HEDG
HEALTH, ECONOMETRICS AND DATA GROUP
THE UNIVERSITY of York

WP 20/08

The effect of education on health policy reform: Evidence from Japan

Masato Oikawa

May 2020

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

The effect of education on health policy reform: Evidence from Japan *

Masato Oikawa †

April 27, 2020

Abstract

This paper analyzes the effect of education on the outcomes of a reform of the health checkups that occur annually at the workplace in Japan. In April 2008, the annual checkup was redesigned to address new scientific concerns about metabolic syndrome. However, as the checkup is mandatory only for salaried workers, their participation rate is significantly higher than other workers such as the self-employed, and so salaried workers were most affected by the reform. Using this institutional information, a difference-in-difference (DID) estimation was conducted with salaried workers being the treatment group and self-employed workers the control group. We found that the reform caused significant changes in health behaviors and outcomes only among university graduates with a relatively high risk of metabolic syndrome. This more highly educated group increased physical activity, reduced energy intake, and achieved a significant weight loss, reducing BMI to a level that minimizes all-cause mortality among middle-aged Japanese. These results imply that a difference in cognitive functioning or educational success may be a key factor in explaining the heterogeneous response to health policy reforms, suggesting that more clearly articulated recommendations for healthy behaviors are needed in order to improve reform uptake.

JEL Classification Numbers: I12, I18, I20
Keywords: health policy, health checkup, health information, education, obesity, BMI, health investment, difference-in-differences (DID) estimation

*I am indebted to Daiji Kawaguchi for his advice and encouragement, and Hidehiko Ichimura and Haruko Noguchi for their advice and support with data access. I also would like to thank Taiyo Fukai, Takuya Hasebe, Yoko Ibuka, Akira Kawamura, Keisuke Kawata, Shiko Maruyama, Hiroyuki Motegi, Ryo Nakajima, Yoshinori Nishimura, Fumio Ohtake, Ryuichi Tanaka, Takahiro Toriyabe, Masanori Tsuruoka, Shintaro Yamaguchi, Mutlu Yuksel, and seminar participants at the 11th Applied Microeconometrics Conference, the 8th Australasian Workshop on Econometrics and Health Economics, the 4th Conference of the International Association for Applied Econometrics, the 12th Annual Meeting of Japan Health Economics Association, the 2017 Japanese Economic Association Autumn Meeting, the Tokyo Labor Economics Workshop, the Empirical Moral Science Workshop 2019, the 2019 Asian Meeting of the Econometric Society, Yokohama National University, Keio University, and Sophia University for their helpful comments. This work was partially supported by the 2017 IAAE travel grant for graduate students and a JSPS Grant-in-aid for Scientific Research [Grant Number:JP17J07578 and 19J01042]. I am grateful for the editing assistance by Philip C. MacLellan. All remaining errors are my own.

†Faculty of Political Science and Economics, Waseda University, Japan, and Japan Society of the Promotion of Science, Japan Email: masato.oikawa1991@gmail.com

1 Introduction

Over the last few decades, economists have become interested in the effect of education on both economic and non-economic outcomes, and one issue attracting much attention is the relationship between education and health. Since Grossman (1972), the relationship has been actively investigated both theoretically and empirically, with the literature reviewed by Grossman (2006) and updated by Eide and Showalter (2011) and Grossman (2015). While most empirical studies to date focus on the causal effect of education on health, less attention has been paid to the specific underlying mechanism, which Grossman (2015) points out is an area of future work that has important implications for effective health policy implementation.

One possible mechanism is that more highly educated individuals might respond to new information and change their behaviors more quickly. In other words, individuals for whom education has been a successful endeavor may be more “teachable”. This possibility has received support from a range of studies on the uptake of newly-approved drugs (Lleras-Muney and Lichtenberg, 2002), a national information campaign on the HIV/AIDS epidemic in Uganda (de Walque, 2007), and several studies of the growing awareness since the 1950s of the negative effects of smoking, as reported in the medical research literature (Price and Simon, 2009), the popular press (de Walque, 2010), and the 1964 U.S. Surgeon General’s Report on Smoking and Health (Aizer and Stroud, 2010; de Walque, 2010).

In all of these studies, however, the interval between the receipt of information and subsequent behavior is long, which makes it difficult to determine the specific mechanism in detail. For example, there are at least two possible paths by which level of education could lead to a heterogeneous response to new health information, each with very different implications for policy. One is that more highly educated individuals might have better access to new health information, but another possibility is that they might respond more quickly or efficiently to new health information even when access is equal. If more highly educated individuals have better access to new health information, then it is important to find ways to disseminate this information more widely, but if access to information is not the problem, then it is important to devise more effective ways to encourage more widespread behavioral change. This might include, for example, more understandable health care campaigns with practical recommendations, specific implications and concrete examples.

Another limitation of the literature is that most studies are not quasi-experimental in that they compare changes in outcome variables among subjects who are either more or less educated without defining clear treatment and control groups. Therefore, while the heterogenous responses to new information received by these two groups may seem intuitive, the studies cannot cannot claim direct causal effects. By contrast, this paper uses a change in the health checkup system in Japan aimed at reducing obesity as the clear mechanism underlying the causal relationship between education and health. Because all workers were aware of the reform and it did not alter participation in the checkup itself, this study discusses the latter possible path by which differences in education leads to heterogeneous responses to policy reform. Further, we use another institutional setting, worker type, to identify the policy effects by comparing the changes in health outcomes and behaviors between two groups with higher and lower proportions of members affected by the policy reform.

While economists have studied the effects of an exogenous variation in health checkups on health outcomes and behaviors in Austria (Hackl et al., 2015), China (Zhao et al., 2013), Japan (Iizuka et al., 2017; Inui et al., 2017), and Korea (Kim et al., 2019), to date there is no unified view of the effects. For example, while Hackl et al. (2015) find that the checkup increases medical expenses but does not change health outcomes, Zhao et al. (2013) find that Chinese receiving a hypertension diagnosis do reduce their fat intake. Meanwhile, while Iizuka et al. (2017) show that Japanese increase their medical expenditures and doctor visits after receiving a diabetes diagnosis but find little or no evidence of any change in smoking, eating, drinking or exercising behaviors or in health outcomes, Inui et al. (2017) find that stress and probability of non-smoker is reduced while Kim et al. (2019) finds evidence of an increase in medication and weight loss. In sum, evidence of the effect of health checkups is equivocal, as Inui et al. (2017) and Kim et al. (2019) find positive health effects but Hackl et al. (2015) and Iizuka et al. (2017) find no such evidence.

One possible reason for the mixed results is the estimation methodology used in previous studies. Most of the literature has applied a regression discontinuity design (RDD) with a biomarker threshold for diagnosing a health condition such as high blood pressure(Zhao et al., 2013), diabetes (Iizuka et al., 2017; Kim et al., 2019), or obesity (Kim et al., 2019). Since biomarkers are affected by various exogenous factors such as timing, measurements just above and below a threshold are likely to be random and so the effect of a diagnosis on subsequent health outcomes and behaviors is estimated around that threshold. Although this strategy has strong identification power, the estimated effects appear to be highly localized. Further, these studies have utilized a variety of thresholds for the identification strategy and moreover, have found heterogenous effects based on numerous factors including level of income(Zhao et al., 2013), age (Hackl et al., 2015), and level of health risk (Iizuka et al., 2017; Kim et al., 2019). Considering this, as it is likely that these studies have investigated different estimands, making the estimated results not necessarily comparable, further investigation of the effects of health checkups is necessary to gain consensus.

Additionally, there are few studies of health diagnoses focusing on heterogeneity by level of education. One study by Zhao et al. (2013) utilizes an RDD framework to analyze the heterogeneous effects of a hypertension diagnosis on nutrition intake by both education and income, finding that the effect on fat intake is stronger among those with a lower education and higher income, which is inconsistent with the literature on the relation between education and health. One possible reason is that the more highly educated might know the diagnosis threshold, and so even if their blood pressure rating is just below the threshold, they may change their behavior because they know that they are still high risk. Therefore, this strategy does not seem ideal for analyzing the heterogeneity of effects by level of education,[1] leaving ample room for more studies of heterogeneous effects of health checkups by level of education.

This paper utilizes a reform of the system of health checkups in Japan to analyze the heterogeneous effects of education level on the health behaviors and outcomes of Japanese workers. Following the reform, health checkups focused on fighting obesity by providing participants with an objective evaluation of risk factors, followed by guidance by health professionals. Recognizing that the reform did not affect the participation rate in checkups but that the proportion of workers affected by the reform differed exogenously according to their work status, the study uses a difference-in-differences (DID) approach to compare the changes in the outcome variables between two groups with higher and lower proportions of members affected by the policy reform.

The results show that while the DID estimates of weight and body mass index (BMI) are statistically significantly negative for university graduates with a higher risk of obesity, there are no significant changes among university non-graduates or individuals with a relatively low risk of obesity. Further, the differences in the DID estimates between university graduates and non-graduates are statistically significant, and among university graduates with a high risk of obesity, health behaviors such as physical activity and eating habits also changed. These results suggest that more highly educated individuals are more likely to respond to a health checkup diagnosis and/or a health guidance to improve their health and, moreover, that cognitive functioning may be a key factor explaining this heterogeneity of response, which is consistent with other recent discussions of the role of cognition in the causal relationship between education and health (e.g., Conti et al., 2010; Bijwaard et al., 2015; Bijwaard and Van Kippersluis, 2016).

The remainder of the paper is organized as follows: section 2 explains the institutional setting and section 3 discusses the data and descriptive statistics. Section 4 describes the

[1] Additionally, the authors note that the sample size of the more highly educated group is relatively small, which may have led to less precise estimates.

identification strategy and estimation model, and section 5 discusses the estimation results. Section 6 provides some additional remarks, and section 7 concludes the paper, with suggestions for further research.

2 Institutional Background

2.1 Japan's Annual Health Checkup

Since the 1970s in Japan, an annual health checkup has been provided to workers as part of the country's health promotion policy. The Industrial Safety and Health Act of 1972 requires all employers to provide a health checkup for their employees, and salaried workers are legally obligated to take it, so virtually all salaried workers in Japan undergo a health checkup each year. Although the checkup potentially includes additional items, a number of tests are required by law, so salaried workers across Japan all receive at least this uniform minimum-level checkup. In addition to the checkup provided at the workplace, local governments also provide checkups for residents over age 40 who are not salaried workers. Thus all middle aged or older Japanese residents have the opportunity to receive the annual health checkup, though it is mandatory for salaried workers and voluntary for others including the self-employed.

However, in the early 2000s, despite these publicly provided checkups having taken place annually for decades, health had not sufficiently improved. Specifically, according to the mid-term evaluation of the "Health Japan 21" promotion policy implemented in 2000 at the turn of the 21st century, the incidence of lifestyle-related health conditions such as diabetes and obesity had increased. As these health conditions comprise a large proportion of public health expenditures, the problems with the health checkups were investigated and summarized in a Council of Governments report. [2] The first problem identified was that intervention does

[2] Further information (in Japanese) can be found at:

not work for people who already have a disease. While screening is most effective for people with high risk of a disease but who do not yet have it, [3] the purpose of the checkups to that point was to select for intervention those patients who were already in the early stages of a disease. The second problem with the health checkups was that the intervention was insufficient. While Knowler et al. (2002), for example, find that for those with high risk of diabetes, intervention to help change lifestyle habits is more effective in preventing the onset of diabetes than medication, the health checkups merely provided those identified as high risk with general information about the disease and a recommendation to see a doctor. The final problem was that the content of the health checkups, conducted by providers across the country under various local laws, was not unified. Addressing these identified inadequacies of the existing health checkups thus required a reform of the system to provide a more substantial intervention targeting those at high risk of disease and implemented uniformly across all institutions nationwide. This new system is described next.

2.2 Specific Health Checkups and Specific Health Guidance

In April 2008, a new health checkup system, the Specific Health Checkups and Specific Health Guidance, was introduced, aimed at preventing lifestyle-related diseases by providing participants objective assessments of their health risks and specific guidance by health professionals. This new health checkup system now focuses on metabolic syndrome, a condition represented by a confluence of biomarkers including excess body fat, high blood pressure and high blood sugar which together identify people at high risk of lifestyle-related diseases. The policy reform was introduced uniformly for the target population of individuals covered by public health insurance and their dependents aged between 40 and 74. Since Japan has a

---

`https://www.wam.go.jp/wamappl/bb14GS50.nsf/vAdmPBigcategory40/98E6F3F836572E8B4925716F0006B833?OpenDocument` and `https://www.mhlw.go.jp/shingi/2005/09/s0915-8.html`.

[3] The Schellenberg et al. (2013) systematic review of diabetes, for example, finds no evidence that intervention is effective for those who already have type 2 diabetes but is effective for those at high risk of getting it.

universal health care insurance system, this target population covers almost all residents of Japan. Importantly for this study, the reform did not alter participation in health checkups. As seen in Panel (a) of Figure 1 [4] , there is no jump in the checkup participation rate of the middle-aged around the policy reform.

The current system is divided into two parts: a health checkup to screen for participants at high risk of metabolic syndrome, followed by face-to-face guidance by a doctor or dietitian aimed at prevention by changing lifestyle habits. The content of the checkup is based on medical and scientific evidence for identifying metabolic syndrome, and includes body measurements, blood tests, and questionnaires about such topics as smoking and medication histories. As excess body fat is a marker of metabolic syndrome, a measure of abdominal girth was added to the new system to estimate the amount of visceral fat. Based on the results of the health checkup, an objective assessment of metabolic syndrome risk is determined.

Participants then receive health guidance specifically tailored to their physical condition. Those at high risk are given health guidance about their lifestyle habits, aimed at informing participants of the benefits and risks of their lifestyle habits and providing support to change behavior.

To sum up, the policy reform provides checkup participants with objective knowledge of the risk associated with their health condition and specific information about the benefits and risks of their health behaviors.

3 Data

The main dataset used in this study is the Japanese Study of Aging and Retirement (JSTAR), a biennial panel survey of elderly Japanese aged over 50 that is a sister dataset of the U.S. Health and Retirement Study (HRS), the English Longitudinal Survey on Age-

[4] I constructed Figure 1 based on information in Kawamura et al. (2019).

ing (ELSA), and the Survey on Health, Ageing, and Retirement in Europe (SHARE). The first wave of JSTAR was conducted in 2007 at five cities in Japan. Although other sample cities were added in later waves, this study employs only the five original cities in order to obtain data both before and after the health policy reform of 2008. The JSTAR includes comprehensive information on demographics, labor force status, economic variables, health investment behaviors, and health outcomes for analyzing the impact of the checkup system reform on health outcomes and behaviors. [5] The purpose of the policy reform was to reduce the number of people at high risk of metabolic syndrome, and one solution is weight loss. This is captured in this study through measures of weight and of body mass index (BMI), or body weight adjusted by height, a common measure of obesity.

In addition to JSTAR, the Comprehensive Survey of Living Conditions, the Longitudinal Survey of Middle-aged and Elderly Persons, and the General Survey on Working Conditions are also used in this study to interpret and provide further context for the main estimation results.

4 Identification Strategy

This paper utilizes heterogeneity in the health checkup participation rate to identify the effects of the policy reform on health outcomes and behaviors. As noted above, the new checkup system was implemented uniformly, making it difficult to assign participants to treatment and control groups on that basis, so a different institutional setting was used based on employment status. As explained above, the annual health checkup has always been mandatory for salaried workers, so their participation rate is higher than others. Panel (b) of Figure 1 shows the participation rate of middle-aged workers according to employment status and indicates that the participation rate of salaried workers is about 90% and constant before

[5] Please see Ichimura et al. (2009) and `http://www.rieti.go.jp/en/projects/jstar/index.html` for more detailed information about the JSTAR.

and after the policy reform. Therefore, about 90% of salaried workers were affected by the policy reform but their participation in health checkups did not change. On the other hand, the participation rate of self-employed workers is substantially lower, at about 50%, and also showing no significant change in participation around the time of the policy reform. We can therefore conclude that about 50% of self-employed workers were affected by the reform but their participation in health checkups also did not change.

We use this stable difference in the proportion of workers affected by the policy reform is used as the identification strategy, applying a difference-in-difference (DID) framework to compare the before-after change in health outcomes between the salaried worker treatment group with a higher proportion of members affected by the reform and the self-employed worker control group with a lower proportion of members affected. [6] Within this DID approach, the estimated effects are deducted by the difference in the participation rate between salaried and self-employed workers, so if the signs of both groups are the same, the DID estimate indicates the lower bound of the magnitude of the effect in absolute value. The interpretation of the DID estimate is discussed further in Section 6.

In order to ensure the validity of the DID approach, a number of assumptions must hold, and these are discussed here. Firstly, it is important for the identification strategy that the relative participation in health checkups does not change between the treatment and control groups. As the checkup has always been mandatory for salaried workers, one would not expect their participation to change, and this has been confirmed above. For self-employed workers, however, the checkup is voluntary and so it is conceivable that their participation in the improved checkups might increase. If their participation rate were to rise to something approximating that of salaried workers, the shrinking difference in participation rates would reduce the effectiveness of the DID approach. However, as discussed above, the participation rate did not change significantly after policy reform for either salaried or self-

[6] Non-working individuals who are retired or disabled in 2007 are not included in the control group because they may have different trends compared to those who are working.

employed workers (Panel (b) of Figure 1).

Secondly, an important assumption for the internal validity of DID is the common trend assumption; namely, that the counterfactual change in outcomes among salaried and self-employed workers must have been the same if policy reform had not occurred. This allows us to measure the effect of the reform. The typical means of testing for this is to check the trends in target outcomes before the reform, but as our dataset includes only one period before the policy reform, we must test this via other means. First, as JSTAR asks respondents to self-report the change in their health from one year before the survey to the survey date, we use this information to assess whether the trends of salaried and self-employed workers are heterogenous. Second, as the dataset is comprised of longitudinal data that includes rich information about demographic, economic, and health related variables, we can control for individual observable characteristics related to health such as age and economic condition, as well as time-invariant individual heterogeneity. Third, in order to provide further evidence of a common trend, a placebo regression using a health variable less related to the newly introduced system was estimated. While the details of these validity checks are discussed below, at this point we can state that there is substantial evidence that the common trend assumption holds.

4.1 Estimation Model

Controlling for any potential bias caused by heterogeneity in the trends of salaried and self-employed workers, the estimation equation is as follows:

$$\begin{aligned} y_{it} &= \beta_0 + \beta_1 SalariedWork07_i + \beta_2 After_t \\ &\quad + \beta_3 SalariedWork07_i \cdot After_t + x'_{it}\delta + \theta_i + \epsilon_{it} \end{aligned} \quad (1)$$

where i and t are indices of individual and time. The dependent variable y_{it} represents health outcomes such as weight and BMI, and health investment behaviors such as physical activity and eating habits. $SalariedWork07_i$ takes one if the respondent is a salaried worker in 2007 before the policy reform, while $After_t$ takes one after the policy reform. The vector x_{it} is a set of control variables that includes age dummy variables, marital status, number of children, household income, house ownership, hours of work, stress condition at e thworkplace, occupation dummy variables, and prefecture-level macroeconomic variables. The parameter θ_i captures the unobserved individual fixed effects and the parameter ϵ_{it} is an unobserved error term. In Equation (1), the parameter β_3 corresponds to the DID estimate and is the parameter of interest in this paper. This captures the difference in the change in the outcome variable between salaried and self-employed workers in 2007. Equation (1) is estimated for both university graduates and non-graduates and the DID estimates are compared in order to discover the heterogeneous effects of education on the outcomes of policy reform.

4.2 Verification of Common Trend Assumption

As discussed above, any causal interpretation of the DID estimate requires the common trend assumption to hold. Although our dataset includes only one period before the reform, we can use the 2007 self-reported change in health from the previous year to check for any heterogeneity in the trends. Figure 3 shows the self-reported change in health from 2006 to 2007 by level of education and employment status. [7] We see that the patterns are not significantly different between salaried and self-employed workers or university graduates and non-graduates, with the proportion answering "same" about 85% for all groups. This suggests that the common trend assumption between salaried and self-employed workers is

[7] The 2007 JSTAR asks following question: "How is your current health compared to one year ago?" (question number: D-003) The options are "much better", "better", "same", "worse", "much worse." In the figure, "better" includes both "much better" and "better" responses to the original question, while "worse" includes both "much worse" and "worse".

satisfied at least by what can be captured by self-reported changes in health.

Next, as an additional measure to control for any heterogeneity in the trends between salaried and self-employed workers, observable characteristics were added to the model and a fixed effects estimation was conducted using the panel structure of the JSTAR. In addition to the demographic variables of age, marital status, and number of children, household income and home ownership economic variables were included as control variables because previous studies have shown a relationship between health and economic conditions (Case et al., 2002; Chetty et al., 2016; Semyonov et al., 2013). Further, because the analysis sample includes the JSTAR data during the financial crisis of 2008 which may have affected workplace and regional economic conditions heterogeneously, workplace-related variables (hours worked, physical stress at workplace, job stress at workplace, occupation category dummy variables), and time-variant regional characteristics (prefecture-level GDP and per capita income macroeconomic variables) were also included as controls. Additionally, although the accumulation of health stock until middle age and health preferences could also cause heterogeneity in the trends, the fixed effects estimation controls for such time-invariant unobserved individual heterogeneity.

As a final measure to ensure that the common trend assumption is satisfied, a placebo regression was run using height as the dependent variable. According to previous studies, significant height shrinkage after middle age is a proxy for health decline. As policy intervention to encourage weight loss would appear to have little relation to height, any health shock experienced differently by the treatment and control groups would be captured by this placebo regression.

4.3 Analysis Sample Restrictions

For the estimation, the analysis sample was first restricted to males aged between 50 and 62 because the identification strategy requires workers and in Japan, males are more likely to be working. Next, although the newly introduced health checkup system was made available to everyone aged 40 and over, a minimum age of 50 was chosen for this study because the JSTAR data included only those aged 50 and over. Meanwhile, the maximum age was restricted to 62 in order to eliminate from the sample those workers who were eligible to retire with a pension, as previous studies have found that this affects health conditions and behaviors (e.g., Kajitani, 2011; Motegi et al., 2016, 2020; Nishimura et al., 2018; Zhao et al., 2017). For the first wave of the JSTAR, 62 was age that people get the eligibility of full amount of pension.

Next, the sample was divided into two groups according to their risk of metabolic syndrome before the policy reform because, as mentioned above, individuals with a high risk received not only the objective results of the checkup but also guidance from a health expert. Moreover, low risk individuals do not need to change their behaviors because they are already healthy. For these reasons, there is a possibility that the reform effects might be heterogeneous according to the individual's potential risk of metabolic syndrome.

BMI was used as the criterion to divide the sample by health condition before the reform. As explained above, a range of measurements are used to evaluate the risk of metabolic syndrome, but as the 2007 JSTAR includes only BMI information, it is difficult from that alone to construct a full picture of the pre-reform risk of metabolic syndrome. Fortunately, however, the 2009 JSTAR also includes information on girth of abdomen and blood pressure that can be used to gauge an individual's eligibility for guidance. [8] A plot of these two pieces of information, eligibility for guidance and BMI, was used to determine the BMI criterion for dividing the sample into high and low risk. From Figure 2, we see that the probability of receiving guidance increases as BMI rises, but there is a jump of about 30 percentage points when BMI reaches 23.5, from 40 % at BMI=23 to 70 % at BMI=23.5. Thus BMI greater than 23.5 became the criterion for defining the high-risk portion of the sample, with others

[8] The evaluation derived here is only an approximation because of a lack of information about blood glucose and triglyceride levels which are also risk factors.

at low-risk of metabolic syndrome. Because those at high risk have an incentive to change their behavior and to reduce their obesity, the policy reform effects should be stronger for these individuals, and so the analysis sample was further restricted to them.

5 Estimation Results

5.1 Descriptive Statistics

Before presenting the estimation results, this section discusses relevant descriptive statistics of the sample. First, looking at males between 50 and 62 with a high risk of metabolic syndrome before the policy reform, Table 1 shows the average values of observable characteristics and the difference in those characteristics between salaried and self-employed workers by level of education in 2007. According to Table 1, after conditioning for health risk before the reform, most of the characteristics are not statistically different between salaried and self-employed workers either for university graduates or university non-graduates (Columns (3) and (6)). Average BMI is similar for all groups, ranging from 25.66 to 25.93. Therefore, while we were concerned that if more obese individuals decreased their weight more, this would lead to heterogeneity in the effects of the policy reform on BMI, this concern appears unfounded because BMI is similar across groups. We also notice that salaried workers with a university degree appear to be taller, but the difference is small (about 1.8 %). Next looking at workplace conditions, we notice that among those with less education, the self-employed work about 10% more hours than salaried workers and are more physically stressed. This means that salaried workers enjoy more leisure time, which can be used to invest in positive health behaviors such as exercise. Additionally, the difference in physical effort required at the workplace may affect weight loss. Therefore, because these two differences can cause the estimates to be biased, it is important to control for these workplace conditions, especially for university non-graduates.

Before presenting the estimation results for Equation (1), we first discuss the DID estimates intuitively via Figure 4, which summarizes the changes in average BMI among individuals at high risk of metabolic syndrome before the policy reform by their employment status and level of education. Panels (a) and (b) show the changes for university graduates and non-graduates, recalling that average BMI was approximately the same for all groups in 2007. From Figure 4, we see that for salaried workers with a university degree, average BMI falls bellow 25 in 2009 and has remained at that level ever since (solid line in Panel (a)), while all other groups show no decrease in average BMI (dashed line in Panel (a) and lines in Panel (b)). This suggests that a goal of the reform to reduce obesity occurred only for salaried workers, for whom a larger proportion of members were affected by the reform, and only if they had a higher level of education.

While this result does not completely rule out the possibility of bias arising from the common trend assumption failing to hold, in the estimations discussed below, an attempt is made to minimize any such bias by controlling for observable characteristics and time-invariant individual heterogeneity.

5.2 Effect on Health Outcomes

This section discusses the estimation results for the effect of the reform on the body measurements of BMI, weight, and height. Table 2 shows the estimation results for those with higher risk of obesity before the policy reform by level of education, with Columns (1), (2), and (3) university graduates and Columns (4), (5), and (6) university non-graduates. Fixed effects (FE) estimation was used in all specifications to control for individual time-invariant heterogeneity, and the DID estimate, β_3 of Equation (1), is reported.

According to Table 2, the DID estimate of BMI for university graduates with a higher risk of obesity before the policy reform is -1.065 and statistically significant at the 1% level after controlling for observable characteristics and unobserved time-invariant heterogeneity (Column (1)). Since average BMI for salaried workers with a university degree before the policy reform was 25.78, this means that their BMI decreased by about 4.1% to about 24.72. A medical study by Tsugane et al. (2002) analyzing the relationship between BMI and all-cause mortality for middle aged Japanese finds that the mortality profile for males has a U-shape and bottoms out at a BMI range of 23.0-24.9. Consequently, we can interpret the observed reduction of BMI among university graduates to within this range as an improvement in their health condition.

Additionally, among university graduates, we find a statistically significant 4.0% weight loss (Column (2)) but no significant change in height (Column (3)), suggesting that the decline in BMI is associated with the weight loss. From this we can further infer, as discussed earlier, that the lack of significant height shrinkage suggests that the BMI and weight estimates are less likely to have suffered from bias due to heterogeneous health shocks. In contrast to university graduates, non-graduates show no statistically significant changes in any of the body measurements (Columns (4), (5), and (6)), indicating that health improvement with weight loss is observed only among university graduates with higher risk of obesity. [9] From this, it appears that the policy reform was effective only for individuals with a higher level of education.

To round out the analysis, for individuals with lower risk of obesity before the policy reform, we found no statistically significant changes in any of the body measurements (Table A.4) for either university graduates or non-graduates. This is not surprising, as individuals who are at lower risk of obesity are less likely to have an incentive to change their health condition because they are already healthy, at least by this measure.

[9] The estimation results are robust for other age ranges as well (Tables A.1, A.2, and A.3).

5.3 Effect on Health Investment Behaviors

This section discusses the estimation results for the effect of the reform on health investment behaviors related to obesity, focusing on physical activity, energy intake, drinking habits and eating habits. Table 3 shows the estimation results for university graduates with higher pre-obesity risk. To measure physical activity, a dummy variable was constructed that takes a value of one if the respondent both walks 90 minutes on a normal day and takes part in some form of physical activity on a holiday (Column (1)). Energy intake (Column (2)) was calculated as the difference between actual energy intake and an ideal level which for this study was the estimated energy requirement (EER), the level of energy intake required to maintain current body weight for a low level of physical activity. [10] To achieve weight loss, however, actual energy intake should be lower than the EER for one's current lifestyle, so if one takes part in a moderate level of physical activity, actual energy intake should be lower than the EER for the moderate level. "Reference 1" of Sasaki (2008) classifies sedentary workers with some movement as moderately physically active and workers with physically demanding positions as highly physically active. On this basis, the analysis sample for the current study includes only workers whose level of physical activity is moderate or above. Therefore, as the ideal level of energy intake for weight loss must be lower, the EER for the low level of physical activity is used in this study. Column (3) reports estimates for a drinking dummy variable which takes value 1 if alcohol intake is greater than zero, and Columns (4) to (8) show the results for the daily intake of staple foods, main dishes, meat dishes, fish dishes, and vegetables, where main dish intake is the sum of meat and fish dish intake. Column (9) is the result for energy intake from fish dishes as a proportion of main dishes.

From Table 3, the DID estimate for the physical activity dummy variable is positive and

[10] According to Sasaki (2008), EER is determined by multiplying the basal metabolic rate (BMR) and the physical activity level (PAL). BMR was calculated using the prediction equations developed by Ganpule et al. (2007), with $BMR = 0.1238 + 0.0481 \times Weight(kg) + 0.0234 \times Height(cm) - 0.0138 \times Age - 0.5473$ for males. Based on "Reference 1" of Sasaki (2008), a PAL value of 1.5 was chosen, corresponding to the level for a sedentary lifestyle, and actual energy intake was obtained from the JSTAR nutrition survey.

statistically significant and energy intake is negative and statistically significant (Columns (1) and (2)), indicating that, for university graduates, energy expenditure increases and energy intake falls to the ideal level. Further, among university graduates, salaried workers as compared to self-employed workers stopped drinking after the policy reform at a significance level of 10 % (Column (3)) and also statistically significantly changed their daily intake of meat and fish dishes, but not of staple foods, main dishes or vegetables (Columns (4) to (8)). As the DID estimates for meat and fish dishes have opposite signs (meat is -63.843 and fish is 38.332) and the estimate for the ratio of fish to main dish intake is positive and statistically significant, this suggests that university graduates replaced meat for fish after the health checkup. In contrast, among non-graduates of university, no systematic changes in energy intake was observed while physical activity decreased (Table A.5). Therefore, systematic changes in health behaviors were observed only among university graduates.

To sum up, we found that among university graduates with a higher risk of obesity prior to the health policy reform, the revised health checkup resulted in changes in both health condition and health investment behaviors. BMI decreased to within the range associated with minimal all-cause mortality for middle aged Japanese men, and physical activity and eating habits improved. For all other groups, however, no change was observed in either health outcomes or behaviors.

6 Discussion

6.1 Interpretation of the DID Estimates and Comparison of University Graduates and Non-Graduates

In the previous section, educational heterogeneity in the response to the policy reform is analyzed by comparing the DID estimates of university graduates and non-graduates. This section discusses three possible situations complicating this comparison of the different effects between university graduates and non-graduates even when the common trend assumption between the treatment and control groups (salaried and self-employed workers) is satisfied. These are 1) when the participation rate of salaried and self-employed workers is different, which is likely because this difference is important for our identification strategy, 2) when the health condition of salaried and self-employed workers is different prior to the reform, and 3) when employment status might be related to company size.

First, as discussed in Section 4, the DID estimate is interpreted as the lower bound of the magnitude of the effect of the policy reform in absolute value. Specifically, the estimate corresponds to the effect of the policy reform deducted by the difference in health checkup participation rates between salaried and self-employed workers. Therefore, it is difficult to compare the DID estimates for university graduates and non-graduates when the participation rates of salaried and self-employed workers among these two groups differ. To see this, suppose that among university non-graduates the health checkup participation rate for salaried and self-employed workers is similar so that the difference in participation rate is close to zero. Suppose however that there is a difference among university graduates. In such a situation, the DID estimate for university non-graduates would be smaller in absolute value than that for university graduates even if the magnitude of the effect of the policy reform is the same for both the university graduates and non-graduates. This occurs because there is no difference in participation rate, which is the source of the difference in treatment intensity among the treatment and control groups, only for university non-graduates.

To investigate this possibility, Table 4 shows the difference in health checkup participation rates by level of education using two additional datasets: the Comprehensive Survey of Living Conditions (Panel A) and the Longitudinal Survey of Middle-aged and Elderly Persons (Panel B). [11] Columns (1) and (2) show the participation rates of salaried and self-employed

[11] I constructed Table 4 based on information in Kawamura et al. (2019).

workers, and Column (3) shows the difference in their participation rates. The three rows of each panel show the results for the whole sample, university graduates, and university non-graduates. Considering both panels, the difference in the participation rate ranges from 29 to 38 percentage points, with the difference for university graduates larger than for non-graduates (0.35 vs 0.29, and 0.38 vs 0.34), a single-digit percentage point difference. However, as seen in Table 2, the DID estimate of BMI for university graduates is several orders of magnitude larger (about 50 times larger) than for non-graduates (Column (1) vs Column (4)). Consequently, we can infer that the difference-in-differences in the checkup participation rates of the two groups is not a major source of the difference in the estimation results.

A second concern in interpreting the DID estimates occurs when there is a difference in the pre-reform health conditions of salaried and self-employed workers because, as explained in Section 2, the newly introduced system is comprised of two programs, health checkup and health guidance, and the specific intervention received depends on the participant's assessed risk of metabolic syndrome. If salaried workers were less healthy than self-employed workers before the policy reform, they would have been more likely to receive health guidance than self-employed workers, causing the intensity of the effect of the policy reform to be larger among salaried workers. This would be problematic because it would conflate the DID estimates, with the difference in the DID estimates of university graduates and non-graduates perhaps only capturing the different health tendencies. For this reason, before conducting the estimation procedure, the sample was divided according to pre-reform health condition in order to address this issue.

A third issue regarding the interpretation of the estimation results involves a possible conflation of employment status and workplace conditions. Recall that we found that only for university graduates did salaried workers significantly change their health conditions relative to self-employed workers. In Japan, salaried workers with a university degree are likely to work at large companies, which often provide more fringe benefits than smaller companies, and this could include extra health checkup items beyond those items required by law. If this were the case, then the significant effects observed for salaried university graduates might merely capture large company trends, so we want to eliminate this possibility. According to the 2007 General Survey on Working Conditions, the ratio of firms providing extra health checkup items in addition to those required by law was almost same regardless of firm size (Figure 5). Similarly, the 2011 report shows that the cost to firms of fringe benefits not obligated by law related to the health checkup is almost the same for all firm size categories ("health checkup" in Figure 6). However, the cost for medical and health services among companies with over one thousand employees is about three times more than that among smaller companies ("medical and health services" in Figure 6). [12] In order to address this issue, the model was estimated again including a firm size variable, [13] and the results were found to be robust after adding the firm size variable to the estimation model (Table A.6).

To sum up, we found that the three potential issues raised in this section are not of concern, and so the DID estimates of the university graduates and non-graduates are comparable for discussing the heterogeneity of the response to the policy reform by level of education. The next section discusses possible reasons why the DID estimates are heterogeneous by level of education.

[12] In Japan, a company with over seven hundred employees can create a health insurance society for their employees. In the General Survey on Working Conditions, if a firm has a health insurance society, the firm's cost for medical and health services can include the society's labor administration costs. Therefore, it is possible that any difference in the cost for medical and health services observed between large and smaller companies may at least partially be explained by the cost of administering the health insurance society rather than any difference in direct support for employees.

[13] The interaction term of the after dummy variable and firm size at the JSTAR 1st wave was chosen as the firm size variable to capture firm size trends.

6.2 Causes of Heterogeneity by Level of Education and Policy Implications

This section discusses possible reasons for the observed heterogeneity by level of education in the responses to the health checkup, followed by policy implications. One possible explanation for the heterogeneous response is a difference in cognitive functioning. In recent studies of the relation between education and health, there is much discussion of the relative contribution of variables at early stages of life, including cognitive and non-cognitive skills and socioeconomic background in childhood (Conti et al., 2010; Bijwaard et al., 2015). According to Bijwaard et al. (2015), at least half of the difference in the survival rates between educational groups can be explained by the selection of education choice, based mainly on cognitive skill. While Bijwaard and Van Kippersluis (2016) finds that people with higher education are more efficient producers of hospitalization to survival probability than people with lower education, when cognitive skill is accounted for, the difference disappears.

In order to test this cognitive functioning hypothesis, the JSTAR memory and numeracy skill tests of cognitive functioning were used. Additionally, the discount rate, health preferences and self-reported probability of living at age 80 are used to discuss other possible explanations of the heterogeneous response.

First, according to Table 5, university graduates have statistically significantly higher cognitive functioning scores than university non-graduates on the word recall test of memory functioning and the serial 7s numeracy test (8.0% and 5.3%, respectively). Moreover, the proportion of individuals whose word recall and serial 7s test scores are greater or equal to the 3rd quartile point is higher for university graduates than non-graduates (48.5% higher and 17.5% higher, respectively). In contrast, differences between university graduates and non-graduates on other characteristics such as the discount rate and self-reported probability of living at age 80 are not significant.

Next, in order to decompose the educational heterogeneity of the policy reform response into cognitive functioning and other factors, the university graduate and non-graduate samples were combined and the model was estimated with the interaction terms of the DID term and variables including university dummy variable, health preferences, cognitive functioning test score, discount rate, and self-rated probability of living at age 80. For this estimation, the latter three dummy variables took a value of one if the response was equal to or greater than the 3rd quartile point.

Columns (1)-(3) of Table 6 show the results of the estimation without any interaction terms other than the university dummy variable. We see that the coefficients for the interaction of DID and university ("$\times$ Univ. = 1") on BMI and weight are negative and statistically significant (Columns (1) and (2)), indicating that the DID estimates of BMI and weight are statistically significantly different between university graduates and non-graduates. However, after adding the other interaction terms, the coefficients for the interactions of DID and university on BMI and weight are no longer statistically significant (Columns (4) and (5)). Further, the coefficients for the interaction of DID and cognitive functioning test scores ("$\times$ Serial7 score $\geq$ 3rd quartile" and "$\times$ Word Recall score $\geq$ 3rd quartile") on BMI and weight are statistically significantly negative while all other interaction terms are insignificant. This implies that cognitive functioning is one of the key factors explaining the educational heterogeneity of the response to the health policy reform. [14]

This discussion concludes with a limitation of the estimation related to cognitive functioning. Recall that the identification strategy relies on the difference in the participation rate between salaried and self-employed workers. As discussed above, if the difference in the participation rate were to differ between people with higher and lower cognitive functioning test scores, it would be difficult to interpret the coefficient of the DID and high cognitive test score interaction term. Because the Comprehensive Survey of Living Conditions and the

[14] In addition to health outcomes, the model was also estimated for health investment behaviors, but no systematic tendencies were observed.

Longitudinal Survey of Middle-aged and Elderly Persons do not include information about cognitive abilities, we were not able to conduct a secondary analysis to check this. Therefore, we cannot unequivocally state that the coefficients of these interaction terms represent heterogeneity in the effects rather than heterogeneity in the participation rate of the health checkup.

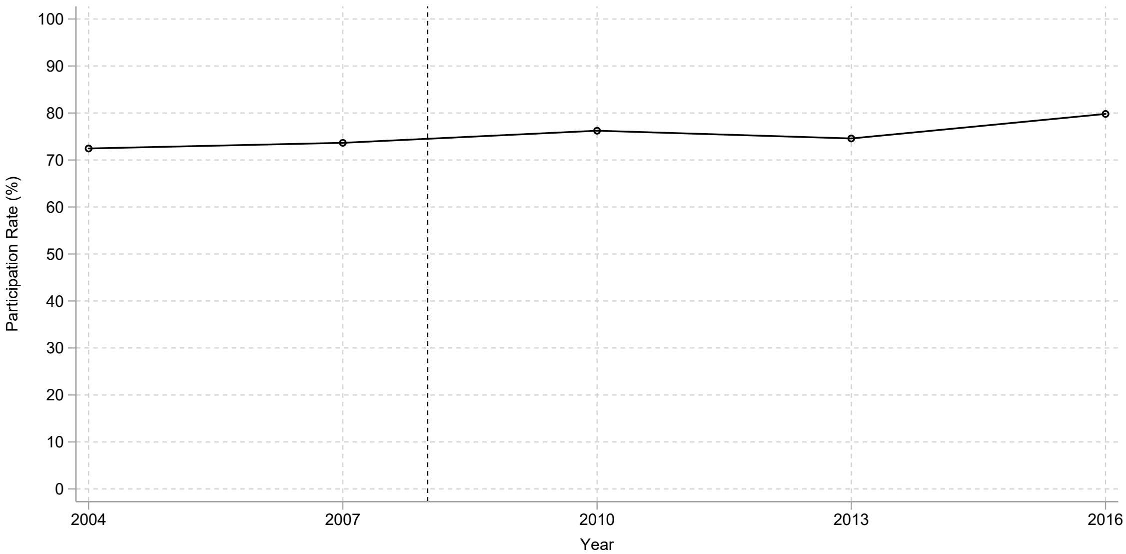
7 Conclusion

This paper analyzes the effects of a policy reform of the health checkup system in Japan aimed at giving participants more objective information about their risk of obesity and health guidance from professionals on the health outcomes and health investment behaviors of participants, finding heterogeneity in the effects by level of education. According to the results of the estimation conducted using a difference-in-differences (DID) framework, the DID estimates of BMI and weight were statistically significantly negative for university graduates with a higher risk of obesity, indicating that the reform of the health checkup produced favorable health outcomes for these participants, who also improved their health investment behaviors including physical activity and eating habits. There were no significant changes observed, however, for university non-graduates or even for graduates with relatively low risk of obesity either in terms of health outcomes or investment behaviors. The differences in the DID estimates between university graduates and non-graduates were statistically significant. These results suggest that more highly educated people are more likely to respond to a health policy reform to change their behavior and improve their health condition.

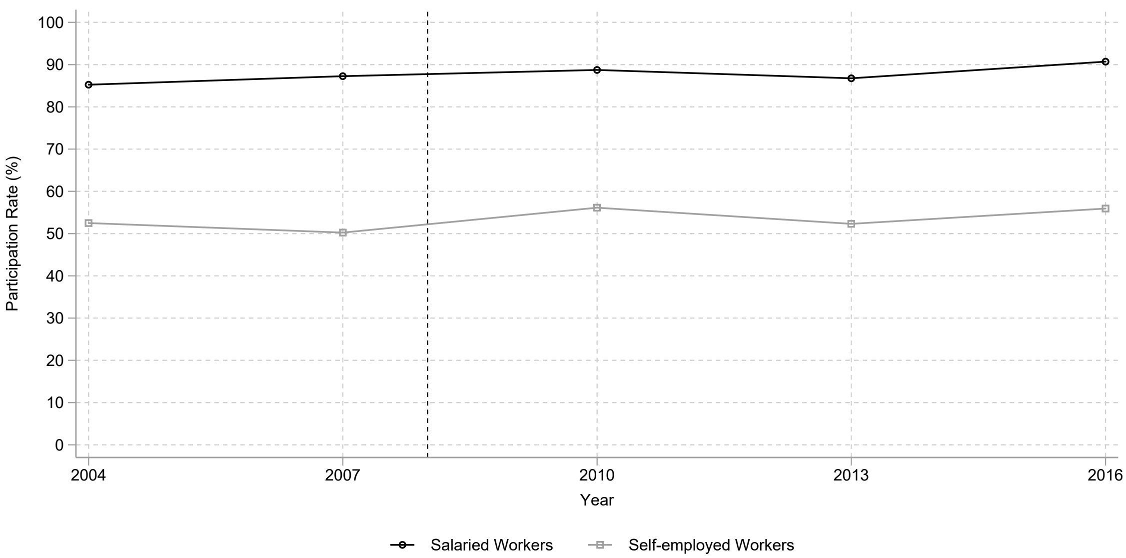
In order to discover why this might be, the study conducted a secondary analysis and found that cognitive functioning appears to be a key factor in explaining the educational heterogeneity of the response to the policy reform. While the descriptive statistics show that the university graduates in the analysis sample have statistically significantly higher cognitive functioning test scores than non-university graduates, a secondary analysis adding the interaction term of DID and high cognitive test scores found this interaction term to be statistically significantly negative. However, the difference in the DID estimates between university graduates and non-graduates was no longer significant, indicating that the reform was effective not strictly for university graduates but only for those graduates with high cognitive functioning.

One possible mechanism is that individuals with high cognitive functioning are more able to efficiently process the information from the checkup and subsequent guidance. The policy implication would thus be that a more effective health checkup with broader uptake and more wide-ranging effects could be provided through a more accessible presentation of health information and guidance that includes a clearly articulated explanation of the effects of specific identified risk factors and directly links these risk factors to a concrete individualized action plan.

Before concluding, there are limitations of the paper to be addressed in the future. Firstly, this paper analyzed holistically the effect of the policy reform of the Japanese checkup system which is now a two-part system of assessment and guidance but did not separate the relative effects of the assessment and guidance. Decomposing the reform effects into these two components is the subject of future work, as it may shed additional light on the observed educational heterogeneity of the reform. Secondly, as the first wave of the JSTAR only included five Japanese cities, we need to pay attention to the external validity of the results. An expansion of this study to a more generalized population is also the subject of future work.



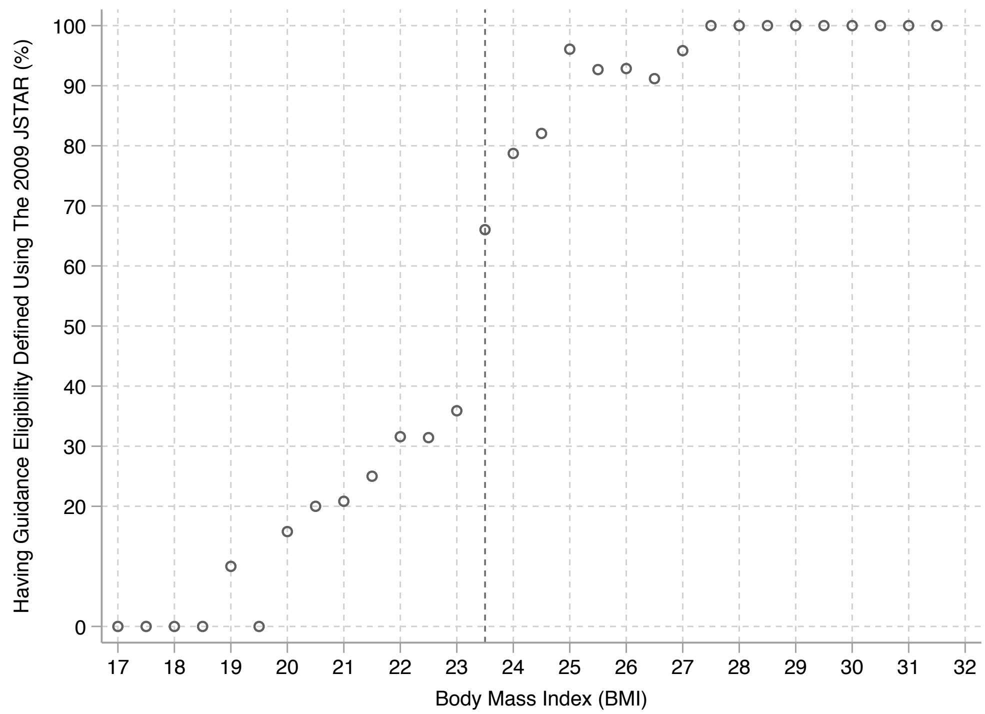
(a) Whole Sample



(b) By Employment Status

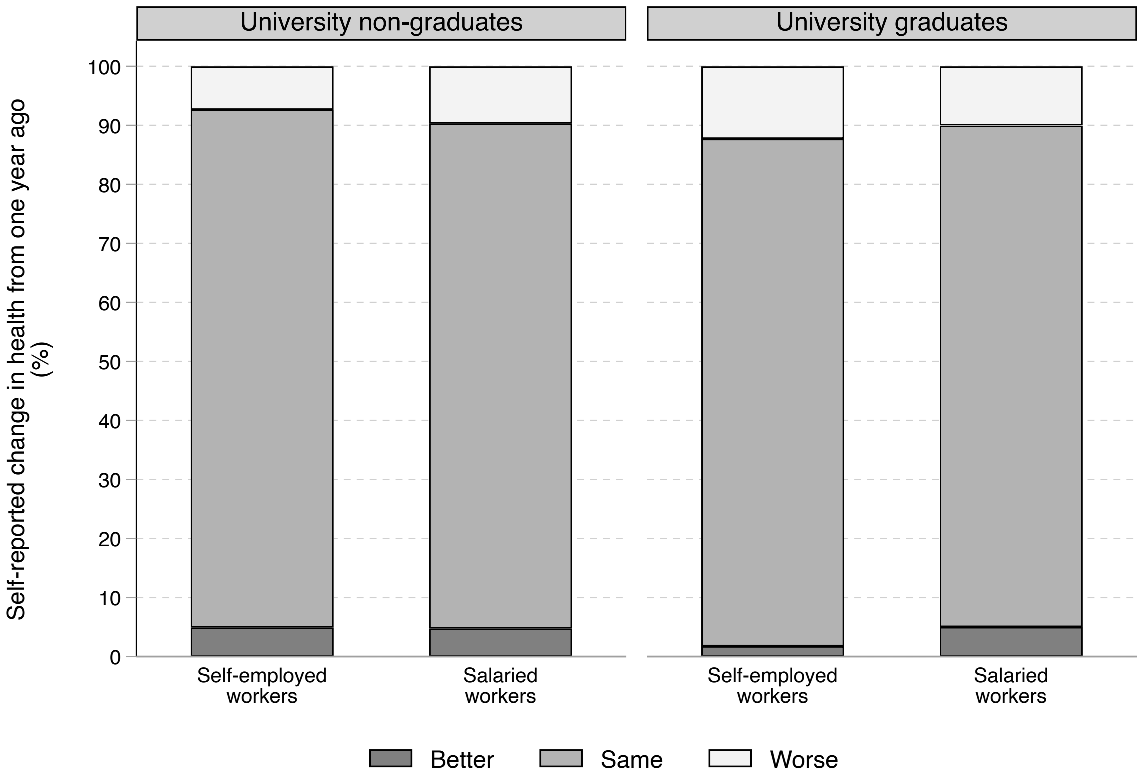
Source: Comprehensive Survey of Living Conditions. Figures are constructed by author based on information in Kawamura et al. (2019). Participation rate is calculated for males aged 50-62.

Figure 1: Participation Rate in Health Checkups of the Middle-Aged



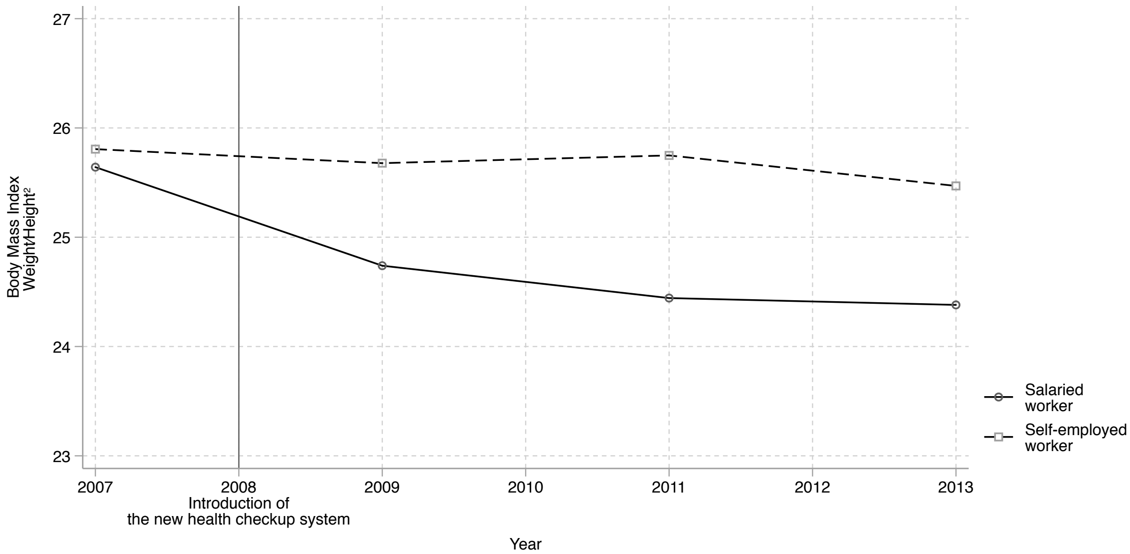
Source: JSTAR 2009. Sample: Males aged over 50.

Figure 2: Relationship between BMI and Eligibility for Health Guidance (2009)

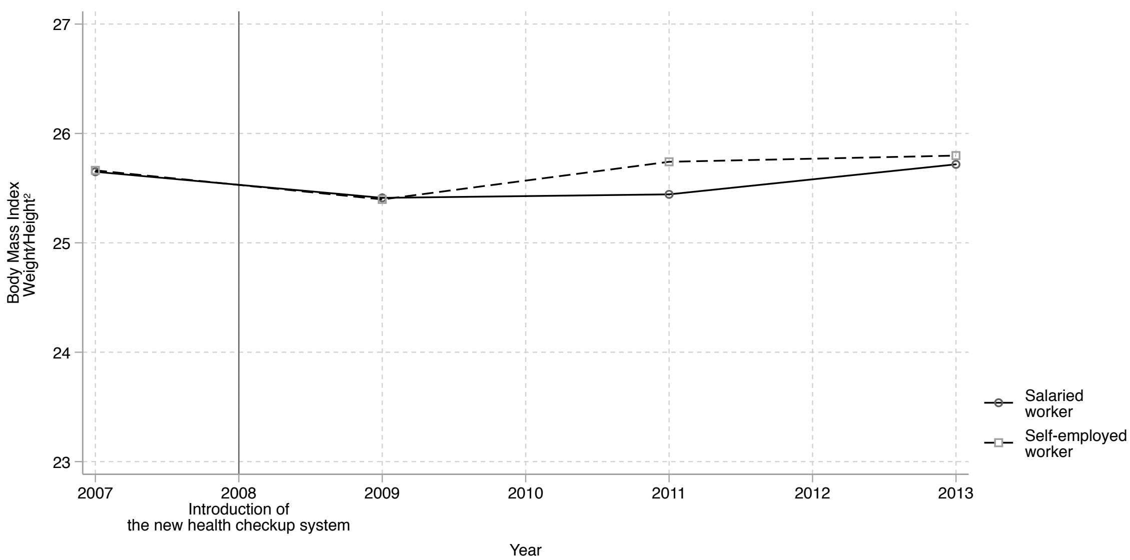


Source: JSTAR 2007. Sample: Males aged 50-62.

Figure 3: Self-Reported Change in Health



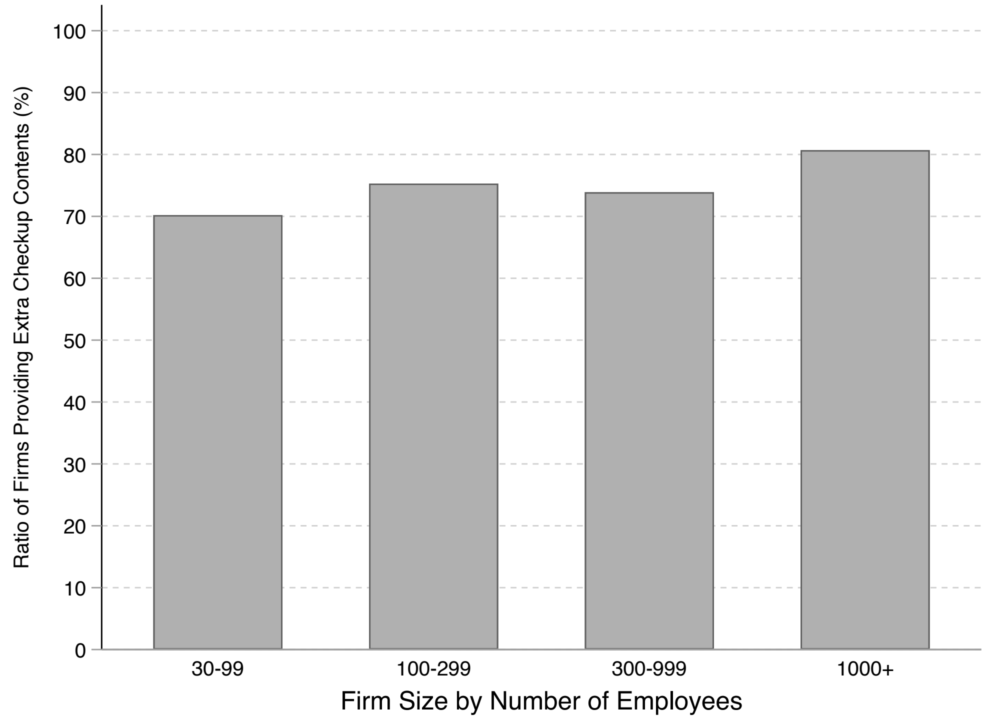
(a) University Graduates



(b) University Non-Graduates

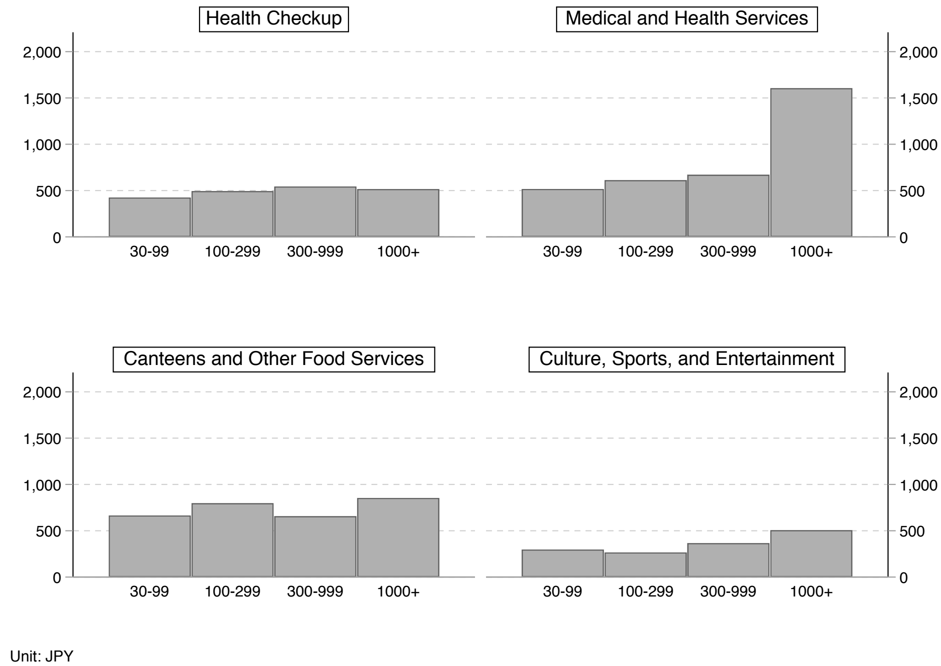
Source: JSTAR. Sample: Males aged 50-62 whose BMI before the policy reform was greater or equal to 23.5.

Figure 4: Changes in BMI among Individuals at High Risk Pre-Period by Employment Status and Level of Education



Source: General Survey on Working Conditions 2007.

Figure 5: Relationship between Firm Size and Provision of Extra Health Checkup Items



Source: General Survey on Working Conditions 2011. Average monthly cost per capita.

Figure 6: Relationship between Firm Size and Benefit Expenses not Required by Law

Table 1: Differences in Characteristics between Salaried and Self-Employed Workers by Education

	University Graduates			University Non-Graduates		
	(1)	(2)	(3)	(4)	(5)	(6)
	Salaried workers	Self-employed workers	Diff.	Salaried workers	Self-employed workers	Diff.
Demographic Variables						
Age	55.27	56.00	-0.73	56.62	56.86	-0.24
= 1 if married	0.92	0.85	0.06	0.86	0.92	-0.06
Number of children	1.80	1.94	-0.15	1.93	2.11	-0.18
Economic Variables						
Household income(10k JPY)	900.07	860.44	39.63	601.42	592.34	9.08
House ownership	0.79	0.74	0.05	0.76	0.84	-0.08
Workplace environments						
Hours worked	47.67	43.97	3.70	45.15	49.66	-4.51***
= 1 if physically stressed	0.20	0.18	0.02	0.42	0.60	-0.18***
= 1 if feeling pressed for time	0.54	0.41	0.12	0.42	0.51	-0.10
Body Measurements						
BMI	25.78	25.93	-0.16	25.68	25.66	0.02
Weight (kg)	74.41	72.59	1.82	72.16	71.67	0.49
Height (m)	1.70	1.67	0.03**	1.68	1.67	0.01
Preference for health						
Be currently interested in own health	0.89	0.85	0.04	0.88	0.81	0.08
Have confidence for own health 3 years later	0.35	0.35	-0.00	0.36	0.31	0.05

1 Source: JSTAR 2007.

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

3 Values are calculated for males aged 50-62 whose BMI before the policy reform was 23.5 or greater.

Table 2: Effects of Policy Reform on Body Measurements by Education (Males Aged 50-62 with High Pre-Obesity Risk)

	University graduates			University non-graduates		
	(1) BMI	(2) Weight (kg)	(3) Height (m)	(4) BMI	(5) Weight (kg)	(6) Height (m)
Salaried before						
Policy Reform × After	-1.065***	-2.987***	0.002	-0.021	0.166	0.003
	(0.321)	(1.071)	(0.004)	(0.287)	(0.740)	(0.002)
Number of observations	202	202	202	480	480	480
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform						
Salaried	25.78	74.41	1.70	25.68	72.16	1.68
Self-employed	25.93	72.59	1.67	25.66	71.67	1.67

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

[2] All specification are estimated using FE model. Clustered robust standard errors are in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table 3: Effects on Health Behaviors (University Graduates with High Pre-Obesity Risk)

				Eating Habits					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Walking and exercise(HD)	Distance btw energy requirement and actual intake	= 1 if drinking alcohol	Staple food (g/d)	Main dishes (g/d)	Meat dishes (g/d)	Fish dishes (g/d)	Vegetables (g/d)	Ratio of fish dishes
Salaried before									
Policy Reform × After	0.277**	-355.942**	-0.223*	-3.290	-25.511	-63.843**	38.332*	39.821	0.103***
	(0.124)	(159.630)	(0.119)	(88.964)	(31.167)	(24.629)	(21.731)	(46.141)	(0.037)
Number of observations	202	202	202	202	202	202	202	202	202
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform									
Salaried	0.14	512.91	0.83	475.64	437.05	270.86	166.19	254.08	0.38
Self-employed	0.06	620.66	0.79	501.78	431.66	263.10	168.56	247.36	0.38

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

[2] All specifications are estimated using FE model. Clustered robust standard errors are in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table 4: Differences in Participation Rate in Health Checkups by Level of Education

Panel A: Comprehensive Survey of Living Conditions	Participation Rate		
	Salaried Workers	Self-employed Workers	Difference
Whole Sample	0.87	0.56	0.31***
University Graduates[2]	0.93	0.58	0.35***
Non-University Graduates[2]	0.85	0.57	0.29***

Panel B: Longitudinal Survey of Middle-aged and Elderly Persons	Participation Rate		
	Salaried Workers	Self-employed Workers	Difference
Whole Sample	0.87	0.51	0.35***
University Graduates	0.91	0.53	0.38***
Non-University Graduates	0.85	0.51	0.34***

[1] Table is constructed by author based on information in Kawamura et al. (2019).

[2] * $p < .1$, ** $p < .05$, *** $p < .01$

[3] As the Comprehensive Survey of Living Conditions asks for level of education only after 2010, data for 2010, 2013, and 2016 are included.

Table 5: Differences in Characteristics of Salaried and Self-Employed Workers by Education

	University Graduates	University Non-Graduates	Difference
Cognitive Functioning Test Score:			
Word recall score (0-20)	11.03	10.21	0.82**
$\geq$ 3rd quartiles	0.48	0.33	0.16***
Serial 7s score (0-5)	4.57	4.34	0.23**
$\geq$ 3rd quartiles	0.74	0.63	0.11**
Discount rate	0.77	0.80	-0.02
$\geq$ 3rd quartiles	0.44	0.44	0.00
Health preference:			
Interested in own health?			
Yes	0.88	0.86	0.02
No	0.01	0.04	-0.03
Confident in own health?			
Yes	0.35	0.34	0.01
No	0.17	0.18	-0.01
Body measurements:			
BMI	25.82	25.68	0.15
Self-reported probability of living at age 80	52.72	50.25	2.48
$\geq$ 3rd quartiles	0.29	0.31	-0.01

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

Table 6: Decomposition

	(1) BMI	(2) Weight (kg)	(3) Height (m)	(4) BMI	(5) Weight (kg)	(6) Height (m)
Salaried before						
Policy Reform × After	-0.020	0.177	0.002	1.314**	3.176*	-0.004
	(0.291)	(0.748)	(0.002)	(0.622)	(1.636)	(0.003)
× Univ.=1	-1.050**	-3.253**	-0.001	-0.475	-1.831	-0.005
	(0.458)	(1.373)	(0.004)	(0.528)	(1.442)	(0.004)
× Serial7 score ≥ 3rd quartile				-1.101**	-2.650*	0.003
				(0.545)	(1.509)	(0.003)
× Word Recall score ≥ 3rd quartile				-1.606***	-4.380***	0.001
				(0.504)	(1.412)	(0.004)
× Discount Rate ≥ 3rd quartile				0.482	1.998	0.009**
				(0.570)	(1.662)	(0.004)
× Not Interested in own health				0.539	1.116	-0.005
				(0.995)	(2.763)	(0.007)
× Not Confident in own health				-0.550	-1.256	0.004
				(0.638)	(1.975)	(0.005)
× Self-report prob. of living at 80 ≥ 3rd quartile				-0.579	-1.582	0.000
				(0.552)	(1.427)	(0.004)
Number of observation	638	638	638	515	515	515

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita, as well as the interaction terms between the control variables and the university dummy variable.

A Appendix

A.1 Screening Procedure for the Specific Health Checkups

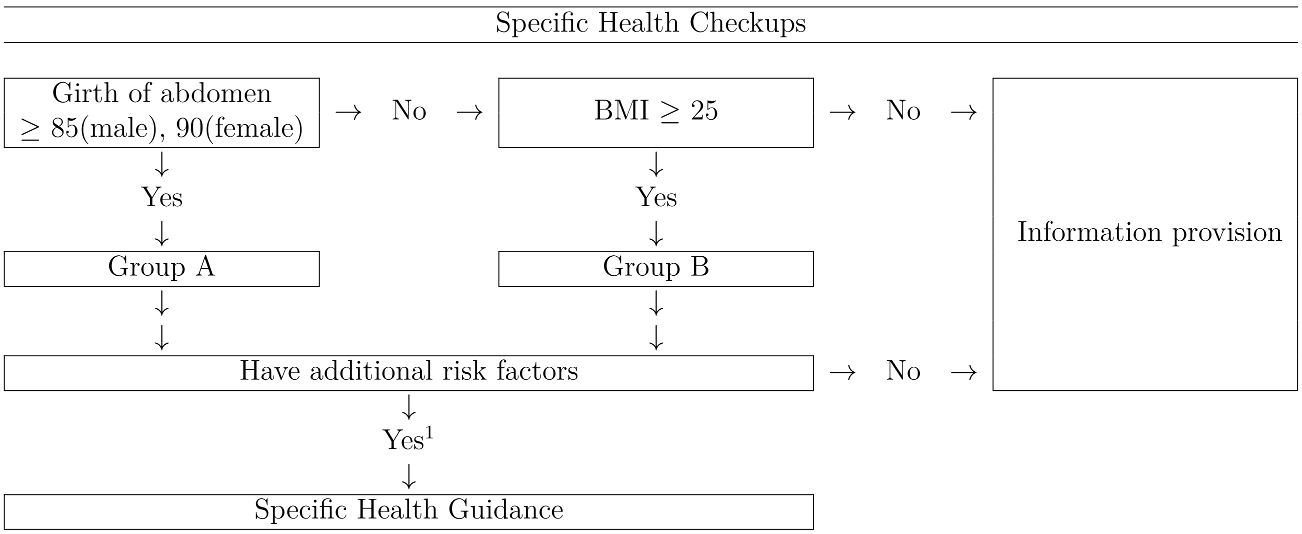
The screening process is divided into two parts: body measurement and additional risk factors. The procedure of the guidance status is as follows:

- First, examinees are divided by their girth of their abdomen, with examinees whose abdomen is over the criteria (male:85cm, female:90cm) assigned to group A.
- Second, those whose abdomen girth is below the criteria (not in group A) but with body mass index (BMI) above 25 are assigned to group B .
- Additionally, the risk level of examinees in group A or B are evaluated by four additional risk factors: high blood sugar, lipid abnormality, high blood pressure and smoking history. [15]
- In group A, examinees with more than two risk factors receive active support guidance, examinees with one risk factor receive motivational support guidance, and examinees without any risk factors are provided information about their health but do not receive any guidance.
- Similarly, in group B, examinees with more than three risk factors receive active support guidance, examinees with one or two risk factors receive motivational support guidance, and examinees without any risk factors are provided information but not guidance.
- Examinees not in group A or B are provided the information about their health but do not receive any guidance.

[15] Smoking history is counted only when examinees also have other risk factors.

The procedure is summarized in Figure A.1. If participants have at least one additional risk factor, then they are categorized as people with high risk. Here participants are able to update their precise risk of metabolic syndrome onset.

Figure A.1: Screening Procedure for Health Guidance



[1] The participants are assigned the level of the guidance depending on the number of additional risk factors. In both groups, examinees with at least one risk factor receive health guidance.

A.2 Other Results

Table A.1: Robustness Check for Age Range (BMI)

	University graduates			University non-graduates		
	(1) Age 50-62	(2) Age 50-61	(3) Age 50-60	(4) Age 50-62	(5) Age 50-61	(6) Age 50-60
Salaried before						
Policy Reform × After	-1.065***	-1.092***	-1.226***	-0.021	0.066	0.188
	(0.321)	(0.328)	(0.371)	(0.287)	(0.303)	(0.334)
Number of observations	202	184	170	480	419	365
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform						
Salaried	25.78	25.77	25.80	25.68	25.63	25.65
Self-employed	25.93	25.97	26.02	25.66	25.74	25.76

[1] * $p < .1$, ** $p < .05$, *** $p < .01$
[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.
[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.2: Robustness Check for Age Range (Weight)

	University graduates			University non-graduates		
	(1) Age 50-62	(2) Age 50-61	(3) Age 50-60	(4) Age 50-62	(5) Age 50-61	(6) Age 50-60
Salaried before						
Policy Reform × After	-2.987***	-3.022***	-3.686***	0.166	0.411	0.667
	(1.071)	(1.061)	(1.223)	(0.740)	(0.791)	(0.884)
Number of observations	202	184	170	480	419	365
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform						
Salaried	74.41	74.48	74.58	72.16	72.11	72.32
Self-employed	72.59	72.88	73.10	71.67	72.02	72.20

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.3: Robustness Check for Age Range (Height)

	University graduates			University non-graduates		
	(1) Age 50-62	(2) Age 50-61	(3) Age 50-60	(4) Age 50-62	(5) Age 50-61	(6) Age 50-60
Salaried before						
Policy Reform × After	0.002	0.002	-0.001	0.003	0.003	0.002
	(0.004)	(0.004)	(0.004)	(0.002)	(0.003)	(0.003)
Number of observations	202	184	170	480	419	365
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform						
Salaried	1.70	1.70	1.70	1.68	1.68	1.68
Self-employed	1.67	1.67	1.68	1.67	1.67	1.67

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.4: Effects of Policy Reform on Body Measurements by Education (Males Aged 50-62 with Low Pre-Obesity Risk)

	University graduates			University non-graduates		
	(1) BMI	(2) Weight (kg)	(3) Height (m)	(4) BMI	(5) Weight (kg)	(6) Height (m)
Salaried before						
Policy Reform × After	0.262	1.378	0.008	0.168	0.422	-0.000
	(0.416)	(1.138)	(0.005)	(0.221)	(0.658)	(0.003)
Number of observations	214	214	214	556	556	556
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform						
Salaried	21.29	60.53	1.69	21.41	59.84	1.67
Self-employed	21.14	60.13	1.69	21.56	59.64	1.66

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.5: Effects on Health Behaviors (University Non-Graduates with High Pre-Obesity Risk)

				Eating Habits					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Walking and exercise(HD)	Distance btw energy requirement and actual intake	= 1 if drinking alcohol	Staple food (g/d)	Main dishes (g/d)	Meat dishes (g/d)	Fish dishes (g/d)	Vegetables (g/d)	Ratio of fish dishes
Salaried before									
Policy Reform × After	-0.168**	151.719	-0.029	36.372	-7.920	7.666	-15.586	8.381	-0.026
	(0.075)	(93.697)	(0.085)	(46.294)	(27.762)	(20.930)	(14.253)	(23.870)	(0.025)
Number of observations	480	480	480	480	480	480	480	480	479
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform									
Salaried	0.12	498.12	0.77	493.38	425.18	255.87	169.30	243.96	0.40
Self-employed	0.10	493.29	0.70	545.23	441.00	282.81	158.19	250.42	0.36

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.6: Robustness Check for Firm Size Effect

	University graduates			University non-graduates		
	(1) BMI	(2) Weight (kg)	(3) Height (m)	(4) BMI	(5) Weight (kg)	(6) Height (m)
Panel A: without firm size variable (Table 2)						
Salaried before						
Policy Reform × After	-1.065***	-2.987***	0.002	-0.021	0.166	0.003
	(0.321)	(1.071)	(0.004)	(0.287)	(0.740)	(0.002)
Number of observations	202	202	202	480	480	480
Panel B: with firm size variable						
Salaried before						
Policy Reform × After	-1.069***	-3.076**	0.001	-0.020	0.177	0.002
	(0.371)	(1.209)	(0.003)	(0.286)	(0.735)	(0.002)
Number of observations	197	197	197	441	441	441

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita.

Table A.7: Decomposition (Health Behaviors)

				Eating Habits					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Walking and exercise(HD)	Distance btw energy requirement and actual intake	= 1 if drinking alcohol	Staple food (g/d)	Main dishes (g/d)	Meat dishes (g/d)	Fish dishes (g/d)	Vegetables (g/d)	Ratio of fish dishes
Salaried before									
Policy Reform × After	-0.340*	346.338*	0.020	-44.387	-75.148	-30.572	-44.575	-29.831	-0.008
	(0.176)	(188.711)	(0.220)	(78.714)	(48.747)	(39.540)	(27.194)	(52.078)	(0.055)
× Univ.=1	0.605***	-264.963	-0.317*	-146.842	-30.906	-69.306*	38.399	15.749	0.099**
	(0.154)	(193.351)	(0.187)	(99.745)	(57.599)	(41.343)	(30.539)	(68.767)	(0.049)
× Serial7 score ≥ 3rd quartile	0.255	-141.547	0.066	76.138	4.984	-28.140	33.124	50.198	0.049
	(0.168)	(189.680)	(0.195)	(96.858)	(52.975)	(40.322)	(28.954)	(52.395)	(0.054)
× Word Recall score ≥ 3rd quartile	-0.097	-258.668	-0.003	-121.411	70.490	57.308	13.183	2.566	-0.036
	(0.155)	(182.529)	(0.201)	(83.779)	(55.855)	(41.652)	(27.354)	(53.702)	(0.046)
× Discount Rate ≥ 3rd quartile	-0.050	-172.280	-0.045	130.728	33.635	25.630	8.005	38.954	0.014
	(0.162)	(183.249)	(0.154)	(94.972)	(63.571)	(50.106)	(25.428)	(58.985)	(0.050)
× Be not intrested in own health	-0.192	762.919***	-0.440	377.059**	96.863	66.764	30.099	55.124	0.020
	(0.236)	(258.496)	(0.305)	(157.371)	(130.020)	(90.535)	(57.001)	(155.656)	(0.086)
× Do not have confidence in own health	-0.272	-204.711	0.078	-83.045	93.269	58.127	35.141	-45.559	-0.063
	(0.225)	(232.663)	(0.171)	(111.396)	(71.347)	(54.920)	(34.478)	(72.632)	(0.064)
× Self-rated probability of living at age 80 ≥ 3rd quartile	0.281*	165.991	-0.198	-59.153	28.946	23.877	5.069	9.383	-0.032
	(0.150)	(193.674)	(0.136)	(80.628)	(61.075)	(47.926)	(28.205)	(53.612)	(0.052)
Number of observations	515	515	515	515	515	515	515	515	515
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean before the reform									
Salaried	0.13	502.43	0.79	488.21	428.63	260.24	168.40	246.90	0.40
Self-employed	0.08	533.76	0.73	531.42	438.03	276.55	161.48	249.45	0.37

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

[2] All specification are estimated using FE model, with clustered robust standard errors in parentheses.

[3] Included are age and marriage dummy variables, number of children, income, home ownership, hours worked, physical stress at the workplace, job stress at the workplace, current occupation dummy variables, cross terms of the occupation dummies at 1st wave and survey year dummy variables, and prefecture level macroeconomic variables such as GDP and income per capita, as well as interaction terms between the control variables and the university dummy variable.

References

Aizer, A. and Stroud, L. (2010). Education, Knowledge and the Evolution of Disparities in Health. *NBER Working Paper Series*, 15840.

Bijwaard, G. E. and Van Kippersluis, H. (2016). Efficiency of Health Investment: Education or Intelligence? *Health Economics*, 25(9):1056–1072.

Bijwaard, G. E., van Kippersluis, H., and Veenman, J. (2015). Education and health: The role of cognitive ability. *Journal of Health Economics*, 42:29–43.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic status and health in childhood: The origin of the gradient. *American Economic Review*, 92(5):1308–1334.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA - Journal of the American Medical Association*, 315(16):1750–1766.

Conti, G., Heckman, J., and Urzua, S. (2010). The Education-Health Gradient. *American Economic Review*, 100(2):234–238.

de Walque, D. (2007). How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda. *Journal of Development Economics*, 84(2):686–714.

de Walque, D. (2010). Education, Information, and Smoking Decisions. *Journal of Human Resources*, 45(3):682–717.

Eide, E. R. and Showalter, M. H. (2011). Estimating the relation between health and education: What do we know and what do we need to know? *Economics of Education Review*, 30(5):778–791.

Ganpule, A. A., Tanaka, S., Ishikawa-Takata, K., and Tabata, I. (2007). Interindividual variability in sleeping metabolic rate in Japanese subjects. *European Journal of Clinical*

Nutrition, 61(11):1256–1261.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):223–255.

Grossman, M. (2006). Chapter 10 Education and Nonmarket Outcomes. *Handbook of the Economics of Education*, 1(06):577–633.

Grossman, M. (2015). The relationship between health and schooling : What ' s new ? *Nordic Journal of Health Economics*, 3(1):7–17.

Hackl, F., Halla, M., Hummer, M., and Pruckner, G. J. (2015). The Effectiveness of Health Screening. *Health Economics*, 24(8):913–935.

Ichimura, H., Hashimoto, H., and Shimizutani, S. (2009). JSTAR First Results 2009 Report: Japanese Study of Aging and Retirement. *REITI Discussion Paper Series*, (09-E-047).

Iizuka, T., Nishiyama, K., Chen, B., and Eggleston, K. (2017). Is Preventive Care Worth the Cost? Evidence from Mandatory Checkups in Japan. *NBER Working Paper Series*.

Inui, T., Ito, Y., Kawakami, A., MA, X. X., Nagashima, M., and Zhao, M. (2017). Empirical Study on the Utilization and Effects of Health Checkups in Japan. *RIETI Discussion Paper Series*, 17-E-082.

Kajitani, S. (2011). Working in old age and health outcomes in Japan. *Japan and the World Economy*, 23(3):153–162.

Kawamura, A., Noguchi, H., and Oikawa, M. (2019). Effects of Expansion of Health Service Expenses by Local Governments on Health : Evidence from Japanese Health Checkup Policy Revision. mimeo.

Kim, H. B., Lee, S. A., and Lim, W. (2019). Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea. *Journal of Health Economics*, 65:1–14.

Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker,

E. A., and Nathan, D. M. (2002). Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *New England Journal of Medicine*, 346(6):393–403.

Lleras-Muney, A. and Lichtenberg, F. (2002). The Effect of Education on Medical Technology Adoption: Are the More Educated More Likely to Use New Drugs. *NBER Working Paper*, (9185).

Motegi, H., Nishimura, Y., and Oikawa, M. (2020). Retirement and Health Investment Behaviors: An International Comparison. *The Journal of the Economics of Ageing*, page 100239.

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

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*, 32(3):792–847.

Price, J. and Simon, K. (2009). Patient education and the impact of new medical research. *Journal of Health Economics*, 28(July 2001):1166–1174.

Sasaki, S. (2008). Dietary reference intakes (DRIs) in Japan. *Asia Pacific Journal of Clinical Nutrition*, 17(SUPPL. 2):420–444.

Schellenberg, E. S., Dryden, D. M., Vandermeer, B., Ha, C., and Korownyk, C. (2013). Lifestyle Interventions for Patients With and at Risk for Type 2 Diabetes. *Annals of Internal Medicine*, 159(8):543.

Semyonov, M., Lewin-Epstein, N., and Maskileyson, D. (2013). Where wealth matters more for health: The wealth-health gradient in 16 countries. *Social Science and Medicine*, 81:10–17.

Tsugane, S., Sasaki, S., and Tsubono, Y. (2002). Under- and overweight impact on mortality among middle-aged Japanese men and women: a 10-y follow-up of JPHC study cohort

I. *International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity*, 26:529–537.

Zhao, M., Konishi, Y., and Glewwe, P. (2013). Does information on health status lead to a healthier lifestyle? Evidence from China on the effect of hypertension diagnosis on food consumption. *Journal of Health Economics*, 32(2):367–385.

Zhao, M., Konishi, Y., and Noguchi, H. (2017). Retiring for better health? Evidence from health investment behaviors in Japan. *Japan and the World Economy*, 42:56–63.