1949.1 ~ 1949.12	66y0m
1950.1 ~ 1950.12	66y0m
1951.1 ~ 1951.12	66y0m
1952.1 ~ 1952.12	66y0m
1953.1 ~ 1953.12	66y0m
1954.1 ~ 1954.12	66y0m
1955.1 ~ 1955.12	66y2m
1956.1 ~ 1956.12	66y4m
1957.1 ~ 1957.12	66y6m
1958.1 ~ 1958.12	66y8m
1959.1 ~ 1959.12	66y10m
1960.1 ~ 1960.12	67y0m

Country: Birth cohort	**the U.K. PEA**
Normal PEA: Male	
~ 1953.12	65y0m
1954.1 ~ 1954.12	66y0m
1955.1 ~ 1959.12	66y0m
1960.1 ~ 1960.12	67y0m
1961.1 ~	67y0m
Normal PEA: Female	
~ 1949.12	60y0m
1950.1 ~ 1950.12	61y0m
1951.1 ~ 1951.12	62y0m
1952.1 ~ 1952.12	63y0m
1953.1 ~	65y0m

Country: Birth cohort	**Germany PEA**
Early PEA: Male	
~ 1952.12	63y0m
1953.1 ~ 1953.12	63y2m
1954.1 ~ 1954.12	63y4m
1955.1 ~ 1955.12	63y6m
1956.1 ~ 1956.12	63y8m
1957.1 ~ 1957.12	63y10m
1958.1 ~ 1958.12	64y0m
1959.1 ~ 1959.12	64y2m
1960.1 ~ 1960.12	64y4m
1961.1 ~ 1961.12	64y6m
1962.1 ~ 1962.12	64y8m
1963.1 ~ 1963.12	64y10m
1964.1 ~ 1964.12	65y0m
Early PEA: Female	
~ 1951.12	60y0m
Normal PEA	
~ 1946.12	65y0m
1947.1 ~ 1947.12	65y1m
1948.1 ~ 1948.12	65y2m
1949.1 ~ 1949.12	65y3m
1950.1 ~ 1950.12	65y4m
1951.1 ~ 1951.12	65y5m
1952.1 ~ 1952.12	65y6m
1953.1 ~ 1953.12	65y7m
1954.1 ~ 1954.12	65y8m
1955.1 ~ 1955.12	65y9m
1956.1 ~ 1956.12	65y10m
1957.1 ~ 1957.12	65y11m
1958.1 ~ 1958.12	66y0m
1959.1 ~ 1959.12	66y2m
1960.1 ~ 1960.12	66y4m
1961.1 ~ 1961.12	66y6m
1962.1 ~ 1962.12	66y8m
1963.1 ~ 1963.12	66y10m
1964.1 ~ 1964.12	67y0m

Country: Birth cohort	**France PEA**
Early PEA	
~ 1951.6	60y0m
1951.7 ~ 1951.12	60y4m
1952.1 ~ 1952.12	60y9m
1953.1 ~ 1953.12	61y2m
1954.1 ~ 1954.12	61y7m
1955.1 ~ 1955.12	62y0m
1956.1 ~ .	62y0m
Normal PEA	
~ 1951.6	65y0m
1951.7 ~ 1951.12	65y4m
1952.1 ~ 1952.12	65y9m
1953.1 ~ 1953.12	66y2m
1954.1 ~ 1954.12	66y7m
1955.1 ~ 1955.12	67y0m
1956.1 ~ .	67y0m

Table A.1.2: Pension eligibility age in Section 5

Country: Birth cohort	Denmark PEA
Early PEA	
~ 1953.12	60y0m
1954.1 ~ 1954.6	60y6m
1954.7 ~ 1954.12	61y0m
1955.1 ~ 1955.6	61y6m
1955.7 ~ 1955.12	62y0m
1956.1 ~ 1956.6	62y6m
1956.7 ~ 1958.12	63y0m
1959.1 ~ 1959.6	63y6m
1959.7 ~ 1964.6	64y0m
1964.7 ~	64y0m
Normal PEA	
~ 1953.12	65y0m
1954.1 ~ 1954.6	65y6m
1954.7 ~ 1954.12	66y0m
1955.1 ~ 1955.6	66y6m
1955.7 ~ 1955.12	67y0m
1956.1 ~ 1956.6	67y0m
1956.7 ~ 1958.12	67y0m
1959.1 ~ 1959.6	67y0m
1959.7 ~ 1964.6	67y0m
1964.7 ~	67y0m

Country: Birth cohort	Switzerland PEA
Early PEA: Male	
~ 1924.12	63y0m
1925.1 ~ 1950.12	63y0m
Early PEA: Female	
~ 1937.12	60y0m
1938.1 ~ 1940.12	61y0m
1941.1 ~	62y0m
Normal PEA: Male	
~ 1924.12	65y0m
1925.1 ~ 1950.12	65y0m
Normal PEA: Female	
~ 1937.12	62y0m
1938.1 ~ 1940.12	63y0m
1941.1 ~	64y0m

Country: Birth cohort	Japan PEA
Normal PEA: Male	
~1941.4.1	60y0m
1941.4.2~1943.4.1	61y0m
1943.4.2~1945.4.1	62y0m
1945.4.2~1947.4.1	63y0m
1947.4.2~1949.4.1	64y0m
1949.4.2~1953.4.1	65y0m
1953.4.2~1955.4.1	65y0m
1955.4.2~1957.4.1	65y0m
1957.4.2~1959.4.1	65y0m
1959.4.2~1961.4.1	65y0m
1961.4.2~	65y0m
Normal PEA: Female	
~1932.4.1	55y0m
1932.4.2~1934.4.1	56y0m
1934.4.2~1936.4.1	57y0m
1936.4.2~1937.4.1	58y0m
1937.4.2~1938.4.1	58y0m
1938.4.2~1940.4.1	59y0m
1940.4.2~1946.4.1	60y0m
1946.4.2~1948.4.1	61y0m
1948.4.2~1950.4.1	62y0m
1950.4.2~1952.4.1	63y0m
1952.4.2~1954.4.1	64y0m
1954.4.2~1958.4.1	65y0m
1958.4.2~1960.4.1	65y0m
1960.4.2~1962.4.1	65y0m
1962.4.2~1964.4.1	65y0m
1964.4.2~1965.4.1	65y0m
1965.4.2~	65y0m

Country: Birth cohort	Korea PEA
Early PEA	
~ 1952.12	55y0m
1953.1 ~ 1956.12	56y0m
1957.1 ~ 1960.12	57y0m
1961.1 ~ 1964.12	58y0m
1965.1 ~ 1968.12	59y0m
1969.1 ~ .	60y0m
Normal PEA	
~ 1952.12	60y0m
1953.1 ~ 1956.12	61y0m
1957.1 ~ 1960.12	62y0m
1961.1 ~ 1964.12	63y0m
1965.1 ~ 1968.12	64y0m
1969.1 ~ .	65y0m

A.2 Detailed Literature Review

The following tables show the detailed information on the literature.

Table A.2.1: Original index and Mental health 1

	Kerkhofs and Lindeboom	Lindeboom et al.	Charles	Bound and Waidmann	Coe and Lindeboom	Dave et al.	Neuman
	1997, Health Economics	2002, Health Economics	2004, Research in Labor Economics	2007, Univ. Michigan WP	2008, IZA DP	2008, Sourthern Economic Journal	2008, J of Labor Research
original index							
CESD		no			no	negative	no (M) no(F)
Depression							
SWB			positive			no(psychological problem	
SR health					positive (restricting within 2 years)	negative	positive(M) positive(F)
health fair poor				positive(M) positive(F)			
HSCL	positive						
Method	Fixed effect method	Fixed effect method	Iinstrumental variable method	pseudo RDD	Iinstrumental variable method	Fixed effect method	Iinstrumental variable method
Method (details)			IVs: Social security normal retirement age		IVs: Pension eligibility age	Restricting sample who has good health before retirement, and retire as of 62	IVs: public and private PEA for respondent and for spouse working more than 10 years
Def. of Retirement	not working for pay	early retirement	not working for pay not seeking work not worked for a year		Report to be out of the labor force or not having any paid employment	complete retirement (retired and not working)	Working less than 1200 hours in a year
Controls(Demog.)	age	age, residential area, marital status, family conditions, severe financial problems, severe conflict with others, victim of crimes	race, education, age, marital status		age, education, marital status, children	age, sex, race, marital status, education	age, education, race, whether parents living or not, children, marital status, region
Controls(Economic)					net worth deciles	income, asset	financial status
Controls(Working.)	labor market status	labor market status			job types (blue and white collar)		occupation
Controls(Health)		health conditions	health conditions		health outcome	lifestyle habits	early factors health beaviors
Data	CERRA 93, 95	Longitudinal Aging Study Amsterdam panel 92, 95, 98	HRS	ELSA 1st wave	HRS 1st-7th wave	HRS 1st-7th wave	HRS 1st-7th wave
Sample			60-69 years old male		male workers aged 55-70 years	50-75 years old and some manipulations	
Country	Netherlands	Netherlands	The U.S.	The U.K.	The U.S.	The U.S.	The U.S.

Table A.2.2: Original index and Mental health 2

	Johnston and Lee	Lee and Smith	Kajitani	Latif	Coe and Zamarro	Behncke	Fonseca et al.	Insler
	2009, Economics Letters	2009, J Population Aging	2011, Japan and the World Economy	2011, J Socio-Economics	2011, J Health Economics	2012, Health Economics	2014, J Population Aging	2014, J Human Resources
original index					positive*[1]	negative*[2]		positive (for the case of long term retirement)*[3]
CESD	positive(M)	no				no		
Depression	positive(mental health)				no(EUROD)			
SWB				positive			no (EUROD)	
SR health			positive(=1 if "excelelent" or "fairly good")		positive	negative		
health fair poor								
HSCL								
Method	RDD	Two-limit tobit and Probit	Probit	Fixed effect method and FE-IV	Iinstrumental variable method	Nonparametric matching	Iinstrumental variable method	FE-IV
Method (details)	Using 65 years as kink points robustness check by changing bandwidth		1st stage: Tobit estimation with the employment status(self-employed or not) and marital status as IVs; 2nd stage: Probit estimation.	IVs: pemsion eligibility age	IVs: eligibility age for early and full retirement	Using state pension eligibility age as IV	IVs: pension eligibility age	IVs; working expectations and preference derived from "workers' self-reported probabilities of working past ages 62 and 65."
Def. of Retirement	Retired from paid work	Answering retired from working, never worked, retired and unemployed	working hours per week	currentlly not working due to retirement	Not in the paid labor force	retired describes her current situation best and not in paid work was her activity in the last month	Answered retired	short retirement; retire at period t, long term retirement; retire before period t-1. self reported retirement (robustness check; Are you currently working for pay)
Controls(Demog.)		age, sex, education, marital status, children	age, age-squared, education, householder, large city, year dummy	age, sex, education, residential area, marital status	education, maritl status, children	many controls are use for estimating propensity score	age, sex, education, marital status	sex, age, marital status, education, race
Controls(Economic)		income, asset		income	income			asset
Controls(Working.)		employment status	longest-held occupation dummy		self employment		unemployed	
Controls(Health)		health condition and lifestyle habits	BMI category, illness of any member of the respondent's family	self reported health			disability and health conditions	
Data	Health Survey for England	KLoSA 1st wave	1990, 1993, and 1996 National Surveys of the Japanese Elderly	Canadian National Population Health Survey 1st-6th wave	SHARE 1st-2nd wave	ELSA 1st-3rd wave	SHARE 2004, 06, 10	HRS 1992˜2010
Sample	Male who do not have degree	50-64 years old	male over 60 years old	over 45 years old	50-69 years old male and some manipulations			restricting elderly working more than 10 years
Country	The U.K.	Korea	Japan	Canada	European countries	The U.K.	European countries	The U.S.

*[1] Predicted value of self-report of health using an ordered probit regression with objective health measures as covariates. See section 3.1.1 of Coe and Zamarro (2011) for more details.
*[2] Predicted value of self-report of health using an ordered probit regression with socio-economic variables and health indicators as covariates. See footnote 3 of Behncke (2012) for more details.
*[3] Predicted value of self-report of health using an ordered probit regression with objective binary response doctor-diagnosed health conditions as covariates. See section III.C of Insler (2014) for more details

Table A.2.3: Cognitive functioning and Physical function 1

	Lindeboom et al.	Bound and Waidmann	Coe and Lindeboom	Dave et al.	Neuman	Johnston and Lee	Rohwedder and Willis
	2002, Health Economics	2007, Univ. Michigan WP	2008, IZA DP	2008, Sourthern Economic Journal	2008, J of Labor Research	2009, Economics Letters	2010, J Econ Perspectives
cognitive functioning	no(MMSE (tests cognitive abilities))						negative
physical performance		no(M) negative(F)					
body nagi limitations		positive (M) positive(F)					
ADL			no	negative	no(M) positive(F)		
Body Mass Index						No	
Method	FE method	pseudo RDD	IV method	Fixed effect method	Iinstrumenta variable method	RDD	IV
Method (details)			IVs: pension eligibility age	Restricting sample who has good health before retirement, and retire as of 62	IVs: public and private PEA for respondent and for spouse working more than 10 years	Using 65 years as kink points robustness check by changing bandwidth	IVs: pension eligibility age for early and full
Def. of Retirement	early retirement		people report to be out of the labor force or not having any paid employment	complete retirement (retired and not working)	elderly working less than 1200 hours in a year	Retired from paid work	not having worked for pay in the last 4 weeks
Controls(Demog.)	age, residential area, marital status, family conditions, severe financial problems, severe conflict with others, victim of crimes		age, education, marital status, children	age, sex, race, marital status, education	age, education, race, whether parents living or not, children, marital status, region		
Controls(Economic)			net worth deciles	income, asset	financial status		
Controls(Working.)	labor market status		job types (blue and white collar)		occupation		
Controls(Health)	health conditions		health outcome	lifestyle habits	early factors health beaviors		
Data	Longitudinal Aging Study Amsterdam panel 92, 95, 98	ELSA 1st wave	HRS 1st-7th wave male workers aged 55-70 years	HRS 1st-7th wave	HRS 1st-7th wave	Health Survey for England	HRS ELSA SHARE at 2004
Sample				50-75 years old and some manipulations		Male who do not have degree	
Country	Netherlands	The U.K.	The U.S.	The U.S.	The U.S.	The U.K.	The U.S.The U.K.EU

Table A.2.4: Cognitive functioning and Physical function 2

	Coe and Zamarro	Behncke	Bonsang et al.	Mazzonna and Peracchi	Coe, Gaudecker, Lindeboom and Maurer	Bingley and Martinello	Godard
	2011, J Health Economics	2012, Health Economics	2012, J Health Economics	2012, European Economic Review	2012, Health Economics	2013, European Economic Review	2016, J Health Economics
cognitive functioning	no	negative	negative	negative	positive (blue collor) no (white collor)	negative	
physical performance							
body nagi limitations							
ADL		negative					
Body Mass Index							positive(BMI,M), no(BMI,F)
Method	Iinstrumenta variable method	Nonparametric matching	FE-IV method	IV method	Generalization of 2SLS	IV method	FE-IV method
Method (details)	IVs: eligibility age for early and full retirement	Using state pension eligibility age as IV	IVs: pension eligibility age	IVs: pension eligibility age for early and full	IVs: pension eligibility age (nonparametric regression of first stage regression)	IVs: pension eligibility age for early and full	IVs: pension eligibility age for early retirement age
Def. of Retirement	someone who is not in the paid labor force	retirede describes her current situation best and not in paid work was her activity in the last month	not having worked for pay in the last 1 year	max {0, current age-age as retirement} including unemployment elderly as retirement	interview year-retirement year (calculating by units of month and convert to the unit of year)	not having worked for pay in the last 4 weeks	self-declared current job situation (whether an individual is retired)
Controls(Demog.)	many controls are use for estimating propensity score	many controls are use for estimating propensity score	age	age and education	education, race, religion and age	age, sex, and education	age, age squared, year dummy, living with partners or not
Controls(Economic)							
Controls(Working.)							
Controls(Health)							
Data	SHARE 1st-2nd wave	ELSA 1st-3rd wave	HRS 1998˜2008 6 waves	SHARE 2004, 06	HRS, only male elderly born after 1931	HRS ELSA SHARE 2004	SHARE 2004, 2006, 2010.
Sample	50-69 years old male		51-75 years old and some manipulations		50-80 yearls old male	Dropping elderly whose educational variables are missing and restricting 60˜64.	restricting 50-69
Country	EU	The U.K.	The U.S.	EU	The U.S.	The U.S.The U.K.EU	European countries

Here, we show the rest of the results on the health indexes, which were not introduced in Section 2. We summarize the rest of the results on health indexes in Table A.2.5.

Table A.2.5: Illness

	Bound and Waidmann	Coe and Lindeboom	Dave et al.	Neuman	Johnston and Lee	Coe and Zamarro	Behncke	Hernaes et al.
	2007, Univ. Michigan WP	2008, IZA DP	2008, Southern Economic Journal	2008, J of Labor Research	2009, Economics Letters	2011, J Health Economics	2012, Health Economics	2013, J Health Economics
Metabolic Syndrome	positive(M) no(F)	positive (diabetes, restricting within 4 years)	negative(diabetes)		no(BMI,diabetes)		negative(metabolic syndrome) no(diabetes)	
heart risk	no(M) no(F)	no(heart attack)	no(heart desease)		no		no(heart attack) negative(again heart attack and stroke)	
mortality		no				no		no(M&F)
SPBB score	positive(M) no(F)							
heart diabeats diagnosis M	no(M) negative(F)							
choronic illness M	positive(M) positive(F)			no(M) no(F) choronic cdns			negative	
plain M	positive(M) positive(F)							
high blood pressure		no	no		no(hypertension)			
cancer		no	no				negative	
mobility			negative	no(M) no(F)			negative(difficulty walking)	
illness			negative					
stroke			no					
arthritis			negative					
difference in self ratings				no(M) no(F)				
large muscle functions				no(M) no(F)				
# days ill					positive			
asthma					no			
arthritis					no		no	
angina							no	
stroke							negative	
psychiatric							no	
health stock							negative	
limitting long standing illness							negative	
seeing difficulties							no	
hearing difficulties							negative	
high C-reactive protein (>3mg/L)							no	
high fibrinogen (7>mmol/L)							negative	
low haemoglobin (<12g/dl)							negative	
Method	pseudo RDD	IV method	Fixed effect method	IV method	RDD	IV method	Nonparametric matching	IV method and hazard model
Method (details)		IVs: pension eligibility age	Restricting sample who has good health before retirement, and retire as of 62	IVs: public and private PEA for respondent and for spouse working more than 10 years	Using 65 years as kink points robustness check by changing bandwidth	IVs: eligibility age for early and full retirement	Using state pension eligibility age as IV	IVs: entitle retirement age
Def. of Retirement		people report to be out of the labor force or not having any paid employment	complete retirement (retired and not working)	elderly working less than 1200 hours in a year	Retired from paid work	someone who is not in the paid labor force	retirede describes her current situation best and not in paid work was her activity in the last month	receving pension, other benefits or sharp drop of income
Controls(Demog.)		age, education, marital status, children	age, sex, race, marital status, education	age, education, race, whether parents living or not, children, marital status, region		education, maritl status, children	many controls are use for estimating propensity score	education, faculty, marital status,
Controls(Economic)		net worth deciles	income, asset	financial status		income		income, pension information
Controls(Working.)		job types (blue and white collar)		occupation		self employment		job industry
Controls(Health)		health outcome	lifestyle habits	early factors health beaviors				
Data	ELSA 1st wave	HRS 1st-7th wave	HRS 92˜05 7 wave	HRS 1992˜2004 7 wave. Only elderly consecutive for 3 years	Health Survey for England	SHARE 1st-2nd wave	ELSA 1st-3rd wave	administrativ data 1992˜2010
Sample		male workers aged 55-70 years	50-75 years old and some manipulations		Male who do not have degree	50-69 years old male		
Country	The U.K.	The U.S.	The U.S.	The U.S.	The U.K.	EU	The U.K.	Norway

A.3 Selection of Analyzed Papers

Table A.3.1 shows the number of studies in each step discussed in section 4.1.

Table A.3.1: Selection of Analyzed Papers (The Number of Studies in Each Step)

		Cognitive score *1	Self-report of health *2	ADL *3	Depression *4	BMI *5
Step 1:		8	6	4	8	2
Step 2:		7	5	4	7	
	Exclusion:	Lindeboom et al. (2002)	Kajitani (2011)		Lindeboom et al. (2002)	
Step 3:		6	4	3	6	
	Exclusion:	Rohwedder and Willis (2010)*6	Coe and Lindeboom (2008)	Coe and Lindeboom (2008)	Coe and Lindeboom (2008)	
Step 4:		5	3	2	5	
	Exclusion:	Behncke (2012)	Behncke (2012)	Behncke (2012)	Behncke (2012)	
Step 5:		2	2		2	
	Exclusion:	Mazzonna and Peracchi (2012) Coe et al. (2012) Bingley and Martinello (2013)	Neuman (2008)		Neuman (2008) Johnston and Lee (2009) Lee and Smith (2009)	

*1 The studies in the row "cognitive functioning" of Tables A.2.3 and A.2.4.
*2 The studies in the row "SR health" of Tables A.2.1 and A.2.2.
*3 The studies in the row "ADL" of Tables A.2.3 and A.2.4.
*4 The studies in the row "CESD" + "depression" of Tables A.2.1 and A.2.2.
*5 The studies in the row "Body Mass Index" of Tables A.2.3 and A.2.4.
*6 Rohwedder and Willis (2010) (published in the Journal of Economic Perspectives) is excluded in Step 3 because it only presents a rough idea for analyzing the effect of retirement on cognitive function.

A.4 Verification Framework 1 Results

We show the detailed results in section 4.2.

Self-report of health (Dave et al. (2008) versus Coe and Zamarro (2011)):

- According to Table A.4.1, when transplanting one factor from Dave et al. (2008) to Coe and Zamarro (2011), the replacement of the analysis method and the surveyed country change from a negative effect to no effect and vice-versa. The sensitivity of replacing the index, the control variables, the analysis method, and the surveyed country are important.

ADL (Dave et al. (2008) versus Neuman (2008)):

- We discuss Table A.4.2. Transplanting one factor from Dave et al. (2008) to Neuman (2008), the replacement of the estimation method and the sample selection method change from a negative effect to no effect, while replacing other factors does not produce such a difference, and vice-versa. This time, the replacement of each factor, except the definition of retirement, produces a change in the results, while the change in the estimation method produces the opposite result for female samples.

Depression (Dave et al. (2008) versus Coe and Zamarro (2011)):

- We discuss Table A.4.3. Transplanting one factor from Dave et al. (2008) to Coe and Zamarro (2011), the replacement of the estimation method and the surveyed country, [14)] change from a negative effect to no effect, while replacing other factors does not produce such a difference, and vice-versa. This time, the replacement of each factor, except the control variables, produces a change in the results.

BMI (Godard (2016) versus Johnston and Lee (2009)):

- We discuss Table A.4.4. Transplanting one factor from Godard (2016) to Johnston and Lee (2009), the replacement of all factors except the definition of retirement and the control variables change from a negative effect to no effect, while replacing other factors does not produce such a difference, and vice-versa. This time, the replacement of each factor does not produce a change in the results.

Table A.4.1: Self-report of health

	Dave et al. (2008)		Coe and Zamarro (2011)
Estimated result in the original paper	0.0268***	→	-0.3545**
Our replication	0.025*** (0.002*2)		
Repracing ...	Def. of retire	→	0.023***
	Controls	→	0.025***
	Method	→	0.02
	Sample	→	0.027***
	Survey country	→	-0.007

	Coe and Zamarro (2011)		Dave et al. (2008)
Estimated result in the original paper	-0.3545**	→	0.0268***
Our replication	-0.368** (0.014*2)		
Repracing ...	Index	→	-0.077
	Def. of retire	→	-0.314**
	Controls	→	-0.147
	Method	→	0.030**(poor health)
	Sample	→	-0.545***
	Survey country	→	-0.121(poor health)

*1 The red (blue) character indicates the positive (negative) impact.

*2 $\frac{\text{(Coeff.: original paper)} - \text{(Coeff. our replication)}}{\text{(maximum value of index)}}$

Table A.4.2: ADL

	Dave et al. (2008)		Neuman (2008)
Estimated result in the original paper	0.0267***	→	-0.025(M)/0.101**(F)
Our replication	0.043*** (-0.003*2)		
Repracing ...	Def. of retire	→	0.021***
	Controls	→	0.029***
	Method	→	0.142
	Sample	→	0.003(M)/0.004(F)

	Neuman (2008)		Dave et al. (2008)
Estimated result in the original paper	-0.025(M)/0.101**(F)	→	0.0267***
Our replication	-0.013(M)/0.211**(F) (-0.012(M)/-0.11(F)*2)		
Repracing ...	Def. of retire	→	-0.03(M)/0.219***(F)
	Controls	→	0.014(M)/0.082(F)
	Method	→	0.029***(M)/0.042***(F)
	Sample	→	0.01

*1 The red (blue) character indicates the positive (negative) impact.

*2 $\frac{\text{(Coeff.: original paper)} - \text{(Coeff. our replication)}}{\text{(maximum value of index)}}$

Table A.4.3: Depression

	Dave et al. (2008)		Coe and Zamarro (2011)
Estimated result in the original paper	0.1157***	→	-0.0691
Our replication	0.116*** (0.000*2)		
Repracing ...	Def. of retire	→	0.165***
	Controls	→	0.109***
	Method	→	-0.132
	Sample	→	0.143***
	Survey country	→	0.005(EURO-D)

	Coe and Zamarro (2011)		Dave et al. (2008)
Estimated result in the original paper	-0.0691	→	0.1157***
Our replication	0.303 (-0.031*2)		
Repracing ...	Index	→	-0.738
	Def. of retire	→	0.267
	Controls	→	2.857***
	Method	→	0.018
	Sample	→	0.715
	Survey country	→	-0.227

*1 The red (blue) character indicates the positive (negative) impact.

*2 $\frac{\text{(Coeff.: original paper)} - \text{(Coeff. our replication)}}{\text{(maximum value of index)}}$

Table A.4.4: BMI

	Godard (2016)		Johnston and Lee (2009)
Estimated result in the original paper	0.115**	→	0.092
Our replication	0.122** (-0.007*2)		
Repracing ...	Index	→	0.371
	Def. of retire	→	0.122**
	Controls	→	0.077***
	Method	→	0.077
	Sample	→	0.072
	Survey country	→	-0.386

	Johnston and Lee (2009)		Godard (2016)
Estimated result in the original paper	0.092	→	0.115**
Our replication	0.118 (-0.001*2)		
Repracing ...	Index	→	-0.018
	Def. of retire	→	0.118
	Controls	→	-0.798
	Method	→	0.728
	Sample	→	0.235
	Survey country	→	0.291

*1 The red (blue) character indicates the positive (negative) impact.

*2 $\frac{\text{(Coeff.: original paper)} - \text{(Coeff. our replication)}}{\text{(maximum value of index)}}$

A.5 Notes on Replication and Replacement

- Replication 1: In this subsection, we explain the details of replication and replacement procedures. Table A.5.1 shows the table number in the original papers we replicate, the number of samples when we replicate the results, and our comments on the replication. In most cases, we can replicate the results in preceding literature with a number of samples similar to the original number of samples.

- Replication 2: We exclude some control variables when we replicate Neuman (2008) because of data limitation. Neuman (2008) uses detailed regional information and the health status when a respondent is a child. We have generated these variables by using the Cross-Wave: Census Region/Division and Mobility File and Aging Trends and Effects (RELATE) Files. However, when we include these generated variables in the estimated model, the sample size significantly decreases. Therefore, we exclude these variables from the control variables in the Neuman (2008) replication.

Table A.5.1: Notes on Replication

	Table we replicate	**Sample replication** (Original) → (Our replication)	**Comment**
Cognition			
Bonsang et al. (2012)	Table 1	54377 → 55564	
Coe and Zamarro (2011)	Table 6 (Memory)	4928 → 4929	
Self-report of health			
Dave et al. (2008)	Table 2 (Poor health)	NA (not mentioned) → 35594	
Coe and Zamarro (2011)	Table 5 (Bad health)	5282 →5284	
Depression			
Dave et al. (2008)	Table2 (Column 3)	NA (not mentioned) → 28420	
Coe and Zamarro (2011)	Table 5 (Euro-D)	5282 → 5284	
ADL			
Dave et al. (2008)	Table 2 (Column 3)	NA (not mentioned) → 30731	
Neuman (2008)	Table 3	7632 → 7655	We omit some control variables.
Obesity			
Johnston and Lee (2009)	Table 1 (Bandwidth 3)	2877 → 2876	
Godard (2016)	Table 9 (Obese)	3951 → 4059	

Tables A.5.2 and A.5.3 summarize the notes on the replacement procedures by each replacement factor. For example, (Bonsang et al. (2012) → Coe and Zamarro (2011)) describe the comments when we carry out the replacement procedure from Bonsang et al. (2012) to Coe and Zamarro (2011). (Controls) describes the comments when we replace control variables.

Table A.5.2: Notes on Replacement 1

Cognition

Bonsang et al. (2012) → Coe and Zamarro (2011)

(Controls)

•We exclude some control variables Coe and Zamarro (2011) include because the variables are not available in all waves used in Bonsang et al. (2012). The problem is that the sample size significantly decreases when we include these variables.*1

(Sample)

•Coe and Zamarro (2011) use health condition variables to restrict the analyzed samples in the SHARE. Since some of these variables are not available in the HRS, we do not apply the same sample restriction procedure in Coe and Zamarro (2011).

Coe and Zamarro (2011) → Bonsang et al. (2012)

(Method and data)

•Since Coe and Zamarro (2011) use only wave 1 of the SHARE, we cannot directly apply the FE-IV estimation for the analysis framework of Coe and Zamarro (2011). Therefore, we use waves 1 and 2 of the SHARE for FE-IV estimation when replacing the method and the dataset. Additionally, since the answer options are different between the waves 1 and 2 of SHARE, we use the chronic diseases number variable that only counts the diseases asked in both waves for replacement, for which we use the wave 2 SHARE.

Self-report of health

Dave et al. (2008) → Coe and Zamarro (2011)

(Method)

•Since Dave et al. (2008) use the FE estimation, they do not use the IVs. Therefore, when applying IV estimation to Dave et al. (2008), we use the same pensionable ages as Bonsang et al. (2012) for the IVs, because Dave et al. (2008) and Bonsang et al. (2012) analyze the USA and the data collection periods roughly overlap.

•We use age and age squared instead of the age dummy when we use the IV estimation. There is a multicollinearity between the IVs (takes the value 1 after a respondent reaches pensionable age) and the age dummy when applying the IV estimation.

Coe and Zamarro (2011) → Dave et al. (2008)

(Index)

•We use "Poor health" (included in wave 1 and 2) as the index for FE estimation because the European scale of self-report of health is asked only in the SHARE wave 1.

(Method and data)

•We use wave 1 and 2 in the SHARE for FE estimation when replacing the method and data because of the same reason in (method and data) of the cognition section.

(Controls)

•We exclude some control variables that are not asked in the SHARE*2 and the health insurance variable that is asked in only several countries, when replacing the control pattern from Coe and Zamarro (2011) to Dave et al. (2008).

(Data)

•We use "Poor health" in the HRS because the European scale of self-report of health is not asked in the HRS when replacing the dataset from the SHARE to the HRS.

*1 e.g., non-professional activities and physical activities.

*2 e.g., race, religious preference.

Table A.5.3: Notes on Replacement 2

Depression
Dave et al. (2008) → Coe and Zamarro (2011)
(Method)
•The same comments as in Self-report of health apply.
Coe and Zamarro (2011) → Dave et al. (2008)
(Method and data)
•The same comments as in Self-report of health apply.
(Controls)
•The same comments as in Self-report of health apply.
(Data)
•We use the CES-D in the HRS because the EURO-D is not asked in the HRS when replacing the dataset from the SHARE to the HRS.
ADL
Dave et al. (2008) → Neuman (2008)
(Method)
•When applying the estimation method by Neuman (2008), we use the same estimation equation and the IVs as Neuman (2008).

Note

1) We omit the literature on the effect of health on retirement. However, a representative paper is McGarry (2004).

2) Lindeboom and Kerkhofs (2009) also specify a model that addresses work decisions, health and health reporting simultaneously.

3) Iparraguirre (2014) broadly reviews some methodological differences found in the literature including public health literature.

4) Charles (2004) surveyed psychological research both theoretically and empirically.

5) Bound and Waidmann (2007), Coe and Lindeboom (2008), Dave et al. (2008), Neuman (2008), Johnston and Lee (2009), Lee and Smith (2009), Latif (2011), Coe and Zamarro (2011), Behncke (2012), Hernaes et al. (2013), Fonseca et al. (2014) and Insler (2014) are representative papers. Furthermore, recently review papers have been published on the impact of retirement on health in other fields. For example, van der Heide et al. (2013) put retirement in the public health context, whereas Wang and Shi (2014) took up retirement in a psychological context.

6) See the website at (http://hrsonline.isr.umich.edu) for more details on HRS.

7) There are two rounds in the Word Recall tests. In the first round (Immediate Word Recall), there is a test to recall the number of words as much as possible. After a while, the second round starts. In the second round (Delayed Word Recall), a respondent is asked to recall the same words as much as possible.

8) "Bad health" is a dummy variable that takes the value 1 if respondents assess their health as fair, bad, and very bad,

and 0 otherwise.

[9]The Gateway to Global Aging Data (http://gateway.usc.edu) provides harmonized versions of data from the international ageing and retirement studies (e.g., HRS, ELSA, SHARE, KLoSA). All variables of each dataset have the same items and follow the same naming conventions. The harmonized datasets enable researchers to conduct cross-national comparative studies. The program code to generate the harmonized datasets from the original datasets is provided by the Center for Global Ageing Research, USC Davis School of Gerontology and the Center for Economic and Social Research (CESR). This code is used to input some variables, such as measures of assets and income.

[10]All models are estimated using the STATA module xtivreg2. See Schaffer (2010) for further details.

[11]Bonsang et al. (2012), Latif (2013), Zhu and He (2015), Zhu (2016), Zhu (2016) and Godard (2016) exploit a similar identification strategy.

[12]For Germany and Denmark (except for females), we use only the dummy variables (e.g. $\{age_{it} \geq A_i^{eb}\}$).

[13]The full results, including control variables, are available on request.

[14]We also change the index of depression (from CES-D to EURO-D) when we change the surveyed country.