What happens when health screenings are subsidized? Evidence from biennial subsidies in South Korea*

Siho Park†

September 20, 2023

This study examines the role of subsidies in health screening participation, selection into screening and disease diagnosis by using biennial subsidies from the South Korean national health insurance. The biennial subsidies provided at even ages increase the take-up and produce two types of spillover. Within an individual, there is positive spillover in take-up from subsidized to not-subsidized screenings and within each screening, there is positive spillover in take-up between spouses. Subsidies for screening also induce participation of compliers with worse health conditions than always-takers due to negative selection on income. Lastly, screenings induced by subsidies increase hospital visits for a new illness.

*I am grateful to Julian Reif, Hyuncheol Bryant Kim, David Molitor, Mark Daniel Bernhardt, and seminar participants at UIUC AMRL, MVEA conference for helpful feedback.

†University of Illinois Urbana-Champaign. sihohp2@illinois.edu.

1 Introduction

Health screening is a set of medical tests on asymptomatic individuals to identify the presence of a disease. An early diagnosis can lead to treatment at a curable stage, which prevents premature death and lowers health care costs.[1] Given the goal of screening is to find a disease, the ideal target is those with bad health conditions who might have a disease but do not yet know about it. However, much of the empirical evidence finds it is healthy people who get screened and unhealthy people more likely to have a disease do not participate and go undiagnosed (Pill et al., 1988; Waller et al., 1990; Bender et al., 2015; Jones et al., 2019). Einav et al. (2020) and Kowalski (2023) examine mammogram recommendations in the US and an influential randomized controlled trial (RCT) in Canada and confirm that screening policies induce the participation of healthy people.[2] This selection pattern is also held responsible for many studies' failure to find significant effect of screenings on mortality and morbidity.[3]

However, previous studies do not pay much attention to the nature of the treatment in the screening program. Different treatments or policies affect different parts of the population and lead to different selection patterns. While Einav et al. (2020) find doctors' recommendations for mammogram induced the participation of healthy people, another policy offering more direct appeals to unhealthy people may perform better and induce participation of less healthy people that are more likely to have a disease. Mapping out different policies and corresponding selection patterns is important to better target those who are most likely to benefit from health screenings.

This motivates my analysis: what is the role of subsidies in health screening? Specifically, I investigate three research questions. First, what is the effect of subsidies on health screening participation? Second, how does the selection pattern change in response to

[1] Cutler (2008) argued that early treatment of cancers through screenings made the most important and cost-effective contribution in the war on cancer since 1990 in the US.

[2] The selection of healthy people in using preventive care is not limited to screening.

[3] One example is the general health screening that consist of tests for basic health conditions like measuring blood pressure or blood glucose levels. A meta-analysis of clinical randomized controlled trials (RCT) concludes with high certainty that general screenings have little or no effect on future mortality and morbidity (Krogsbøll et al., 2012). One potential reason the meta-analysis provides is sample selection of healthy people. Another example is the delayed the mammogram starting age from 40 to 50 by the United States Preventive Services Taskforce (USPSTF).

subsidies? Is there same selection of healthy people into screening or does offering direct monetary incentives appeal more to poor people with worse health conditions? Lastly, what is the effect of subsidies and subsequent screenings on disease diagnoses?

To answer these question, I use the research design that exploits biennial subsidies provided by the Korean National Health Insurance Service (NHIS).[4] The NHIS subsidizes 90 to 100 percent of the costs for 4 types of screenings (general, stomach, breast, cervical) biennially and 2 types of screenings (liver, colorectal) annually from age 40.[5] The biennial subsidies are provided in a year when one's age is even numbered.[6] I show this subsidy rule is random conditional on age and creates large variation in screening take-up between the even age group and the odd age group. Hence, the research design is to compare the even age group (treatment) with the odd age group (control) conditional on age.

This research design is unique in that it features multiple experiments and everyone is ultimately treated. The variation comes not from treatment status, but from treatment timing. Every year, treatment and control groups switch and a new experiment is conducted. This creates an incentive to reallocate screenings intertemporally such that people delay or receive screenings earlier to be eligible for subsidies.[7] While intertemporal substitution is an important mechanism through which biennial subsidies generate variation in take-up, I present evidence that it has limited impact on the results of selection and diagnosis analyses.

This study uses a Korean health panel dataset that includes information on social demographics, health care usage and health behaviors. It is annual individual level survey data from 2008 to 2018 based on the sample of about 7,000 households and 21,300 individuals in the starting year. Despite yearly data collection, the health care usage section includes visit level information including detailed records of health screenings.[8]

[4] I use the word 'biennial' in the sense that it is provided once every two years.

[5] The subsidy starting age is 30 for cervical screening and 50 for colorectal screening. For detail, refer to section 2.

[6] The year refers to calendar year. The policy rule verbatim states that those born in an even numbered year are eligible for subsidies in an even numbered year and those born in an odd numbered year are eligible for subsidies in an odd numbered year. An equivalent way to phrase this rule is that one is eligible for subsidies during a calendar year when one's age turns an even number.

[7] Using the data on screening dates, I show that delaying is not as common as hastening screening.

[8] This was made possible through survey participants recording specifically designed health diaries and keeping receipts from every visit to a hospital. See section 4 for detail.

This enables me to observe the exact date and type of screenings, medical tests performed and diseases found through screening.

This paper presents three findings. First, screening subsidies significantly increase the take-up of subsidized screenings. All 4 types of biennial screenings show higher take-up at even ages than at odd ages due to subsidy eligibility. The biennial subsidies also generate two types of spillover. First, within an individual, there is positive spillover onto annual-subsidy and no-subsidy screenings. Annual-subsidy screenings (liver and colorectal) and no-subsidy screenings (lung and prostate) have no reason to show systematic difference in take-up between even and odd ages. However, they also exhibit significantly higher take-up at even ages, suggesting positive spillover from biennial subsidies. Second, within each screening, there is positive spillover between spouses. One's subsidy eligibility increases not only one's own, but also spouse's probability of screening. The spousal spillover, however, works only when wives are eligible for subsidies, not when husbands are eligible.

Analysis of selection into screening in response to subsidies reveals that compared to always-takers, compliers have worse health conditions and are more likely to be diagnosed with a disease through screening. This is in contrast to the findings of Einav et al. (2020) that show US mammogram recommendation at age 40 induce participation of healthier compliers with the lower rate of true positive breast cancer. One reason is that compliers are negatively selected on income. I show that direct monetary incentives induce compliers from lower socioeconomic background with lower income and less education. The comparison of health behaviors also show compliers are less likely to smoke, drink and exercise. This is consistent with negative selection on income, given many previous studies showing the consumption of alcohol, cigarette and exercising are normal goods (Cawley and Ruhm, 2011; Gallet and List, 2003; Gallet, 2007; Apouey and Clark, 2015; Armstrong et al., 2018; Thibaut et al., 2017). This implies that providing subsidies for health screening performs better than mere recommendations in the sense that it better targets people with low income and worse health conditions. Selection with respect to never-takers suggests compliers are positively selected in health behaviors, consistent with previous studies (Oster, 2020; Cutler and Lleras-Muney, 2010).

Lastly, I find screenings induced by subsidies lead to diagnoses of a new disease and seeking medical care. Despite the information on screening results, it cannot be used to estimate the effect of screening on diagnoses, since diagnoses are unobserved for nonparticipants. To overcome this econometric challenge, I infer diagnoses from outpatient care data that recorded outpatient hospital visits regardless of participation in screening. For each hospital visit, the dataset contain information on the reason of visit coded using International Classification of Diseases (ICD) 10 and whether the visit was a first visit for a new illness or a recurring visit. I use first hospital visits for a new illness as a proxy for new diagnoses and combine with ICD-10 codes to use first hospital visits for a cancer as a proxy for new diagnoses of a cancer. Using subsidy eligibility as an instrument, I find that screening leads to 7.6 percent significant increase in first hospital visits for a new illness and 89.3 percent significant increase in first hospital visits for a cancer. The results are mostly driven by diagnoses in breast cancer.

This study makes contribution to three strands of literature. The most directly related strand of literature is the work on health screening policies. Previous studies examined age recommendation, expansion in health insurance, subsidies or workplace wellness program to encourage participation in health screening (Einav et al., 2020; Kadiyala and Strumpf, 2016; Bitler and Carpenter, 2016; Kim and Lee, 2017; Jones et al., 2019). They show that these policies raise screening take-up and lead to diagnoses of cancers and other adverse health conditions. My study considers substantial subsidies in screening and reveals that it not only increases take-up of directly subsidized screenings, but also has positive spillover on the take-up of other screenings as well. Examining the effect of the screening policies on diagnoses can be empirically challenging, since in most of the datasets on health screening, diagnoses are available only conditional on getting screened. Abaluck et al. (2016) and Einav et al. (2020) overcome this challenge by using the claims and cancer registry data. Similarly, I observe diagnoses conditional on getting screened, but use the unique survey question on whether a hospital visit is a first visit for a new illness or a recurring visit to estimate the effect on diagnoses.[9]

[9]Another strand of literature on health screening focuses on health indicators reported during screenings and examine the discontinuity of indicators in defining health conditions, such as BMI cutoff for

Next, it relates to the literature on spillover in health behaviors. People rarely make decisions in isolation, but with each other, and there can be many ways through which one's behavior can be affected by peers (Leibenstein, 1950; Manski, 1993, 2000). Using exogenous variation in peers, many empirical studies document spillover effects in health behaviors or conditions, such as obesity (Christakis and Fowler, 2007; Cohen-Cole and Fletcher, 2008; Kling et al., 2007), male circumcision (Kim et al., 2018), substance use (Argys and Rees, 2008; Lundborg, 2006), exercise (Carrell et al., 2011), and so forth. This study presents evidence of spousal spillover and show that screening behavior is also subject to peer effects.

Lastly, this study contributes to the literature on selection in health screening and preventive services in general. This study investigates selection into screening given subsidies and suggests selection of compliers can vary with different treatment. Two directly related papers that compare always-takers and compliers are Einav et al. (2020) and Kim and Lee (2017). Considering mammogram age recommendation at 40, Einav et al. (2020) finds that compliers display lower rate of true positive cancer and have better health conditions than always-takers. In contrast, this study compares even and odd age groups and finds compliers are more likely to find a disease in screening throughout the age range from 40 to 89 and have worse health conditions than always-takers. I argue that a reason for opposite findings is different treatment and support it with the evidence showing that treatment directly offering monetary incentives induce participation of compliers with lower income and subsequently worse health conditions. Kim and Lee (2017) also consider subsidies for screening and discover that compliers have worse health conditions than always-takers, consistent with my findings. The opposite selection pattern discovered for different treatments suggests each selection pattern is specific to the treatment and can vary with different treatment. This calls for the need to investigate selection patterns for different treatments and presents the possibility for better targeting people most likely to benefit from screening.

defining obesity or blood pressure cutoff for hypertension (Almond et al., 2010; Kim et al., 2019; Iizuka et al., 2021). They estimate the marginal value of health information or medical care and impact on future health outcomes and behaviors.

The rest of the paper is organized as follows. Section 2 introduces the Korean health screening program and the screening subsidies. Section 3 outlines the identification strategy using subsidy eligibility as a source of exogenous variation. Section 4 discusses the Korean health panel dataset used for this study. Section 5 presents the results and section 6 concludes.

2 Institutional background

The National Health Screening Program in Korea is composed of 3 main programs: general health screening, cancer screenings and infant/children health screening. This study focuses on the general and cancer screenings. General health screening comprises of tests to measure basic health conditions, such as the measurement of weight, height, and blood pressure, chest X-ray, dental test, blood test, uroscopy and health risk evaluation. Among the general screening participants, those diagnosed to be at high risk of hypertension, diabetes or cognitive dysfunction are asked to participate in additional tests with consultation. There are 5 types of cancer screenings covered by the National Health Insurance Service (NHIS) during the study period: stomach, breast, cervical, liver and colorectal screenings. This study focuses on these 5 types of cancer screenings and additionally on lung and prostate cancer screenings that are not subsidized, but captured in the dataset.[10]

Based on the frequency of subsidies, screenings can be grouped into 3 categories: biennial, annual, and no-subsidy screenings. Table 1 summarizes the subsidy rules. Biennial screenings are general, stomach, breast and cervical screenings, and they are subsidized every other year. The biennial schedule is that those who are born in even (odd)-numbered years are entitled to a subsidized screening at any time during the even (odd)-numbered calendar years. For instance, a woman born in 1968, an even birth year, is entitled to get a subsidized stomach screening any time during 2022, an even year, but one must pay the full cost for the same stomach screening in 2023, an odd year. If we define age as the difference between the current year and the year of birth, that is, current

[10]Lung cancer screening began to be subsidized right after the study period in 2019.

year - year of birth, then one is entitled to subsidized screenings during a calendar year when the age is even-numbered. This is because the difference between an even and an even number, or an odd and another odd number, is always even. The group of annual screenings consists of liver and colorectal screenings. These screenings are subsidized every year.[11] The no-subsidy screenings are lung and prostate screenings and they are not covered by the NHIS.

For each subsidized screening, irrespective of biennial or annual, there is a cutoff age where the subsidies start to kick in. For general, stomach, breast and liver screenings, the subsidies start at age 40. There are no subsidies before age 40, even if one's age is even. Hence, general, stomach and breast screenings that follow the biennial schedule are subsidized at ages 40, 42, 44 and onwards. Liver screening that follows the annual schedule is subsidized at ages 40, 41, 42, and onwards. The age cutoff for cervical screening is 30 and the one for colorectal screening is 50.[12] Since cervical screening belongs to biennial group, it is subsidized at ages 30, 32, 34 and onwards, while colorectal screening following annual schedule is subsidized at ages 50, 51, 52 and onwards. There is no age cutoff for lung and prostate screenings, since they are not subsidized at all.[13]

The amount of subsidies is full coverage for the general health screening and 90% subsidy for 5 types of cancer screenings. So, one pays nothing for the general health screening, but needs to pay 10% of the cost for cancer screenings. The subsidies became more generous with time. During our study period, cervical screening was fully subsidized, and colorectal screening also became fully subsidized right after the study period. For those with low income, this 10% of the copays are also subsidized, making cancer screenings free.[14] The rough copay amount for each screening is also reported ranging

[11] Liver screening is subsidized only for those classified as high risk for the liver cancer and is subsidized up to twice a year.

[12] Starting from year 2016, the age cutoff for cervical screening was lowered to 20. Since the age cutoff was 30 for the majority of the study period, I use age 30 as the cutoff for cervical screening.

[13] There are 2 exceptions to the subsidy rules for the general health screening. First, one can be entitled to biennial subsidy for general screening if one is formally employed or head of the household. Second, those with a "non-office" job are entitled to free general screening every year, not every other year. Note that these rules apply only to the general health screening, not any other cancer screening. My data show that these two rules are not strictly followed. Therefore, this study abstracts away from these exceptions.

[14] Kim and Lee (2017) exploits this income cutoff for free cancer screenings to examine take-up and cancer detection.

from $0 for free screenings to $16 in the case of colorectal screening. Note that they are 10% of the full costs and are the amount that screening participants need to pay.[15]

3 Identification

This study exploits quasi-random variation in screening take-up generated by biennial subsidies from age 40. The variation stems from a birth year being even or odd, and I show this is a plausibly random variation that is orthogonal to important covariates. It divides the sample into roughly equally sized treatment and control group as in randomized controlled trials (RCT).

Figure 1a motivates the identification strategy. It plots the average take-up rate of general screening for each age that follows biennial subsidy schedule from age 40. Before age 40, there is a negligible difference in screening rates between the red even age group eligible for the subsidy and the blue odd age group ineligible for the subsidy. However, with the subsidy starting at 40, the screening rate jumps to over 30 percent at even ages, while the screening rate at odd ages remains around 12 percent. This study uses this variation and compares the even age group, also called treatment group, with the odd age group, control group, to study the impact of screening subsidies.[16]

While the distinction between even and odd ages seems plausibly exogenous, simple comparison between the two groups suffers from confounding factors. Table 2 checks the balance between the even age group (treatment group) and the odd age group (control group) among individuals with age 40 to 89. Column 3 shows the treatment group is younger.[17] Along with the age difference follows imbalances in other covariates correlated with age. They are less likely to be female, more educated, has higher income and more likely to live in a larger household.

This is a mechanical effect due to the subsidy design that starts to provide subsidy

[15]Depending on the type of tests received and year, the copay amount varies. The reported amount is based on 2018 data using an exchange rate of 1 USD = 1,000 KRW.

[16]An alternative research design is regression discontinuity design that takes advantage of age cutoff at 40. This study does not take this road due to the coarseness of the running variable, but focuses on the variation between even and odd ages.

[17]Note that the age variable is the difference between current year and year of birth. My dataset does not contain true age, but it is reasonable to expect similar difference in true age.

at age 40, an even number. Given that the age distribution is roughly declining from age 40, the even age group is always younger than the odd age group in an analytical sample starting from age 40, regardless of the ending age. This creates imbalances in age between the treatment and the control group and other covariates correlated with age.

I argue that the only difference between the treatment and the control group is age and all the other differences are the result of imbalance in age. If true, controlling for the age effect should restore the balance between the two groups. Hence, my identifying assumption is that conditional on flexible functions of age, the assignment mechanism, that is, whether one is born in odd or even year and hence has odd or even age, is random. Column 4 in Table 2 shows the conditional difference between the treatment and the control group after including the linear splines of age with 5 years interval as control variables. Conditional on the function of age, the differences between the two groups become muted and statistically insignificant. Note that this is not an artifact of standard errors becoming larger. The precision barely changes, but it is the point estimates that become smaller after controlling for age effect. This implies that age was indeed the only factor leading to an imbalance between the treatment and the control group. Conditional on age, the two groups are otherwise similar and any difference between the two groups can be attributed to the causal effect of screening subsidy. Table A6 in the Appendix shows that the balance is robust to using linear splines of age with different lengths of intervals.

This motivates the following econometric specification. The analytical sample is those with age from 40 to 89.

$$y_{it} = \alpha + \gamma \cdot age\_even_{it} + \mathbf{f}(\mathbf{age_{it}}) + \varepsilon_{it} \tag{1}$$

The outcome variable, y_{it}, is screening take-up of individual i in year t. age_even_{it} is the indicator variable for whether individual i has even age in year t, and hence belongs to the treatment group. $\mathbf{f}(\mathbf{age_{it}})$ is a function of age flexible enough to remove the age effect between the even and odd group. I use linear splines with 5 years interval as the main specification and provide robustness checks for using different length of interval in the

Appendix section F. Conditional on this term, the parameter of interest, γ, estimates the average treatment effect of screening subsidy on the outcome variable. The error term, ε_{it}, is clustered at the individual level.

The usage of panel data suggests we have repeated experiments. The comparison between even and odd age groups does not require panel data. One year of cross sectional data are sufficient for current research design. Hence, this study uses uses panel data as pooled cross sectional data and account for the fact that the same individual appears multiple times in different years by clustering the standard error at the individual level. Despite the usage of panel data, I do not use any panel method to make it clear that my identification strategy does not require panel structure. In the Appendix section F, I present robustness checks where I estimate models with individual fixed effects and the results are robust to the usage of panel method.

This research design is unique in that the treatment and the control group are all treated, but at different times. The two groups switch every year and this is different from a typical RCT setting where control group is never treated.[18] This can result in intertemporal substitution in screening timing, that is, people delay screening or get it earlier such that they get it when it is subsidized. This intertemporal reallocation is an important mechanism through which subsidies generate variation in screening take-up. It has implications on the interpretation of the results and I discuss them along the way in each section. Furthermore, I discuss the intertemporal substitution in more detail in Appendix section A by examining the distribution of screening months over the years.

Another way to show that starting age is the driver of difference in age and subsequently other covariates is to adjust the analytical sample such that it starts with an odd number. If the treatment group is younger because the sample starts from age 40, an even number, then for an analytical sample starting from either 39 or 41, the treatment group should be older than the control group. Table A7 in the Appendix explores the balance with adjusted samples and confirms this is true. Average age is larger for the treatment group when the analytical sample starts from an odd age and the imbalances in other

[18]The biennial design is quite common in health screening context. Many screenings, including mammogram guidelins in the US, are recommended once every two years.

covariates also switch signs. This confirms that age is indeed the driver of imbalance in other covariates and the difference in age stems from the starting age of the sample. This also motivates a robustness check where I estimate Equation (1) with samples starting from age 39, 40 and 41, and examine robustness of the estimates. The imbalance runs in opposite directions depending on whether the starting age is even or odd, and this provides nonparametric bounds of the estimates. The results for bounding robustness checks are presented in Appendix section F.

4 Data

This study utilizes the collection of Korean Health Panel Study datasets spanning the years 2008-2018.[19] It is a collection of yearly panel datasets that started with about 7,000 households and 21,300 individuals in 2008. To guarantee national representativeness in response to gradual attrition, the second cohort of 1,800 households with 5,000 individuals was added to the sample in 2014. Data were collected through face-to-face interviews. All household members were surveyed every year by survey enumerators using computer-assisted personal interviewing (CAPI), and hence, all variables are self-reported.

The collection of datasets includes rich information on demographic and socioeconomic characteristics, health care usage, health behaviors and so forth. The demographic and socioeconomic status variables, such as income and education level, are defined on a yearly basis. They will be used as control variables and also to characterize compliers.

There are 3 separate health care usage datasets: outpatient, inpatient and emergency care. Although data collection was carried out annually, the unit of observations in these datasets is at the visit level.[20] For each visit to a hospital, the dataset contains the date of visit, hospital bills and drug expenditures incurred, diagnosis and whether the hospital visit was a first visit for a new illness as opposed to a recurring visit for an illness one

[19] It is version 1.7.1 made jointly by Korean Institute for Health and Social Affairs (KIHSA) and National Health Insurance Services (NHIS).

[20] This was possible because survey participants were asked to keep a specifically designed health diary and leave detailed records of every visit to a hospital or a pharmacy with receipts. The enumerators collected the health diary and all the receipts at annual visit, compared each entry in the diary with receipts, and recorded the data.

is already aware of.[21] Records on health screenings can be found in the outpatient care dataset. For each screening, it includes the type of screening, medical tests performed, and diagnoses if any disease was found.

The health behavior dataset includes annual information on smoking, drinking and exercising behavior. Each behavior contains information on the intensity. For smoking and drinking, I consider current drinker or smoker status, and additionally everyday drinker or everyday smoker to take into account the intensity of the behavior. These variables were defined based on the action in the previous one month of the time of interview. Similarly for exercise, there are 3 types by intensity: vigorous exercise, moderate exercise and walking. Specific examples for such activities were given at the survey. They are defined based on the action in the previous one week of the survey.

5 Results

5.1 Effect of subsidy on screening take-up

Figure 1a, 1b, 1c and 1d present average screening rates for 4 types of biennial screenings. There is a large divergence in screening rate between the treatment and the control group after the cutoff age. Using Equation (1), I estimate the effect of biennial screening subsidy on the take-up of four biennial screenings in the same year conditional on a flexible function of age. Table 3 presents the estimation results. Column 1 shows that the participation rate for any type of biennial screenings is 12 percent in the absence of subsidy, but it increases by 20 percentage point (169 percent) given subsidies. The corresponding estimates for each type of biennial screening are also significantly positive suggesting the large impact of subsidy on take-up. The reported F-statistics show the first stage is strong (Angrist and Pischke, 2009; Bound et al., 1995).

An important mechanism behind the effect of subsidy is intertemporal substitution. The knowledge of biennial subsidy allows individuals to temporally allocate screenings to take full advantage of the financial incentives. People can either delay or move screenings

[21] Diagnoses were coded using International Classification of Diseases 10 (ICD-10).

to an earlier date so as to receive them when they are entitled to subsidies. My first stage estimates measure the effect of biennial subsidy program that includes both the price effect of subsidy itself and also intertemporal reallocation. Appendix section A provide evidence of intertemporal substitution by examining the month of screening take-up when one is eligible or ineligible for subsidies. The month distribution of take-up suggests strong tendency to get it early before the year of subsidies pass by. However, the opposite pattern, delaying the screening to next year when one becomes eligible for subsidies, is hardly noticeable.

An interesting age pattern of screening rates is that the take-up for all types of screening declines at old ages. Howard et al. (2009) reveal the same pattern exists in the US and European countries for mammogram and colorectal screening. This is because many screening guidelines come with upper bounds for recommended age.[22] For example, due to insufficient evidence to weigh benefit against harm, USPSTF guideline does not make any recommendation of (nor discourage) mammogram for women aged 75 or above. The benefits from screening drop at old ages as the chance of death from other competing causes rises. In contrast, the potential harm from medical tests rises with age (Howard, 2005). Many screening tests can be physically demanding, especially for old people, and it is exacerbated in case of false positive results. These are potential reasons why screening rates drop at old age.[23] Note that for our screening subsidy, there is no upper bound in age.

Table A8 in the Appendix show that the first stage estimates are robust to different sets of control variables. Adjusting the linear splines of age function to have 3 or 7 years interval, instead of 5, does not affect the resulting estimates. Including additional control variables shown in Table 2 on top of linear splines of age hardly affects the magnitude but increases the precision of the estimates, boosting confidence in the exogeneity of the biennial subsidy criterion. The estimates are also robust to the inclusion of individual

[22]Refer to https://www.uspreventiveservicestaskforce.org/uspstf/recommendation-topics/uspstf-a-and-b-recommendations for the screening guidelines recommended by the US Preventive Services Task Force.

[23]As people live longer and longer, the need to provide preventive screenings for older adults may become stronger. USPSTF has more than 30 recommendations that are relevant to older adults, but considers it an area that requires more research (Moyer et al., 2013; Sazlina, 2015; Eckstrom et al., 2013).

fixed effect terms on top of age function.[24]

The first stage estimates are also robust to imbalances that run in opposite directions. Table A9 in the Appendix reports the estimation result using samples with starting age 39, 40 and 41 without the age control variables. Even though imbalances in covariates run in opposite directions depending on the starting age as shown in Table A7, the first stage estimates are highly robust. This implies that the imbalance between the treatment and the control group is minimal and hence hardly affects the first stage estimates.

5.2 Inter-screening spillover

This section investigates spillover in screening take-up within an individual from biennial to annual and no-subsidy screenings. I show that annual and no-subsidy screenings display biennial pattern in take-up starting from age 40 as a result of positive inter-screening spillover.

Liver and colorectal screenings are subsidized every year. Hence, there is no reason to expect systematic difference in take-up between even and odd ages. However, Figure 2a and 2b show that take-up is much higher at even ages indicating positive spillover from biennial screenings. Similar pattern can be observed for no-subsidy screenings. Lung and prostate screenings were not subsidized at any year during the study period. This implies there should be no systematic difference in take-up between even and odd ages in the absence of spillover. Figure 2c and 2d show that although the means are noisier, people are more likely to get them at even ages.

Table 4 presents estimation results of the spillover effect using analytical sample of age range [40, 89]. Being eligible for biennial subsidy at even ages increases the take-up of liver screening by 2.7 percentage points (96 percent), colorectal screening by 3.3 percentage points (127 percent), lung screening by 0.6 percentage points (67 percent) and prostate screening by 0.7 percentage points (78 percent). They correspond to intent-to-treat (ITT) estimates. In Panel B, I report the local average treatment effect (LATE)

[24]Note that even though I am using panel data, no panel structure is required to answer the research question. Hence, including individual fixed effect terms should help with the precision, but should not affect the point estimates.

by instrumenting the take-up of any biennial screening on even age dummy. The ITT estimates are scaled by the first stage coefficient and the resulting estimates are both statistically and economically significant (Angrist et al., 1996; Imbens and Angrist, 1994).

The spillover can also be observed at the subsidy starting age. All but cervical and colorectal screenings have an age cutoff at 40. Cervical screening subsidy starts at age 30. Figure 1d show that despite subsidy starting age at 30, the increase in take-up is larger at age 40 than at 30. This is due to the fact that most of screenings have subsidy starting age at 40 and spillover effect kicks in at age 40. Similarly, as shown in Figure 2b, annual subsidy for colorectal screening starts at age 50, but the take-up increases sharply at even ages from age 40 indicating positive spillover from other screenings that have cutoff at 40. This spillover stemming from subsidy starting age is estimated using regression in Table A1 in the Appendix.

One mechanism of inter-screening spillover is the chance to receive multiple screenings at the same time in one hospital visit. This can be explained by fixed costs in visiting a hospital, such as psychological toll, administrative or travel costs. Or it could be that receipt of one screening can remind participants of the importance of regular health screenings and induce participation in other types as well. Table A2 in the Appendix explores this possibility. It shows that people are less likely to receive a single screening, but more likely to receive annual screenings and biennial screenings together when they are eligible for subsidies.

Take-up of all screenings including annual- and no-subsidy screenings closely follows the biennial pattern from age 40. This provides justification for using the one rule of biennial pattern and age cutoff at 40 for the following analyses of selection and causal effects of health screening.

5.3 Spousal spillover

This section examines spillover in screening take-up across individuals. Using household members relationship information, I show that the screening participation of a spouse increases the probability of one's own screening take-up as a result of positive spousal

spillover.

The analytical sample is adjusted to currently married couples both of whom are included in the dataset. There are 4 types of couples by the combination of their ages; (i) even-even, (ii) even-odd, (iii) odd-even, and (iv) odd-odd. Equation (2) examines their participation rate.

$$y_{it} = \theta \cdot age\_even_{it} + \sigma \cdot spouse\_age\_even_{it} + \tau \cdot age\_even_{it} \times spouse\_age\_even_{it} + \phi_{it} \quad (2)$$

The outcome variable, y_{it}, is screening take-up of individual i in year t. Our parameter of interest is σ that captures whether one is more likely to get screening when the spouse has even age and is eligible for screening subsidies. The coefficient of the interaction term, τ, examines if there is any additional increase in take-up when both the husband and the wife are eligible for subsidies in the same year. The error term, ϕ_{it}, is clustered at the couple level.

Table 5 shows there is positive spousal spillover in screening take-up. The outcome variable is one's own take-up of any type of screening. Column 1 presents positive and significant estimates for both one's own (θ) and the spouse's subsidy eligibility (σ). However, there is no interaction effect (τ), that is, spouse's subsidy eligibility increases one's own take-up by the same magnitude regardless of one is eligible or ineligible for the subsidy.

To account for the possibility that 4 couple types by their age combinations might be correlated with factors that can affect screening participation, I control for their characteristics in two ways. First, column 2 controls for one's own and the spouse's demographic characteristics, but the resulting estimates are robust. Next, leveraging the panel structure of the dataset, I include the couple fixed effects instead of demographic controls. This specification compares screening rates within each couple as their subsidy eligibilities change over the years. Column 3 reports highly robust estimates giving confidence to our original specification.

To translate the effect of subsidy into the effect of screening, I instrument the spouse's screening participation on the spouse's subsidy eligibility. Column 4 measures the effect

of spouse's screening take-up on one's own take-up using the two stage least square estimator. Spouse's take-up increases one's own take-up probability by 6.8 percentage point. Column 5 and 6 show that estimates are robust to adding demographic control variables or couple fixed effect terms.

Appendix section C presents further analyses on the spousal spillover. Instead of using the take-up of any type of screening, I estimate the spousal spillover effect for each screening type in Table A5. I also investigate the direction of spillover by splitting the sample into husbands or wives. Table A4 reveals the spousal spillover shows up only from wife to husband. Husbands are more likely to get screening when their wives receive screening, but the opposite direction hardly works.

5.4 Selection into screening

A common argument against health screening is that it is usually healthy people who receive screenings. It leads to a waste of precious medical resources, false positives on healthy people and overdiagnoses (Rubin, 2019; Kowalski, 2023). What matters in terms of policy effectiveness is the average individuals who respond to screening subsidy and participate. They are the ones who would not have taken up screening without subsidy, but are affected by the policy and subsequently participate. Drawing on the terminology of Imbens and Angrist (1994) and Angrist et al. (1996), they are called "compliers" and their characteristics determine the effectiveness of the policy. Given the goal of screening is to detect disease at an early stage, the policy should ideally induce unhealthy compliers who are likely to have a disease, but do not yet know about it.

It is important to distinguish the reference group when characterizing compliers, that is, whether the comparison is made with respect to always-takers or never-takers. Previous studies examining selection into screening find consistent characteristics of compliers relative to never-takers. Compliers are positively selected in that they show higher socioeconomic status, positive health behaviors that lead to better health, and better health outcomes in terms of cancer incidence and mortality (Einav et al., 2020; Kim and Lee, 2017; Kowalski, 2023). This is consistent with the empirical evidence on the correlations

in various positive health behaviors (Oster, 2020; Cutler and Lleras-Muney, 2010).

However, the evidence on the comparison between compliers and always-takers is mixed. Einav et al. (2020) show that women responding to mammogram age recommendation in the US have lower rate of true positive cancer diagnosis than always-takers, implying compliers have better health than always-takers. On the other hand, Kim and Lee (2017) exploit income cutoff for stomach and breast screening subsidy in Korean national cancer screening program and find that both the cancer incidence and all-cause mortality are lower for always-takers than compliers and are lower for compliers than never-takers.[25]

The problem in complier characterization is that individual categorization into either always-takers, never-takers or compliers is generally not possible, since the definitions of these concepts include unobservable potential outcomes.[26] However, average group characteristics can still be derived to make comparisons among three groups. This study employs methods used by Einav et al. (2020), Kim and Lee (2017) and Kowalski (2023) to characterize compliers in the health screening context.[27]

5.4.1 Methodology for characterizing compliers

To formally estimate the average group characteristics and make comparisons between always-takers, compliers and never-takers, I follow the approach used in Einav et al. (2020); Kim and Lee (2017); Almond and Doyle (2011).[28] The key idea is to infer always-

[25] Kim and Lee (2017) and my study consider the same screening program in Korea, but I use different exogenous variation that requires different research design. In my setting, I exploit the biennial subsidy rule that covers 90 percent of the cancer screening costs. While there is remaining copay of 10 percent when one is entitled for subsidies, those with income below the cutoff were further subsidized and waived the 10 percent copay. Kim and Lee (2017) uses this income cutoff for the waiver of 10 percent copay.

[26] It is straightforward in settings with only one-sided noncompliance. In an RCT setting with only always-takers but no never-takers, one can use the control group to compare always-takers to compliers. They are observable, since those who receive treatment in a control group are always-takers and the rest are compliers. Similar comparison can be made in the treatment group when there are only never-takers but no always-takers.

[27] Another way to characterize complier is to use Abadie's Kappa weighting (Abadie, 2003). However, I do not take that approach for two reasons. First, the goal is to compare compliers to always- and never-takers, not just to characterize compliers. The method of my choice provides a consistent framework to characterize and compare different groups. Second, I use the same method as used in previous studies that examine selection in health screening. To ensure comparability of the results, I follow the methods used by Einav et al. (2020), Kim and Lee (2017) and Kowalski (2023). The Abadie's Kappa weighting method produces similar characteristics of compliers as my method (Results not included).

[28] Marbach and Hangartner (2020) gives a nice summary of the methodology.

takers' characteristics from those who participate in screening in the control group and infer never-takers' characteristics from those who do not participate in the treatment group. The exogeneity of the assignment mechanism guarantees that the characteristics of always- and never-takers will be the same in both the treatment and the control group. Finally, the compliers characteristics can be backed out from the equation where the characteristic of screening participants (non-participants) in the treatment (control) group is a weighted sum of always-takers' (never-takers') and compliers' characteristics.

I estimate the following equation to estimate group characteristics.

$$y_{it} = \beta_1 age\_even_{it} + \beta_2 screen_{it} + \beta_3 age\_even_{it} \times screen_{it} + \boldsymbol{\beta_4'}\mathbf{f}(\mathbf{age_{it}}) + \nu_{it} \quad (3)$$

Given a characteristic variable y_{it}, the above equation can be used to estimate the average characteristic of always-takers by imposing the condition that $age\_even_{it} = 0$ and $screen_{it} = 1$. This is because always-takers are the only group that gets screened even in the absence of subsidies. The estimates are given by $g_{AT}(y) = \hat{\beta}_2 + \hat{\boldsymbol{\beta}}_{\mathbf{4}}{}'\mathbf{f}(\mathbf{age_{it}})$. Similarly, noting that never-takers are the ones who do not get screened despite the presence of subsidies, I impose $age\_even_{it} = 1$ and $screen_{it} = 0$, and the estimates are given by $g_{NT}(y) = \hat{\beta}_1 + \hat{\boldsymbol{\beta}}_{\mathbf{4}}{}'\mathbf{f}(\mathbf{age_{it}})$.

Compliers characteristics can be derived from either the screening participants in the treatment group as a weighted sum with always-takers or the screening non-participants in the control group as a weighted sum with never-takers. While random assignment mechanism implies their average characteristics will be the same in both the treatment and the control group, we differentiate them by denoting compliers in the treatment group as treated compliers and compliers in the control group as untreated compliers.

To calculate the characteristic of treated compliers, denote the average characteristic of treated compliers as $g_C^1(y)$, untreated compliers as $g_C^0(y)$, and the screening participants in the treatment group as $g_T(y)$. The screening participants in the treatment group are always-takers and treated compliers whose average is given by $g_T(y) = \frac{\pi_{AT}}{\pi_{AT}+\pi_C} g_{AT}(y) + \frac{\pi_C}{\pi_{AT}+\pi_C} g_C^1(y)$, where π_{AT} and π_C are share of always-takers and compliers in the entire sample. Imposing $age\_even_{it} = 1$ and $screen_{it} = 1$, we get $g_T(y) = \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 +$

$\hat{\boldsymbol{\beta}}_{\mathbf{4}}{}'\mathbf{f}(\mathbf{age_{it}})$. The share of always-takers and compliers can be calculated from the first stage regression given in (1) as $\pi_C = \hat{\gamma}$, $\pi_{AT} = \hat{\alpha}$ adjusting for age.[29] Inserting all the estimated terms, the average value for treated compliers can be calculated as $g_C^1(y) = [(\pi_{AT} + \pi_C)g_T(y) - \pi_{AT}g_{AT}(y)]/\pi_C$. The average characteristics for untreated compliers can be calculated in a similar way.[30]

To characterize compliers in reference to always-takers and never-takers, ratios are calculated between treated compliers and always-takers, $\frac{g_C^1(y)}{g_{AT}(y)}$, and between untreated compliers and never-takers, $\frac{g_C^0(y)}{g_{NT}(y)}$.[31] Standard errors are calculated using bootstrap with 500 replications clustered at the individual level. The null hypothesis used for ratio is that ratio is equal to one.

5.4.2 Comparison between compliers and always-takers

Before diving into a regression analysis, I provide descriptive characteristics of compliers relative to always-takers. I adjust the sample to screening participants and compare the treatment and the control group in terms of disease diagnoses, socioeconomic status and health behaviors. After the sample adjustment, the treatment group consists of always-takers and treated compliers, whereas the control group is only composed of always-takers.[32] The difference between the two groups comes from composition, that is, compliers in the treatment group. Any treatment effect of screening cancels out, since they are all screening participants. Hence, by comparing the treatment and the control group among screening participants, I can compare treated compliers with always-takers.

Figure 4 shows that compliers are more likely to find a disease through screening than always-takers. The screening participants were asked if they found any disease

[29] Under monotonicity, the share for never-takers is $\pi_{NT} = 1 - \pi_{AT} - \pi_C = 1 - \hat{\alpha} - \hat{\gamma}$

[30] I examine selection pattern at age 40 by imposing $age_{it} = 40$. Appendix section G provides selection results at age 50 and 60. The results are robust to different choices of age.

[31] The reason I differentiate treated and untreated compliers is the possibility that health screening may affect health behaviors. While demographic variables are in general pre-specified and not likely to be affected by health screening, health behaviors such as smoking or drinking can be affected by health screening. Comparing treated compliers with never-takers may be contaminated since treated compliers have received screening while never-takers have not. Same applies to comparison between untreated compliers with always-takers. Therefore, I compare always-takers with treated compliers, both of whom participate in screening, and never-takers with untreated compliers, both of whom did not participate in screening.

[32] Monotonicity assumption rules out defiers.

through screening. The figure plots the share of screening participants who reported they were diagnosed with or showed manifestations of a disease. The higher diagnosis rate in the treatment group throughout the age range compared to the control group implies compliers are more likely to be diagnosed with a disease than always-takers. Note that since I restricted the sample to screening participants, the difference does not reflect the different propensity to participate in health screening between the treatment and the control group, but rather the different health conditions of screening participants between treated compliers and always-takers. One can also detect a sudden increase in the diagnosis rate at age 40 in the treatment group compared to before 40. This also implies compliers have worse health conditions and the diagnosis rate jumps as they start to participate at age 40. This is the opposite finding from Einav et al. (2020) who found that true positive breast cancer rate drops immediately at age 40 cutoff in response to mammogram age recommendation.

Table 6 presents average characteristics of always-takers, never-takers, treated compliers and untreated compliers estimated using Equation (3). Column 5 reports the ratio of treated compliers to always-takers and Figure 5 plots the point estimates and 95 percent confidence intervals. The ratio of diagnoses shows that compliers are 55 percent more likely to be diagnosed with a disease through screening than always-takers. This confirms the descriptive evidence shown in Figure 4. By adjusting outcome variable and $screen_{it}$ variable in Equation (3) to be specific to each screening, I examine in which type of screening did compliers have more diagnoses than always-takers.[33] The ratio estimates show that compliers had particularly more diagnoses than always-takers in stomach screening.[34]

The worse health conditions of compliers can be explained by their socioeconomic characteristics. Table 6 and Figure 5 reveal that compliers have lower individual and

[33]There were too few diagnoses from liver, lung and prostate screenings to conduct statistical analysis. So, I only consider the rest.

[34]One reason for the lack of same pattern in breast, cervical and colorectal screenings may be statistical power. There should be large enough take-up and diagnoses to conduct statistical analysis. Stomach screening showed the highest take-up and also high enough diagnosis rate. Colorectal screening also shows similar level of diagnosis rate to stomach screening, but the average take-up is much lower (18% vs 4%).

household income than always-takers. Part of lower individual income can be explained by lower probability to work. Furthermore, they are less educated. Compliers have slightly lower years of education and are less likely to be college graduates. The lower levels of income and education are consistent with the worse health conditions of compliers compared to always-takers (Lindahl, 2005).

Comparison of health behaviors make it clear that compliers are negatively selected on income. If education drove selection pattern, then lower education level should imply higher probability of smoking (Gallet and List, 2003; De Walque, 2007; Grimard and Parent, 2007; Kenkel et al., 2006). However, despite having lower income, compliers are less likely to be a current smoker and everyday smoker. This contradicts negative selection on education. In addition to lower likelihood of being a smoker, compliers are less likely to be a current drinker and everyday drinker. They are also less likely to engage in physical exercise. This pattern is consistent with negative selection on income. Previous studies on health behaviors mostly find positive income elasticity of demand for cigarette, alcohol and exercise (Cawley and Ruhm, 2011; Gallet and List, 2003; Gallet, 2007; Apouey and Clark, 2015; Armstrong et al., 2018; Thibaut et al., 2017). Since compliers have lower income than always-takers, they have worse health conditions and less likely to smoke, drink, and exercise.

One important finding different from previous studies is the worse health conditions of compliers compared to always-takers. Compliers are more likely to be diagnosed with a disease, particularly stomach related disease, than always-takers throughout the entire age range. Einav et al. (2020) analyze the selection into breast screening in response to screening recommendation at age 40. Comparing before and after age 40, they find that among mammogram participants, the share of true positive breast cancer drops sharply at age 40, implying compliers are less likely to have true positive cancer and have better health conditions than always-takers.

An explanation for the opposite selection pattern is the different nature of treatment. Using different instrumental variable (IV) or treatment may result in different pool of compliers and different local average treatment effects (LATE). The treatment considered

by Einav et al. (2020) is mammogram age recommendation at age 40 provided by the U.S. Preventive Service Task Force (USPSTF).[35] Compliers in this setting are women who pay attention to these guidelines and follow the recommendation. On the other hand, treatment in this study is financial assistance in the form of screening fee waiver. Then, the compliers are the ones who decide to get screening only when it is subsidized. Those who take heed of medical guidelines and follow doctor's recommendation could be quite different from those who respond to monetary incentives.[36] My analysis shows that compliers are negatively selected on income and this could be due to the nature of the treatment that directly provides monetary incentives.[37]

A policy implication from different treatments and subsequently different compliers is that screening subsidies can be used as a tool to better target those with worse health conditions. Given the goal of screening is to find a disease at an early stage, it is important to induce participation of those who potentially have a disease and not yet know about it. This study shows that providing subsidies not only increase participation, but more importantly, induce the participation among people with lower income and worse health conditions. The change in the composition of screening participants should be taken into consideration when evaluating the value of the screening subsidies.

5.4.3 Comparison between compliers and never-takers

The comparison between compliers and never-takers shown in Table 6 and Figure 6 confirms positive selection in health behaviors, consistent with previous studies (Kowalski, 2023; Oster, 2020; Cutler and Lleras-Muney, 2010). Selection in terms of socioeconomic status is not strong, since compliers have slightly higher household income and years of education, but are less likely to work and have lower individual income. However, they are clearly positively selected in health behaviors. Compliers are less likely to smoke, less likely to drink and more likely to exercise. Kowalski (2023) confirms similar positive

[35] This guideline was changed in 2009 and reflected in 2016 guideline (Siu and Force, 2016). Women of age 40 to 49 may choose to begin biennial mammogram based on individual risk factors.

[36] Einav et al. (2020) do not find significant differences in social demographics between compliers and always-takers.

[37] Kim and Lee (2017) also consider screening fee subsidy and find corroborating evidence that compliers have higher cancer incidence and all-cause mortality compared to always-takers.

selection in health behaviors in the Canadian National Breast Screening Study, an influential RCT on mammogram that informed the USPSTF guideline. Jones et al. (2019) also find similar selection pattern in a workplace wellness program that includes screening component. Note that it is not possible to infer the health outcomes of never-takers with screening results, since, by definition, they do not get screened.

5.5 Effect of screening

This section estimates the causal effect of health screening on health care utilizations, using the exogenous variation in screening take-up between even and odd ages. I estimate the effect of screening by instrumenting the take-up with the even age variable conditional on the linear splines of age with 5 years interval. The regression specification is given as follows.

$$y_{it} = \eta + \lambda \cdot screen_{it} + \mathbf{f}(\mathbf{age_{it}}) + \mu_{it} \tag{4}$$

The outcome variable y_{it} is the number of outpatient, inpatient and emergency hospital visits of individual i in year t.[38] The main independent variable, $screen_{it}$, is the take-up of any biennial screening and it is instrumented by the even age group dummy. To cope with mechanical imbalance, I include the flexible function of age as control variables. The standard errors μ_{it} are clustered at the individual level. To account for multiple hypotheses testing, I also report Westfall-Young adjusted p-values that take into account testing for 12 hypotheses with 5,000 bootstrap replications (Jones et al., 2019).

This study also measures the effects of screening on disease diagnoses. Our research design where treatment and control group switch every year can only capture the short term effects of health screening that appear in the same year when one is eligible for subsidies. Hence, it is not suitable for measuring long term health outcomes, such as mortality or morbidity. Instead, I examine whether screening leads to diagnoses of new diseases as an intermediate step for long term health outcomes.

An econometric challenge is that screening results which include cancer diagnoses cannot be used as outcome variables. It is because the results are defined conditional on

[38] For outpatient care, I exclude the number of hospital visits for screening.

getting screened. To overcome this problem, I infer cancer diagnoses not from screening results data, but from outpatient hospital visit records that are collected regardless of screening take-up.[39] I define hospital visit that meets two criteria as a cancer diagnosis. First, the visit has to be a first visit for a new illness, which is inferred from a survey question that asked whether it was a first visit for a new illness or a recurring visit. Second, the diagnosis code for the visit needs to be cancer codes in ICD-10.[40] The outpatient data whose unit of observation is at visit level make it possible to count the number of first hospital visits for cancers each year, and I use it as a proxy for cancer diagnoses.

Panel A in Table 7 shows health screening has limited impact on the number of hospital visits for outpatient, inpatient and emergency care. The number of outpatient hospital visits are hardly affected, since the local average treatment effect is less than 2 percent of the control group mean. For inpatient and emergency care, there might be an increase or decrease, respectively, but the estimates are imprecisely estimated.

However, screening leads to more first hospital visits for a new illness. Panel B shows first outpatient hospital visits increases by 0.30 (7.6 percent) and the first visits for cancers increase by 0.014 (89 percent). This suggests that screening subsidies induce people to receive health screening, and it subsequently makes people go to hospitals more for a new illness they found through screening. This confirms the primary goal of health screening; inducing people to find a disease one is not initially aware of and seek medical care. To examine which cancer diagnosis is driving the result, I further break it down to different types of cancer diagnoses.[41] The results show the breast cancer diagnosis is the main one driving the result. The point estimates for other types were generally positive, but imprecisely estimated.

[39]Two studies on medical tests that share the same problem are Einav et al. (2020) and Abaluck et al. (2016). Einav et al. (2020) have access to mammogram tests and results from both the claims and cancer registry data. Abaluck et al. (2016) can also measure computerized tomography (CT) scan results from the claims data. This study also overcome the challenge by using visit level outpatient data that are similar to claims data in that diagnoses are recorded regardless of screening status.

[40]I only consider malignant neoplasms and do not include in-situ and benign ones. This corresponds to ICD-10 codes with a first letter starting with C.

[41]I used the corresponding cancer screening take-up, instead of any screening take-up, in estimation of equation (4).

6 Conclusion

This paper studies the Korean health screening program that offers subsidies for various screenings biennially at even ages. The analysis finds significantly higher participation when subsidies are provided. I also document positive spillover in take-up across screenings that are not directly subsidized. This stems from the tendency to receive multiple screenings at one hospital visit. This implies that one strategy to increase participation in health screenings is to bundle screenings together to make it easier to receive them on the same day. Furthermore, I document spillover between spouses, especially from wife to husband. One is more likely to receive screening when the spouse also receive it.

The screening subsidies create a selection pattern that differs from other screening incentives. Those who respond to subsidies and participate have lower income than those who voluntarily choose to get screened in the absence of subsidies. Subsequently, the compliers show lower education level and worse health conditions. This could be due to the fact that the treatment is monetary incentives and it is those with low income that are most responsive to the treatment. The comparison with those who do not participate in screening in the presence of subsidies reveals that compliers are positively selected in health behaviors. This is in line with ample empirical evidence that compliance with healthy habits and behaviors are highly correlated.

The selection into screening decision showed that with different treatment comes different selection. This calls for the need to examine selection for various interventions to map out the relationship between an intervention and a group of people most sensitive to the intervention. From policy perspective, it is directly related to targeting. For health screenings and preventive care in general, it is important to induce unhealthy people to participate, since they are the ones most likely to benefit from additional care. This requires knowledge of the intervention and the corresponding complier groups that are affected.

The analysis on the causal effects of screening finds significant impact on new diagnoses. The increase in first hospital visits, especially for cancers, shows that screening fulfills its primary goal of inducing new diagnoses and people subsequently seeking med-

ical care. However, it may not always be an improvement in welfare, since Kowalski (2023) suggests many of screening-induced diagnoses can be overdiagnoses.

This paper does not directly examine the causal effect of screening on mortality due to the research design that only allows capturing short term effects. However, it highlights an important caveat in such analysis. The local average treatment effect is driven by compliers. Understanding the treatment and corresponding selection pattern is needed to not only explain the results, but also to design an effective policy.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of econometrics*, 113(2):231–263.

Abaluck, J., Agha, L., Kabrhel, C., Raja, A., and Venkatesh, A. (2016). The determinants of productivity in medical testing: Intensity and allocation of care. *American Economic Review*, 106(12):3730–3764.

Almond, D. and Doyle, J. J. (2011). After midnight: A regression discontinuity design in length of postpartum hospital stays. *American Economic Journal: Economic Policy*, 3(3):1–34.

Almond, D., Doyle Jr, J. J., Kowalski, A. E., and Williams, H. (2010). Estimating marginal returns to medical care: Evidence from at-risk newborns. *The quarterly journal of economics*, 125(2):591–634.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Apouey, B. and Clark, A. E. (2015). Winning big but feeling no better? the effect of lottery prizes on physical and mental health. *Health economics*, 24(5):516–538.

Argys, L. M. and Rees, D. I. (2008). Searching for peer group effects: A test of the contagion hypothesis. *The Review of Economics and Statistics*, 90(3):442–458.

Armstrong, S., Wong, C. A., Perrin, E., Page, S., Sibley, L., and Skinner, A. (2018). Association of physical activity with income, race/ethnicity, and sex among adolescents and young adults in the united states: findings from the national health and nutrition examination survey, 2007-2016. *Jama Pediatrics*, 172(8):732–740.

Bender, A. M., Jørgensen, T., and Pisinger, C. (2015). Is self-selection the main driver of positive interpretations of general health checks? the inter99 randomized trial. *Preventive medicine*, 81:42–48.

Bitler, M. P. and Carpenter, C. S. (2016). Health insurance mandates, mammography, and breast cancer diagnoses. *American Economic Journal: Economic Policy*, 8(3):39–68.

Bound, J., Jaeger, D. A., and Baker, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American statistical association*, 90(430):443–450.

Carrell, S. E., Hoekstra, M., and West, J. E. (2011). Is poor fitness contagious?: Evidence from randomly assigned friends. *Journal of public Economics*, 95(7-8):657–663.

Cawley, J. and Ruhm, C. J. (2011). The economics of risky health behaviors. In *Handbook of health economics*, volume 2, pages 95–199. Elsevier.

Christakis, N. A. and Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4):370–379.

Cohen-Cole, E. and Fletcher, J. M. (2008). Is obesity contagious? social networks vs. environmental factors in the obesity epidemic. *Journal of health economics*, 27(5):1382–1387.

Cutler, D. M. (2008). Are we finally winning the war on cancer? *Journal of Economic Perspectives*, 22(4):3–26.

Cutler, D. M. and Lleras-Muney, A. (2010). Understanding differences in health behaviors by education. *Journal of health economics*, 29(1):1–28.

De Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the vietnam draft as an instrument for college education. *Journal of health economics*, 26(5):877–895.

Eckstrom, E., Feeny, D. H., Walter, L. C., Perdue, L. A., and Whitlock, E. P. (2013). Individualizing cancer screening in older adults: a narrative review and framework for future research. *Journal of general internal medicine*, 28:292–298.

Einav, L., Finkelstein, A., Oostrom, T., Ostriker, A., and Williams, H. (2020). Screening and selection: The case of mammograms. *American Economic Review*, 110(12):3836–70.

Gallet, C. A. (2007). The demand for alcohol: a meta-analysis of elasticities. *Australian Journal of Agricultural and Resource Economics*, 51(2):121–135.

Gallet, C. A. and List, J. A. (2003). Cigarette demand: a meta-analysis of elasticities. *Health economics*, 12(10):821–835.

Grimard, F. and Parent, D. (2007). Education and smoking: Were vietnam war draft avoiders also more likely to avoid smoking? *Journal of health economics*, 26(5):896–926.

Howard, D. H. (2005). Life expectancy and the value of early detection. *Journal of health economics*, 24(5):891–906.

Howard, D. H., Richardson, L. C., and Thorpe, K. E. (2009). Cancer screening and age in the united states and europe. *Health Affairs*, 28(6):1838–1847.

Iizuka, T., Nishiyama, K., Chen, B., and Eggleston, K. (2021). False alarm? estimating the marginal value of health signals. *Journal of Public Economics*, 195:104368.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica: journal of the Econometric Society*, pages 467–475.

Jones, D., Molitor, D., and Reif, J. (2019). What do workplace wellness programs do? evidence from the illinois workplace wellness study. *The Quarterly Journal of Economics*, 134(4):1747–1791.

Kadiyala, S. and Strumpf, E. (2016). How effective is population-based cancer screening? regression discontinuity estimates from the us guideline screening initiation ages. In *Forum for Health Economics and Policy*, volume 19, pages 87–139. De Gruyter.

Kenkel, D., Lillard, D., and Mathios, A. (2006). The roles of high school completion and ged receipt in smoking and obesity. *Journal of Labor Economics*, 24(3):635–660.

Kim, H. B., Kim, B., Pop-Eleches, C., and Jung, J. (2018). Male circumcision, peer effects and risk compensation. *Unpublished Manuscript, Department of Policy Analysis and Management, Cornell University*.

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.

Kim, H. B. and Lee, S.-m. (2017). When public health intervention is not successful: Cost sharing, crowd-out, and selection in korea's national cancer screening program. *Journal of health economics*, 53:100–116.

Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kowalski, A. E. (2023). Behaviour within a clinical trial and implications for mammography guidelines. *The Review of Economic Studies*, 90(1):432–462.

Krogsbøll, L. T., Jørgensen, K. J., Larsen, C. G., and Gøtzsche, P. C. (2012). General health checks in adults for reducing morbidity and mortality from disease: Cochrane systematic review and meta-analysis. *Bmj*, 345.

Leibenstein, H. (1950). Bandwagon, snob, and veblen effects in the theory of consumers' demand. *The quarterly journal of economics*, 64(2):183–207.

Lindahl, M. (2005). Estimating the effect of income on health and mortality using lottery prizes as an exogenous source of variation in income. *Journal of Human resources*, 40(1):144–168.

Lundborg, P. (2006). Having the wrong friends? peer effects in adolescent substance use. *Journal of health economics*, 25(2):214–233.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of economic perspectives*, 14(3):115–136.

Marbach, M. and Hangartner, D. (2020). Profiling compliers and noncompliers for instrumental-variable analysis. *Political Analysis*, 28(3):435–444.

Moyer, V., LeFevre, M., and Siu, A. (2013). Third annual report to congress on high-priority evidence gaps for clinical preventive services. *Rockville, MD: US Preventive Services Task Force*.

Oster, E. (2020). Health recommendations and selection in health behaviors. *American Economic Review: Insights*, 2(2):143–60.

Pill, R., French, J., Harding, K., and Stott, N. (1988). Invitation to attend a health check in a general practice setting: comparison of attenders and non-attenders. *The Journal of the Royal College of General Practitioners*, 38(307):53–56.

Rubin, R. (2019). Debating whether checkups are time wasted or time misused. *JAMA*, 322(2):101–102.

Sazlina, S. (2015). Health screening for older people—what are the current recommendations? *Malaysian family physician: the official journal of the Academy of Family Physicians of Malaysia*, 10(1):2.

Siu, A. L. and Force, U. P. S. T. (2016). Screening for breast cancer: Us preventive services task force recommendation statement. *Annals of internal medicine*, 164(4):279–296.

Thibaut, E., Eakins, J., Vos, S., and Scheerder, J. (2017). Time and money expenditure in sports participation: The role of income in consuming the most practiced sports activities in flanders. *Sport Management Review*, 20(5):455–467.

Waller, D., Agass, M., Mant, D., Coulter, A., Fuller, A., and Jones, L. (1990). Health checks in general practice: another example of inverse care? *British Medical Journal*, 300(6732):1115–1118.

7 Tables

Table 1: Screening schedule

	Biennial				Annual		No-subsidy	
	General	Stomach	Breast	Cervical	Liver	Colorectal	Lung	Prostate
Recommended frequency	2 years	2 years	2 years	2 years	0.5 year	1 year		
Subsidy starting age	40	40	40	30	40	50		
Amount of subsidy	100%	90%	90%	100%	90%	90%	0%	0%
Copay ($)	0	7	3.5	0	10	16		

Notes: This table summarizes the subsidy schedule for health screenings in Korea. Biennial screenings are subsidized once every two years when one's age (=current year - birth year) is even-numbered. Annual screenings are subsidized every year. No-subsidy screenings are not subject to any subsidy by the Korean National Health Insurance Service (NHIS) during the study period. Liver screening is subsidized up to twice a year.

Table 2: Balance table

	(1)	(2)	(3)	(4)
	Treatment group	Control group	Unconditional difference	Conditional difference
Age	58.697	59.240	–0.543***	-
	(12.532)	(12.353)	(0.026)	-
Female	0.530	0.532	–0.002**	–0.002*
	(0.499)	(0.499)	(0.001)	(0.001)
Currently married	0.799	0.798	0.001	–0.001
	(0.401)	(0.402)	(0.001)	(0.001)
Years of education	10.320	10.227	0.093***	–0.003
	(4.510)	(4.538)	(0.009)	(0.008)
Working status	0.610	0.608	0.001	–0.003*
	(0.488)	(0.488)	(0.002)	(0.001)
Individual income	1446.3	1425.7	20.6***	2.8
	(2081.6)	(2068.1)	(5.5)	(5.2)
Household income	4104.4	4086.7	17.7	3.2
	(3708.6)	(3737.9)	(14.6)	(14.3)
Own a house	0.734	0.737	–0.002*	–0.000
	(0.442)	(0.441)	(0.001)	(0.001)
Number of household members	3.067	3.051	0.016***	–0.004
	(1.317)	(1.317)	(0.003)	(0.003)
N	54274	52909		
Share	(0.51)	(0.49)		

Notes: This table reports the unconditional and conditional balance check between the treatment group (even age group) and the control group (odd age group). The sample consists of those with age in [40, 89]. Column 3 reports the unconditional difference between the treatment and the control group. Column 4 reports the difference conditional on linear splines of age with 5 years interval. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 3: Effect of biennial subsidy on take-up

	(1)	(2)	(3)	(4)	(5)
	Any	General	Stomach	Breast	Cervical
Age even	0.204***	0.187***	0.189***	0.191***	0.164***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
N	107183	107183	107183	56923	56923
Adj R^2	0.068	0.061	0.069	0.080	0.074
F-statistic	5022	4804	4835	2908	2520
Sample age range	[40, 89]	[40, 89]	[40, 89]	[40, 89]	[40, 89]
Subsidy starting age	40	40	40	40	30
Age controls	Y	Y	Y	Y	Y
Control group mean	0.121	0.102	0.082	0.067	0.055

Notes: This table reports the effect of biennial subsidy on 4 types of screening take-up. Screenings reported in column 2 to 5 are subject to biennial subsidy when ages are even. Column 1 uses the take-up of any type of screening reported in column 2 to 5. The estimates measure the effect of screening subsidy on take-up conditional on linear splines of age with 5 years interval. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 4: Inter-screening spillover

	(1)	(2)	(3)	(4)
	Annual subsidy		No subsidy	
	Liver	Colorectal	Lung	Prostate
Panel A. ITT				
Age even	0.027***	0.033***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B. LATE				
Biennial screening	0.131***	0.162***	0.031***	0.046***
	(0.006)	(0.006)	(0.003)	(0.006)
N	107183	107183	107183	50260
Sample age range	[40, 89]	[40, 89]	[40, 89]	[40, 89]
Subsidy starting age	40	50	40	40
Age controls	Y	Y	Y	Y
Control group mean	0.028	0.026	0.009	0.009

Notes: This table reports the spillover effect from biennial screenings take-up to the take-up of annual-subsidy and no-subsidy screenings. Panel A shows intent-to-treat (ITT) estimates by regressing take-up on even age group dummy. Panel B shows local average treatment effect (LATE) estimates from two-stage least square regressions using even age group dummy as the instrumental variable for the take-up of any type of biennial screening. All specifications are conditional on the linear splines of age with 5 years interval. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 5: Spillover between spouses

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome var: Own screening take-up					
Age even	0.181***	0.179***	0.182***	0.179***	0.179***	0.180***
	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Spouse age even	0.012***	0.010***	0.013***			
	(0.004)	(0.004)	(0.003)			
Age even	0.000	0.003	0.000			
× Spouse age even	(0.008)	(0.008)	(0.000)			
Spouse screening				0.068***	0.066***	0.076***
				(0.018)	(0.018)	(0.018)
N	101726	101493	101726	101726	101493	101726
Demographic controls		Y			Y	
Couple fixed effects			Y			Y
Estimator	OLS	OLS	OLS	2SLS	2SLS	2SLS

Notes: This table reports the spillover effect in screening take-up between spouses. Outcome variable is one's own screening take-up of any kind. The sample consists of currently married couples. Demographic control variables include age, gender, years of schooling, health insurance type, handicap status, working status, household income decile and survey year of oneself and the spouse. In column 4 to 6, spouse checkup variable is instrumented by spouse age even variable. Standard errors are clustered at the couple level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 6: Complier selection

	(1)	(2)	(3)	(4)	(5)	(6)
	Average value				Ratio	
	Always-takers	Treated compliers	Untreated compliers	Never-takers	CP_1/AT	CP_0/NT
Panel A. Diagnoses						
Diagnosed with a disease	0.258	0.399	-	-	1.550***	-
	(0.006)	(0.007)	-	-	(0.053)	-
Stomach disease diagnosis	0.171	0.273	-	-	1.594***	-
	(0.006)	(0.006)	-	-	(0.074)	-
Breast disease diagnosis	0.022	0.026	-	-	1.180	-
	(0.004)	(0.003)	-	-	(0.263)	-
Cervical disease diagnosis	0.069	0.065	-	-	0.950	-
	(0.007)	(0.006)	-	-	(0.149)	-
Colorectal disease diagnosis	0.190	0.202	-	-	1.061	-
	(0.011)	(0.011)	-	-	(0.104)	-
Panel B. SES						
Individual income	2614	1991	1949	2205	0.762***	0.884***
	(55)	(50)	(53)	(40)	(0.018)	(0.018)
Household income	5568	5030	4997	4778	0.903***	1.046**
	(74)	(69)	(95)	(54)	(0.014)	(0.019)
Years of education	14.174	13.880	13.888	13.560	0.979***	1.024***
	(0.066)	(0.068)	(0.081)	(0.046)	(0.006)	(0.005)
College graduate	0.388	0.333	0.335	0.333	0.859***	1.006
	(0.011)	(0.010)	(0.011)	(0.009)	(0.021)	(0.021)
Working status	0.796	0.713	0.724	0.754	0.896***	0.960***
	(0.010)	(0.010)	(0.012)	(0.008)	(0.013)	(0.014)
Panel C. Health behaviors						
Current smoker	0.233	0.191	0.171	0.298	0.819***	0.573***
	(0.010)	(0.010)	(0.011)	(0.009)	(0.033)	(0.028)
Everyday smoker	0.218	0.181	0.157	0.284	0.830***	0.554***
	(0.010)	(0.009)	(0.011)	(0.009)	(0.034)	(0.028)
Current drinker	0.849	0.812	0.797	0.816	0.957***	0.977*
	(0.009)	(0.009)	(0.012)	(0.007)	(0.012)	(0.013)
Everyday drinker	0.043	0.028	0.035	0.051	0.641***	0.691***
	(0.005)	(0.005)	(0.006)	(0.004)	(0.115)	(0.119)
Vigorous exercise	0.310	0.268	0.283	0.264	0.866***	1.071*
	(0.009)	(0.009)	(0.012)	(0.007)	(0.030)	(0.039)
Moderate exercise	0.456	0.426	0.426	0.389	0.934***	1.096***
	(0.010)	(0.010)	(0.015)	(0.008)	(0.023)	(0.034)
Walking	0.791	0.786	0.778	0.753	0.994	1.033**
	(0.008)	(0.008)	(0.012)	(0.007)	(0.011)	(0.016)

Notes: This table reports the average values of screening diagnoses, socioeconomic status, and health behaviors among always-takers, treated compliers, untreated compliers and never-takers. Treated compliers are compliers in the treatment group who participate in screening. Untreated compliers are compliers in the control group who do not participate. The average value is calculated using Equation (3). Diagnoses are not reported for untreated compliers and never-takers, since by definition, they do not receive screening. The null hypotheses used for ratios are $H_0 : CP_1/AT = 1$ and $H_0 : CP_0/NT = 1$ for comparison with always-takers and never-takers, respectively, where AT = Always-takers, NT = Never-takers, CP_1 = Treated compliers and CP_0 = Untreated compliers. All the average values and ratios are calculated at age 40. Standard errors are calculated using bootstrap with 500 replications and are clustered at the individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

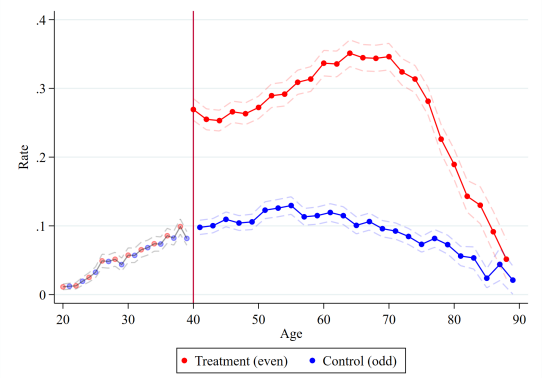
Table 7: Effect of health screening

	(1)	(2)	(3)	(4)	(5)
	Control group mean	ITT	LATE	Adjusted p-values	N
Panel A. Health care utilizations					
Outpatient visit	20.7796	0.0713	0.3494	0.886	107183
		(0.0757)	(0.3713)		
Inpatient visit	0.2329	0.0056	0.0272	0.702	107183
		(0.0039)	(0.0192)		
Emergency visit	0.1252	–0.0024	–0.0117	0.886	107183
		(0.0026)	(0.0127)		
Panel B. First outpatient visits					
First outpatient visit	3.9160	0.0607***	0.2976***	0.003	107183
		(0.0152)	(0.0747)		
First outpatient visit for a cancer	0.0160	0.0029***	0.0143***	0.123	107183
		(0.0011)	(0.0055)		
Stomach cancer	0.0023	0.0007	0.0037	0.675	107183
		(0.0004)	(0.0024)		
Breast cancer	0.0033	0.0016**	0.0086**	0.352	56923
		(0.0008)	(0.0042)		
Cervical cancer	0.0003	–0.0001	–0.0007	0.886	56923
		(0.0002)	(0.0009)		
Liver cancer	0.0009	0.0002	0.0071	0.886	107183
		(0.0003)	(0.0104)		
Colorectal cancer	0.0019	0.0005	0.0150	0.807	107183
		(0.0004)	(0.0126)		
Lung cancer	0.0020	0.00001	0.0011	0.988	107183
		(0.00042)	(0.0666)		
Prostate cancer	0.0032	–0.0011	–0.1525	0.675	50260
		(0.0007)	(0.0994)		

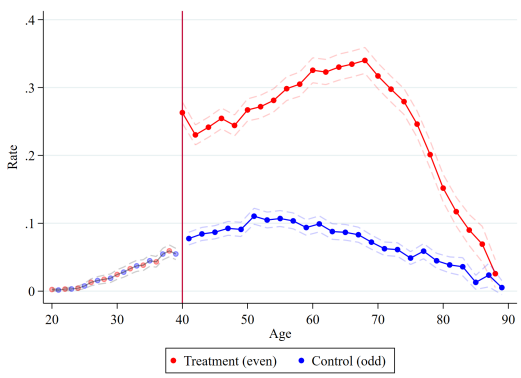
Notes: This table reports the effect of health screening on health care utilizations. Panel A uses the number of outpatient, inpatient and emergency hospital visits except for screening purpose as outcome variables. Panel B uses the total number of first outpatient hospital visits and for specific cancers as outcome variables. First visit is inferred from the visit-level outpatient dataset asking whether the hospital visit was a first visit for a new illness or a recurring visit for an illness one is already aware of. The type of cancer is inferred from ICD-10 diagnosis code recorded for each visit to a hospital. Column 2 ITT estimates report the effect of biennial subsidy on outcome variables. Column 3 LATE estimates report the effect of screening take-up on outcome variables using biennial subsidy as IV. For specific types of cancer screenings shown in Panel B, the take-up of corresponding cancer screening was used to calculate the LATE estimates. Westfall-Young adjusted p-values from second stage regressions are reported in column 4. They account for all 12 hypotheses with 10000 bootstrap replications. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

8 Figures

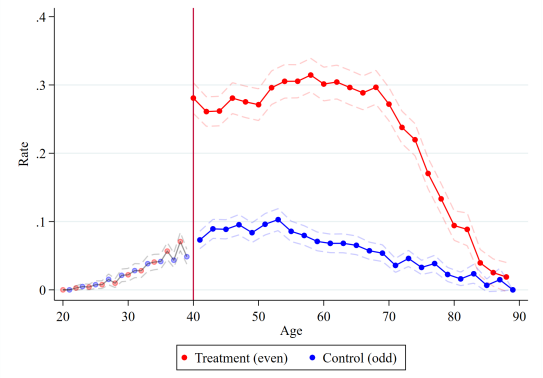
Figure 1: Screening rates by age for biennial screenings



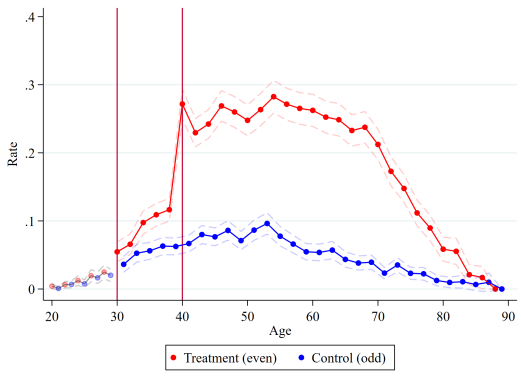
(a) General screening



(b) Stomach screening



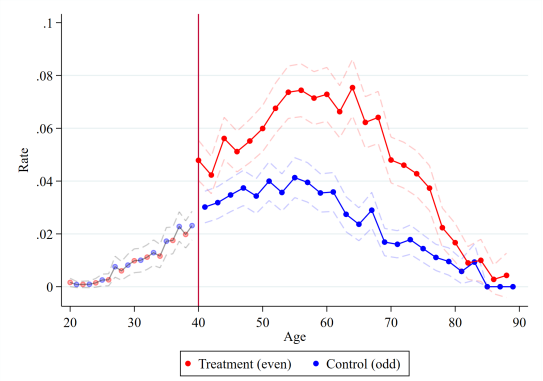
(c) Breast screening



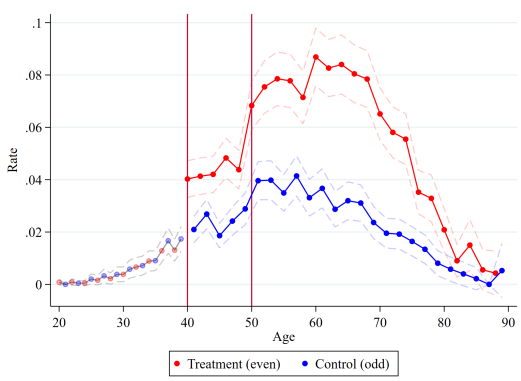
(d) Cervical screening

Notes: Figures show the average screening rates by age for 4 types of screening that are subsidized biennially. Even ages are colored in red and odd ages are colored in blue. Red vertical line shows the subsidy starting age and age 40. Only cervical screening has subsidy starting age at 30 and the rest starts from age 40. Confidence intervals at 95 percent are shown in dashed line, separately for even and odd age group from age 40.

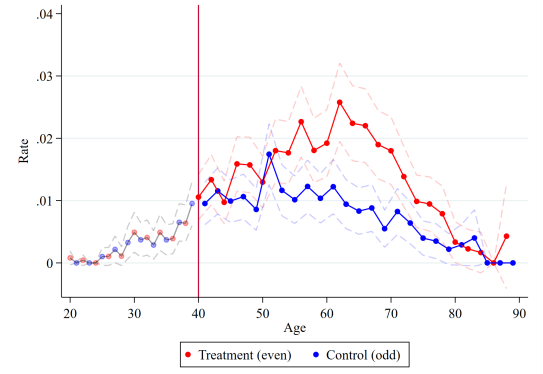
Figure 2: Screening rates by age for annual- and no-subsidy screenings



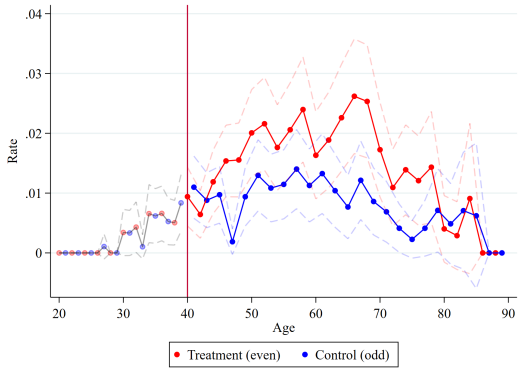
(a) Liver screening



(b) Colorectal screening



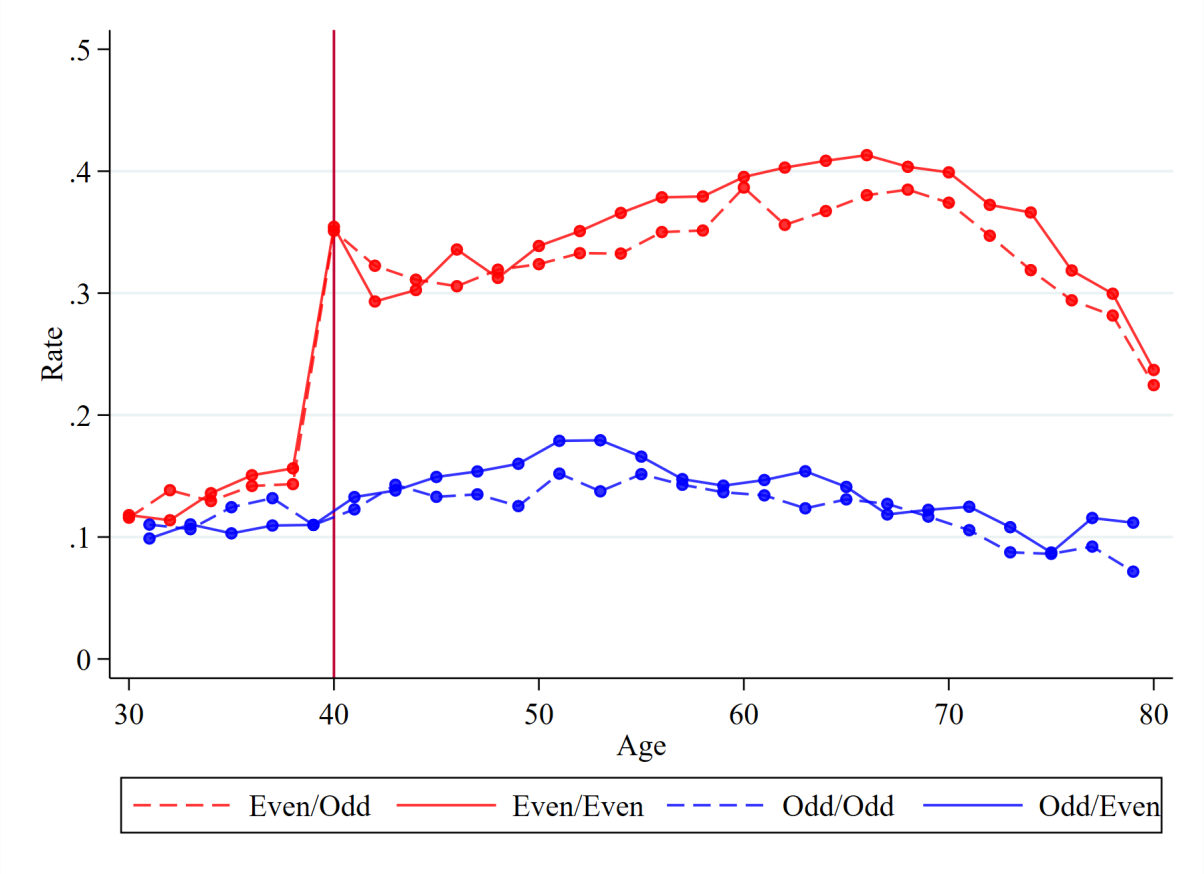
(c) Lung screening



(d) Prostate screening

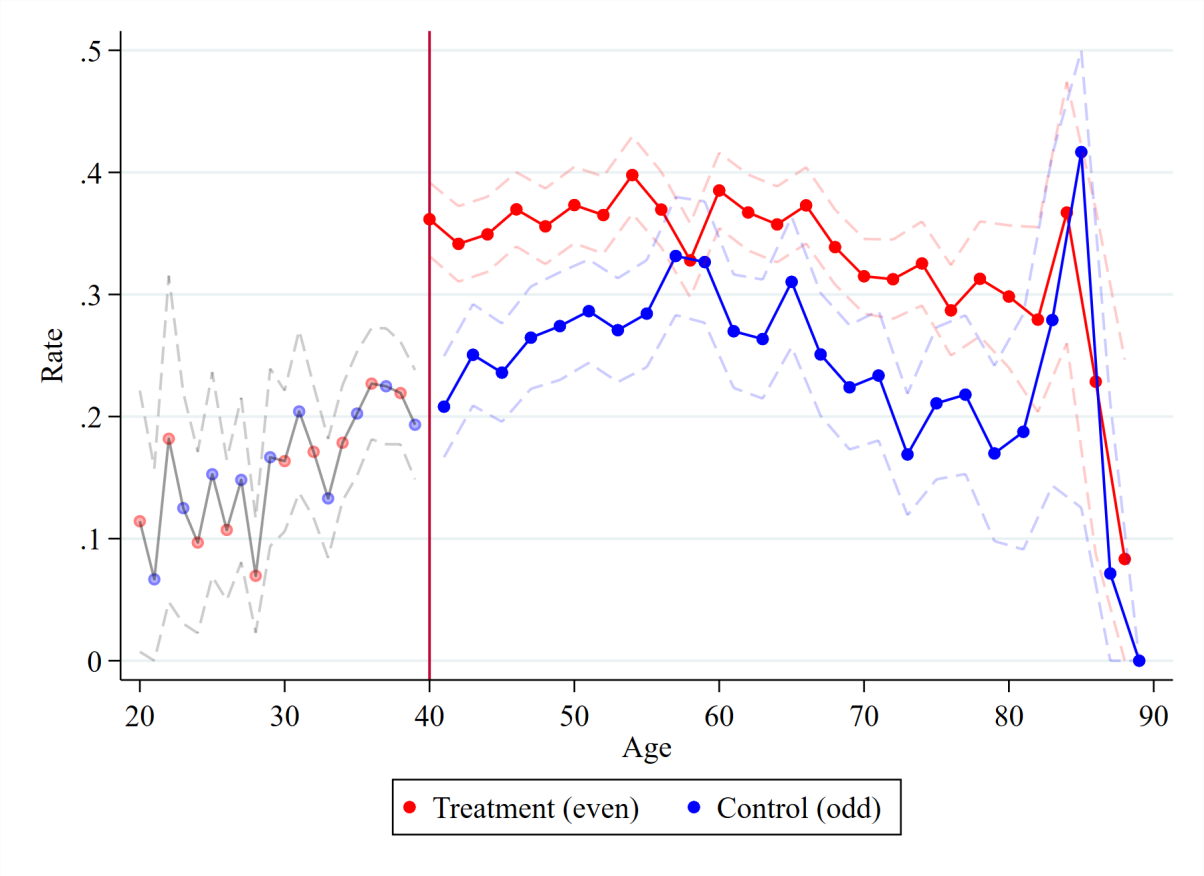
Notes: Figures show the average screening rates by age for liver and colorectal screening that are subsidized annually and lung and prostate screenings that are not subsidized. Even ages are colored in red and odd ages are colored in blue. Red vertical line shows the subsidy starting age and age 40. Colorectal screening has subsidy starting age at 50 and the liver screening starts from age 40. Confidence intervals at 95 percent are shown in dashed line, separately for even and odd age group from age 40.

Figure 3: Spillover in screening take-up between spouses



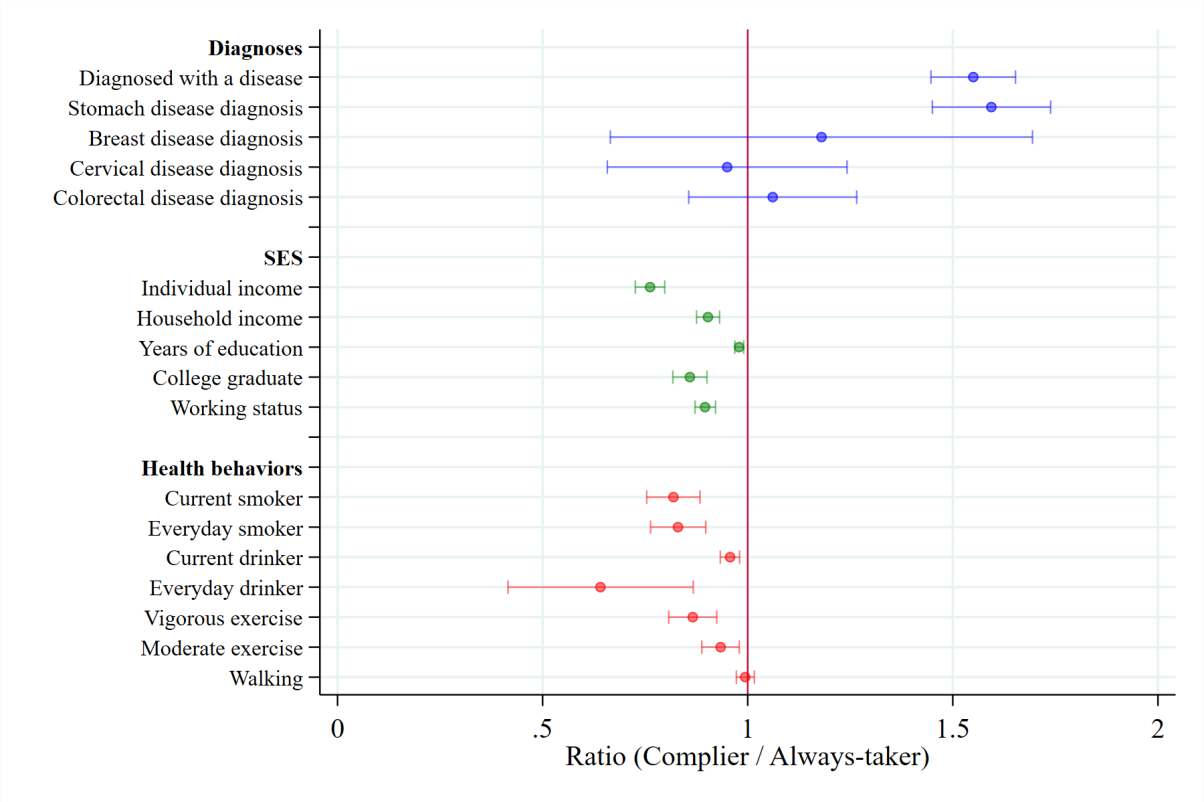
Notes: The figure plots the average screening rates for 4 groups of the currently married people. The sample consists of individuals who are currently married and whose age is in the range [30, 80]. Those whose age is even are colored in red, while those whose age is odd are colored in blue. Those whose spouse's age is even are shown in solid lines, while those whose spouse's age is odd are shown in dashed lines.

Figure 4: Share of screenings with disease diagnoses or manifestations



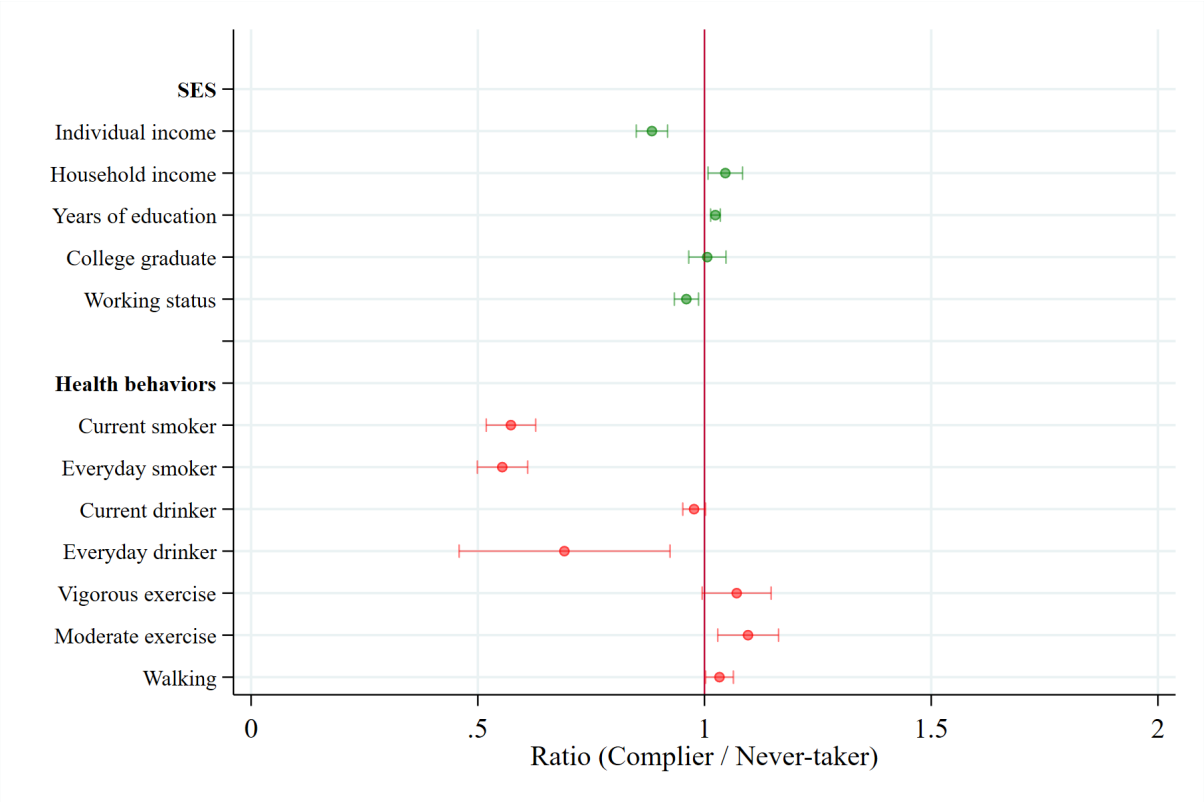
Notes: The figure plots the share of screenings where participants were diagnosed with or showed manifestations of a disease. The sample is restricted to screening participants. For each age, the figure plots the share of screening participants who reported they were diagnosed with a disease or had manifestations of a disease that require further examinations. The types of diseases found were coded using ICD-10 and some examples are reported in Appendix section E. Even ages are colored in red and odd ages are colored in blue. Red vertical line shows the subsidy starting age of 40. Confidence intervals at 95 percent are shown in dashed line, separately for even and odd age group from age 40.

Figure 5: Comparison between treated compliers and always-takers



Notes: The Figure plots the ratio of average characteristics of treated compliers to always-takers with 95 percent confidence intervals. Treated compliers are compliers in the treatment group who participate in screening. The average values and ratios are calculated from the estimation of equation (3) and are reported in Table 6. The standard errors are calculated using bootstrap with 500 replications clustered at the individual level. Diagnoses refer to whether a screening participant was diagnosed with or showed manifestations of a disease that require further examinations. Health behavior variables are dummy variables for engaging in such activities.

Figure 6: Comparison between untreated compliers and never-takers



Notes: The Figure plots the ratio of average characteristics of untreated compliers to never-takers with 95 percent confidence intervals. Untreated compliers are compliers in the control group who do not participate in screening. Diagnoses are not defined for untreated compliers and never-takers, since they do not participate in screening. The average values and ratios are calculated from the estimation of equation (3) and are reported in Table 6. The standard errors are calculated using bootstrap with 500 replications clustered at the individual level. Health behavior variables are dummy variables for engaging in such activities.

Appendix A Intertemporal substitution

This section presents further analyses on intertemporal substitution of screening timing. The research design where the treatment and control group switch every year can induce individuals to receive screening when one is eligible for subsidies. There are two ways substitution can appear. First, in a year of subsidies, agents can move the next year's screening to an earlier date so that they receive it this year when they are eligible for subsidies. Or, in a year of no subsidy, people can delay this year's screening to the next year so that they receive it at subsidized prices. The key is to move the screening to a subsidized year, whether it is delaying it or getting it early.

Using the exact date of screening, I find strong evidence of receiving screening at an earlier date to be eligible for subsidies, but not delaying. Figure A1 plots the month distributions of screening take-up for even and odd age groups separately and puts them adjacent to each other. There are three important patterns. First, screening take-up is larger in the even age group than in the odd age group for all months. This corresponds to the effect of subsidies on take-up as shown in Table 3. Second, the large spike in take-up at December for even age group suggests the tendency to hasten screening before the year passes by and one is no longer eligible for subsidies. This is the first form of intertemporal substitution, that is, receiving screening earlier.[42] Third, delaying the screening does not seem to be common. If people are strategically delaying screening, then the take-up in November or December of odd age group should be declining, accompanied by a large increase in January and February of even age. However, if anything, the take-up in December of odd age is highest and the take-up in January of even age is lowest. This suggests there is less strategic delaying of screening to be eligible for subsidies.

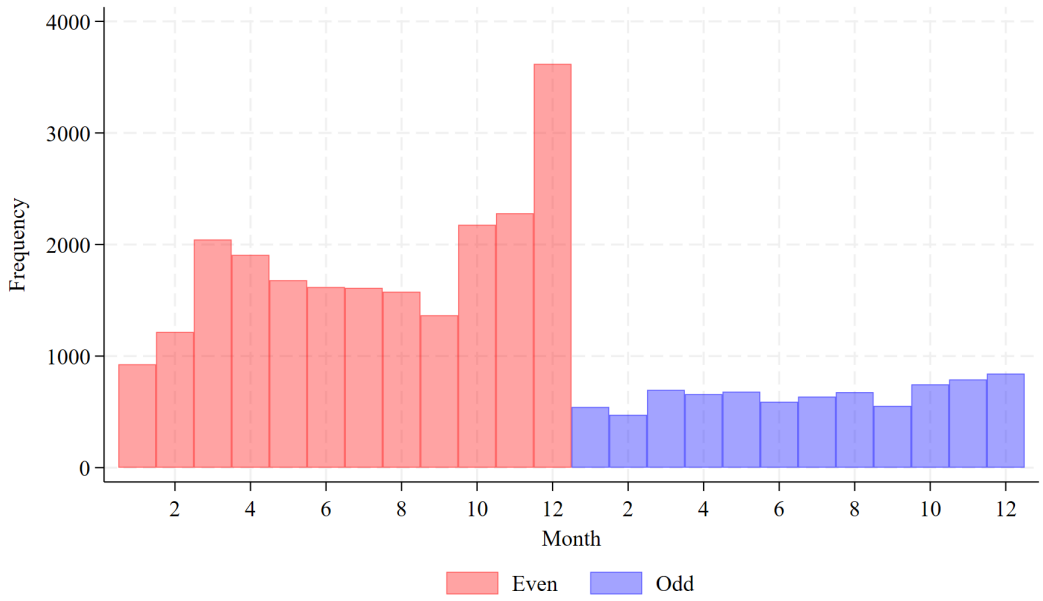
The intertemporal substitution does not affect the results of selection analyses. Section 5.4 showed compliers have worse health conditions than always-takers by comparing the share of screenings with disease diagnoses between the even and odd age group. Given

[42]Some firms enforce employees to receive general screening when one is eligible, since workers failure to receive general screening when one is eligible for subsidies can lead to government fine on firms or employees. Workers putting off might be forced to receive it in December. This could be another reason why the take-up soars in December.

that intertemporal substitution is predominantly getting screenings at an earlier date, if intertemporal substitution is driving this results, we should see that those who receive screening in December at even age have substantially higher disease diagnosis rate than all the other months of even age and those who receive screening in January at odd age have substantially lower disease diagnosis rate. Figure A2 plots the share of screenings with disease diagnoses for each month. It shows that the disease diagnosis rate is consistently higher for all months in even age group, not particularly higher for December. Furthermore, the diagnosis rate is low for all months in odd age group. Odd age January, if anything, seems to have higher diagnosis rate. This disproves the possibility that intertemporal substitution is driving the selection results.

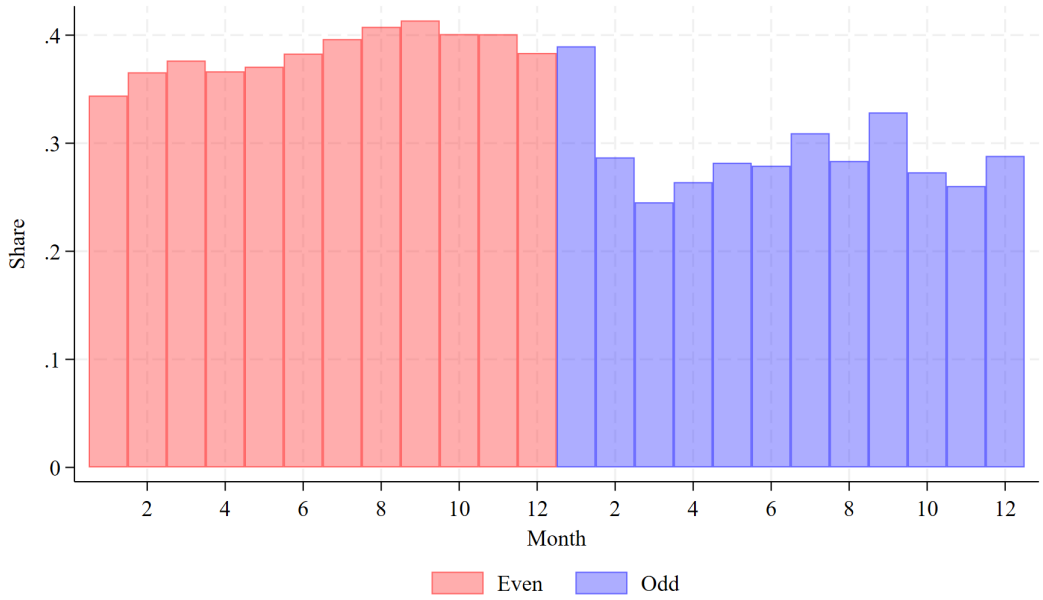
The intertemporal substitution also does not drive the results of diagnosis analysis. Section 5.5 showed that screening induces diagnoses of new diseases by examining the number of first hospital visits for a new illness. Figure A3 plots the number of first hospital visits for a new illness by months separately for even and odd age groups. Even age group shows more first hospital visits for every month and this suggests subsidy-induced screenings lead to discovery of new diseases. Given that intertemporal substitution is primarily happening on December of even age group, if intertemporal substitution is driving the results, we should see substantially larger first hospital visits of even age group in December. However, the difference between the even and odd age group in December is not particularly higher than all the other months. This confirms that intertemporal substitution is not driving the diagnosis results.

Figure A1: Screening take-up by months



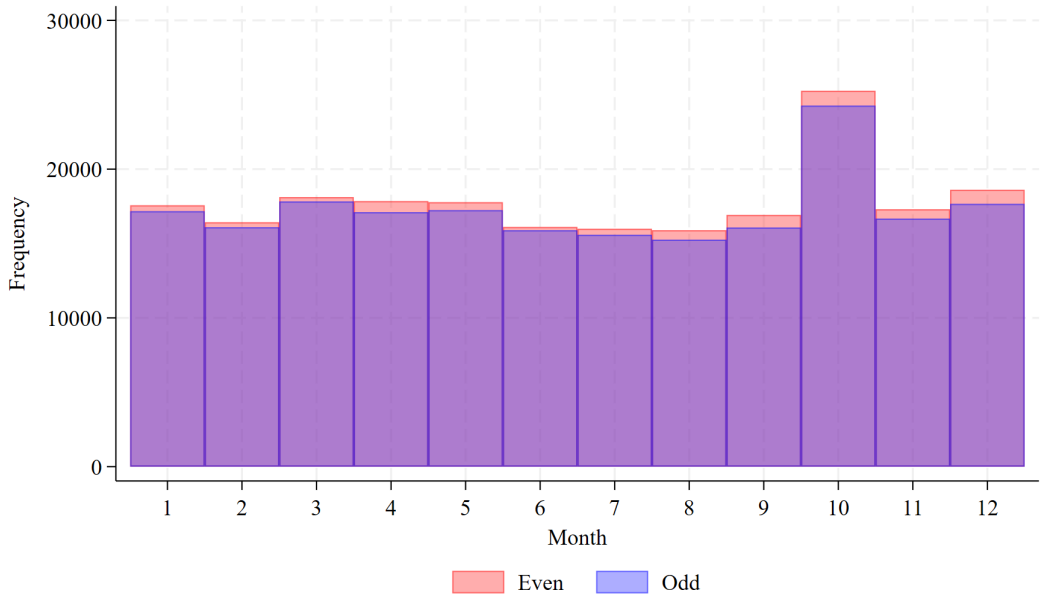
Notes: The figure plots screening take-up by months separately for even and odd age groups. Take-up of any type of screening is considered. They are positioned adjacent to each other to emphasize the moment when the month changes from December to January.

Figure A2: Share of screenings with disease diagnoses by months



Notes: The figure plots the share of screenings with disease diagnoses within each month, separately for even and odd age groups. Take-up of any type of screening is considered. Even and odd groups are positioned adjacent to each other to emphasize the moment when the month changes from December to January.

Figure A3: First hospital visits for a new illness by months



Notes: The figure plots first hospital visits for a new illness by months, separately for even and odd age groups. Even group is colored in red and the odd group is colored in blue.

Appendix B Inter-screening spillover

This section presents additional evidence and potential mechanisms of inter-screening spillover. As explained in section 5.2, the biennial subsidies provided to four types of screenings at even ages affect the take-up pattern for annual- and no-subsidy screenings. Liver and colorectal screenings are subsidized every year, but they show significantly higher participation at even ages when subsidies are provided for biennial screenings as presented in Table 4. Lung and prostate screenings are not subsidized at all, but they also show higher take-up when there are subsidies for biennial screenings.

First, I present evidence that inter-screening spillover can also be found using different subsidy starting ages. All screening subsidies start at age 40, except cervical and colorectal screenings. Cervical screening subsidies start at age 30. As shown in Figure 1d, while take-up jumps at age 30, there is a much larger jump at age 40. This is because all the other screening subsidies start at age 40. Similarly, colorectal screening subsidies start at age 50. However, one sees a jump in screening take-up at age 40, due to the same spillover effect from other screenings with age cutoff at 40. Table A1 presents regression results of spillover using different subsidy starting age for cervical and colorectal screenings.

Next, I explore mechanisms behind inter-screening spillover. One mechanism is that people receive multiple screenings on the same day during one hospital visit, rather than randomly spreading the take-ups. This can be explained by fixed costs in visiting a hospital, such as psychological toll, administrative or travel costs. Or it could be that receipt of one screening can remind participants of the importance of regular health screenings and induce participation in other types as well.

A resulting empirical prediction is that for annual- and no-subsidy screenings, they should be more likely to happen on the same day with biennial screenings when one receives subsidies. Using the exact date of screenings, I test the prediction. I first restrict the sample to screening participants of liver, colorectal, lung and prostate screenings, in each regression. Then, I compare the probability of the screening happening on the same day with general and stomach screenings, two biennial screenings, between even ages and

odd ages.[43] Panel A in Table A2 shows that liver and colorectal screenings are more likely to be received on the same day during the same hospital visit with general and stomach screenings when one is eligible for subsidy. Panel B shows the mirror image. When one is eligible for subsidy, people are less likely to receive one screening on a day apart from other screenings.[44] The estimates were not significant for lung and prostate screenings, possibly due to small take-up.

Lastly, to examine which of the 4 biennial screenings is generating spillover effect, I examine gender difference in spillover effect. Among 4 biennial screenings, general and stomach screenings are subsidized for everyone, but breast and cervical screenings are provided only for women. If the two female screenings are generating spillover, we should see larger spillover effect for women compared to men.

The heterogeneous treatment effect by gender shown in Table A3 does not support this hypothesis. If anything, they seem to be slightly smaller for colorectal screenings. This implies that general and stomach screenings are the ones that generate spillover effects, not breast and cervical screenings. This could be due to the fact general and stomach screenings are the most commonly received types of screenings.

[43] While breast and cervical screenings are also biennially subsidized, I exclude them since they are only for women.

[44] The reason I consider screenings happening 30 days away from any other screening is that for an intricate screenings like colorectal screening that require dietary restrictions and advanced preparation, it is common to receive it on a separate day, but probably in the same week.

Table A1: Inter-screening spillover using different subsidy starting age

	(1)	(2)
	Cervical	Colorectal
Age even	0.041***	–0.002
	(0.005)	(0.001)
Age even × age ⩾ 40	0.123***	0.022***
	(0.006)	(0.002)
Age even × age ⩾ 50		0.018***
		(0.002)
N	69236	130736
Adj R^2	0.072	0.016
Sample age range	[30, 89]	[30, 89]
Subsidy starting age	30	50
Age controls	Y	Y
Control group mean	0.039	0.016

Notes: This table reports inter-screening spillover results for cervical and colorectal screenings using different subsidy starting ages. The subsidy starting age for cervical and colorectal screenings are 30 and 50, respectively. The rest of the screenings covered by the NHIS have subsidy starting age at 40. The sample age range used in both regressions are [30, 89]. Cervical screening regression contains only female. All specifications are conditional on linear splines of age with 5 years interval. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A2: Inter-screening spillover by screening day difference

	(1)	(2)	(3)	(4)
	Annual subsidy		No subsidy	
	Liver	Colorectal	Lung	Prostate
Panel A. Outcome: 1 = conducted on the same day with general & stomach screening				
Age even	0.0399***	0.0480***	0.0146	–0.0433
	(0.0144)	(0.0158)	(0.0283)	(0.0385)
N	4471	4609	1314	634
Control group mean	0.6867	0.5943	0.5788	0.7252
Panel B. Outcome: 1 = conducted alone and more than 30 days away from any other screening				
Age even	–0.0174***	–0.0229***	–0.0066	–0.0002
	(0.0051)	(0.0076)	(0.0085)	(0.0066)
N	4471	4609	1314	634
Control group mean	0.0528	0.2573	0.0353	0.0045

Notes: This table examines if annual- or no-subsidy screenings happen on the same day with biennial screenings. All the coefficients are from separate regressions. For each screening, the sample consists of those who participated in that screening with age $\in$ [40, 89]. Outcome variables used in Panel A regressions are whether the screening happened on the same day with two biennial screenings, general and stomach screenings. Outcome variables used in Panel B regressions are whether only that particular screening happened on that day and more than 30 days away from any other screening. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A3: Gender difference in spillover

	(1)	(2)	(3)
	Liver	Colorectal	Lung
Age even	0.025***	0.036***	0.007***
	(0.002)	(0.002)	(0.001)
Age even × Female	0.002	–0.005*	–0.002
	(0.003)	(0.003)	(0.001)
Female	–0.017***	–0.012***	–0.008***
	(0.002)	(0.002)	(0.001)
N	107183	107183	107183
Control group mean	0.028	0.026	0.009

Notes: This table reports estimates of inter-screening spillover for men and women. The sample consists of those with age from 40 to 89. Those with even age are entitled to free general screening and 90% subsidized stomach screening. Women are additionally entitled to subsidized breast and cervical screenings. Liver screening is subsidized every year from age 40. Colorectal screening is subsidized every year from age 50. Lung screening is not subsidized. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Appendix C Spousal spillover

This section presents the direction of spousal spillover and estimates of spousal spillover in each type of screening. Section 5.3 provided evidence of spousal spillover in screening take-up. When one is eligible for screening subsidies, it not only increases one's own probability of getting screened, but also the spouse's probability of screening as well. Using spouse's subsidy eligibility as an instrument, I estimated the spillover effect of spouse's screening take-up on one's own probability of take-up.

Table A4 shows the spillover effects run from wives to husbands. To investigate the direction of spillover, I split the sample into husbands and wives and estimate the equation (2) again. Table A4 presents the estimation results without using any control variable. As presented in the first two columns, husband's being eligible for screening subsidies have statistically negligible effect on wife's screening take-up. While the coefficient is positive, it is imprecisely estimated. However, the last two columns show that the opposite direction is strong. When wife is eligible for screening subsidies, it leads to significant increase in husband's likelihood of receiving screenings.

Table A5 presents the estimates of spillover effect for each screening type. It shows that spousal spillover exists for most types of screenings. The coefficient estimate of spouse screening instrumented by the spouse's eligibility for subsidies is significantly positive for all screenings except liver and lung screenings.

Table A4: Spousal spillover directions

	(1)	(2)	(3)	(4)
	Among wives		Among husbands	
Even age	0.220***	0.218***	0.142***	0.141***
	(0.004)	(0.004)	(0.004)	(0.004)
Spouse even age	0.007		0.017***	
	(0.004)		(0.004)	
Spouse screening		0.048		0.079***
		(0.031)		(0.018)
N	50863	50863	50863	50863
Estimator	OLS	2SLS	OLS	2SLS

Notes: This table reports the direction of spousal spillover in screening take-up. Column 1 and 2 examine spousal spillover among married females, while column 3 and 4 examine spillover among married husbands. In column 2 and 4, spouse checkup variable is instrumented by spouse even age variable. No control variables are included in the regressions. Standard errors are clustered at couple level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A5: Spillover between spouses by screening types

	(1)	(2)	(3)	(4)	(5)
	General	Stomach	Liver	Colorectal	Lung
Age even	0.162***	0.163***	0.023***	0.029***	0.005***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)
Spouse screening	0.059***	0.065***	0.010	0.022***	0.004
	(0.017)	(0.017)	(0.008)	(0.008)	(0.004)
N	101726	101726	101726	101726	101726

Notes: This table reports the spousal spillover effect for different types of screenings. Outcome variable is one's own screening take-up. The sample consists of currently married couples. The spouse screening variable is instrumented by the spouse having even age. Standard errors are clustered at the couple level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Appendix D Graphical evidence on complier characteristics

This section provides graphical evidence of complier characteristics shown in Figure 5. As shown in Figure 4, simple sample adjustment can be used to compare compliers to always-takers and never-takers. The key is to adjust the sample such that the comparison between the treatment and the control group, an observable characteristic, yields comparison between compliers and always-takers or never-takers.

First, to provide descriptive characteristics of compliers relative to always-takers, I adjust the sample to screening participants and compare the treatment and the control group. After the sample adjustment, the treatment group consists of always-takers and compliers, whereas the control group is only composed of always-takers.[45] The difference between the two groups comes from composition, that is, compliers in the treatment group. Any treatment effect of screening cancels out, since they are all screening participants. Hence, by comparing the treatment and the control group, I can compare the average characteristics of compliers relative to always-takers.

Figures A4a, A4b, and A4c plot the average individual income, share of college graduates and smokers for each age among screening participants. Treatment group has lower income, lower share of college graduates and lower share of smokers compared to control group. This implies that compliers have lower income and they are less likely to be college graduates and smokers compared to always-takers.

A similar comparison can be made between compliers and never-takers by adjusting the sample to screening non-participants. The sample restriction leaves never-takers in the treatment group, but both the compliers and never-takers in the control group. Hence, the difference in composition can be used to infer the difference between compliers and never-takers.

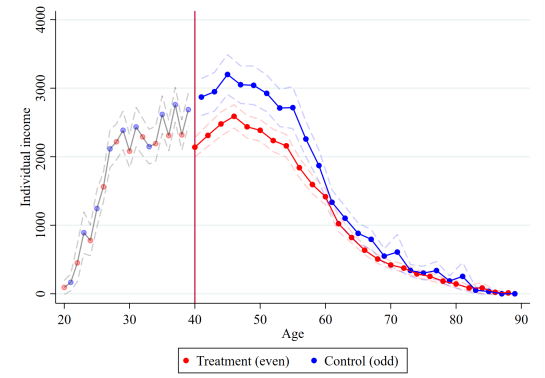
Figure A4d, A4e and A4f plot the same individual income, share of college graduates and smokers for each age among screening-nonparticipants. While the differences in

[45] Monotonicity assumption rules out defiers.

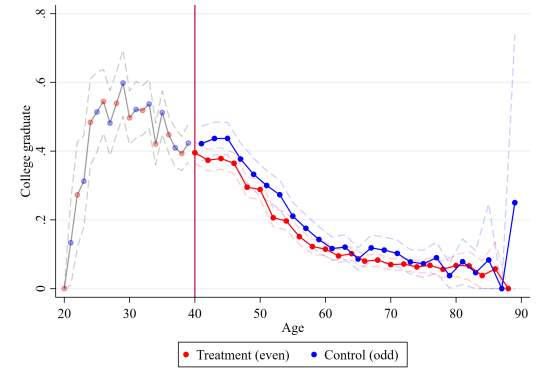
individual income or college graduate rate are not as clear as in comparison between compliers and always-takers, Figure A4f display clear contrast. Control group is less likely to be smoker than the treatment group, suggesting compliers are less likely to be smokers than the never-takers.

Figure A4: Complier characteristics

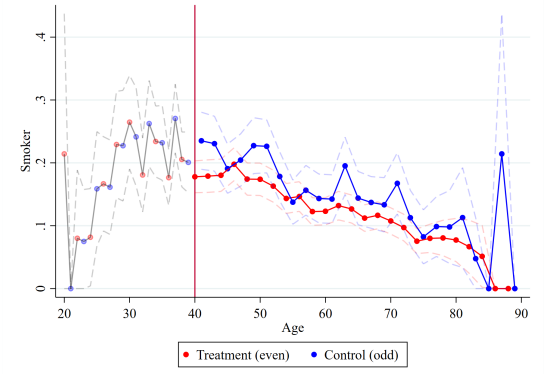
Panel A. Screening participants



(a) Individual income

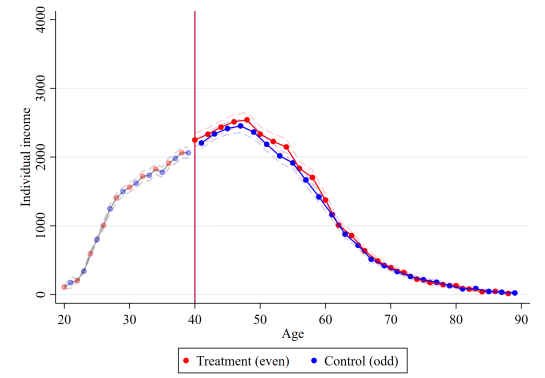


(b) College graduate

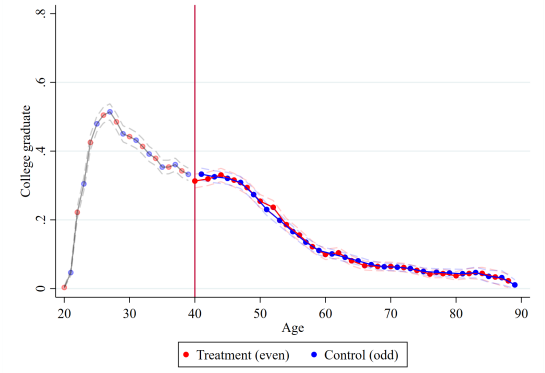


(c) Smoker

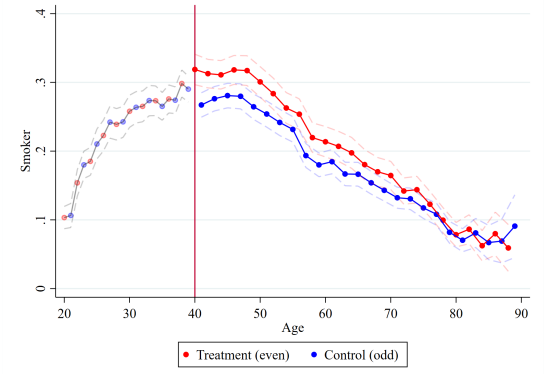
Panel B. Screening non-participants



(d) Individual income



(e) College graduate



(f) Smoker

Notes: Figures plot the average characteristics by age for screening participants and nonparticipants. Panel A figures use samples that are restricted to screening participants. They plot average individual income, share of college graduate and share of smokers. Panel B figures use samples that are restricted to screening nonparticipants. Individual income is denoted in 10,000 Korean Won. Even ages are colored in red and odd ages are colored in blue. Red vertical line shows the subsidy starting age of 40. 95 percent confidence intervals are shown in dashed line, separately for even and odd age group from age 40.

Appendix E Classification of disease diagnosis

This section provides a list of ICD-10 codes used for classifying disease diagnoses. The survey asked screening participants if they had found any disease through screening, and if so, the diagnoses were coded using International Classification of Diseases (ICD) version 10. I list below the ICD-10 codes that were classified as pertaining to each organ. The diagnoses are listed in the order of frequency, so this provides a list of common diagnoses made through screening.

- Stomach diseases
  - K29 Gastritis and duodenitis
  - K52 Other and unspecified noninfective gastroenteritis and colitis
  - K21 Gastro-esophageal reflux disease
  - K25 Gastric ulcer
  - B98 Helicobacter pylori
  - K31 Other diseases of stomach and duodenum
  - K20 Esophagitis
  - C16 Malignant neoplasm of stomach
  - K26 Duodenal ulcer
- Colorectal diseases
  - K63 Other diseases of intestine
  - D12 Benign neoplasm of colon, rectum, anus and anal canal
  - D13 Benign neoplasm of other and ill-defined parts of digestive system
  - R19 Other symptoms and signs involving the digestive system and abdomen
  - C18 Malignant neoplasm of colon
- Breast diseases

  - N63 Unspecified lump in breast
  - N64 Other disorders of breast
  - D24 Benign neoplasm of breast
  - N60 Benign mammary dysplasia
  - C50 Malignant neoplasm of breast

- Female reproductive part diseases
  - N76 Other inflammation of vagina and vulva
  - N71 Inflammatory disease of uterus, except cervix
  - N85 Other noninflammatory disorders of uterus, except cervix
  - N83 Noninflammatory disorders of ovary, fallopian tube and broad ligament

Appendix F Robustness check for econometric specifications and bounding estimates

This section presents the robustness check for econometric specifications. First, I use a different function of age. As discussed in section 3, the identifying assumption is that the treatment and the control group is balanced conditional on a flexible function of age. All the regressions shown in the main paper use linear splines of age with 5 years interval on the age range [40, 89]. This section additionally provides the same regression results using linear splines with 3 years and 7 years interval.

Second, I present regression results with additional control variables. Given age controls are enough to guarantee balance, no more control variables are needed for inference. However, they can be useful by absorbing variations and increasing precision of the estimates. I present the estimation results using linear splines of age with 5 years interval and additionally full control variables shown in the balance table, Table 2, or individual fixed effect terms. These additional control variables should not affect the point estimates. The results are provided in Table A6, A8, A10, A12, and A14.

Lastly, I present bounding estimates using three different samples: (i) age $\in$ [39, 89], (ii) age $\in$ [40, 89], and (iii) age $\in$ [41,89]. As argued in section 3, the sample starting age is creating mechanical imbalance between the treatment and the control group. Table A7 confirms this argument. To gauge the extent to which the imbalances affect the main estimates, I run all the analyses without the age control variables, but with three different samples. The imbalances running in opposite directions depending on the starting age provide nonparametric bounds for the estimates. This obviates the need to specify any functional form of age control variables. The results are provided in Table A9, A11, A13, and A15.

Table A6: Robustness check for balance table

	(1)	(2)	(3)
	3 years	5 years	7 years
Female	–0.002*	–0.002*	–0.002*
	(0.001)	(0.001)	(0.001)
Currently married	–0.001*	–0.001	–0.001
	(0.001)	(0.001)	(0.001)
Years of education	–0.003	–0.003	–0.003
	(0.008)	(0.008)	(0.008)
Working status	–0.003**	–0.003*	–0.003*
	(0.002)	(0.001)	(0.001)
Individual income	1.1	2.8	1.2
	(5.3)	(5.2)	(5.2)
Household income	0.6	3.2	–4.5
	(15.4)	(14.3)	(14.1)
Own a house	–0.000	–0.000	0.000
	(0.001)	(0.001)	(0.001)
Number of household members	–0.004	–0.004	–0.004*
	(0.003)	(0.003)	(0.003)

Notes: This table reports the difference between the treatment group (even age group) and the control group (odd age group) conditional on linear splines of age with 3, 5 and 7 years interval. The sample consists of those with age in [40, 89]. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A7: Bounding estimates for balance check

	(1)	(2)	(3)
	Age ∈ [39, 89]	Age ∈ [40, 89]	Age ∈ [41, 89]
Age	0.521***	–0.543***	0.562***
	(0.026)	(0.026)	(0.025)
Female	–0.001	–0.002**	–0.001
	(0.001)	(0.001)	(0.001)
Currently married	–0.001	0.001	–0.002**
	(0.001)	(0.001)	(0.001)
Years of education	–0.094***	0.093***	–0.107***
	(0.009)	(0.009)	(0.010)
Working status	–0.006***	0.001	–0.007***
	(0.001)	(0.002)	(0.002)
Individual income	–16.235***	20.607***	–24.789***
	(5.470)	(5.508)	(5.618)
Household income	–23.153	17.735	–31.182**
	(14.766)	(14.555)	(14.995)
Own a house	0.003***	–0.002*	0.003**
	(0.001)	(0.001)	(0.001)
Number of household members	–0.027***	0.016***	–0.034***
	(0.003)	(0.003)	(0.003)
N	110121	107183	104153

Notes: This table reports the balance check between the treatment group (even age group) and the control group (odd age group) using samples with different starting age (39, 40, 41). It does not include any control variable. The coefficients report the average value of the treatment group relative to the control group. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A8: Robustness check for first stage

	(1)	(2)	(3)	(4)	(5)
	Any	General	Stomach	Breast	Cervical
Panel A. Linear splines of age					
Interval 3	0.204***	0.187***	0.189***	0.190***	0.163***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Interval 5	0.204***	0.187***	0.189***	0.191***	0.164***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Interval 7	0.204***	0.187***	0.189***	0.190***	0.164***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Panel B. Linear splines with 5 years interval plus additional covariates					
Full controls	0.204***	0.187***	0.190***	0.191***	0.164***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Individual FE	0.206***	0.189***	0.191***	0.192***	0.165***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)

Notes: This table reports the effect of biennial subsidy on 4 types of screening take-up with different control variables. Screenings reported in column 2 to 5 are subject to biennial subsidy when ages are even. Column 1 uses the take-up of any type of screenings reported in column 2 to 5. Panel A uses linear splines of age with 3, 5, and 7 years intervals as controls. Panel B uses linear splines of age with 5 years interval plus additional covariates. Full controls specification includes all the variables reported in balance table (Table 2) as controls. Individual FE specification includes individual fixed effects as controls. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A9: Bounding estimates for first stage

	(1)	(2)	(3)
	Age ∈ [39, 89]	Age ∈ [40, 89]	Age ∈ [41, 89]
Any	0.205***	0.204***	0.204***
	(0.003)	(0.003)	(0.003)
General	0.187***	0.186***	0.188***
	(0.003)	(0.003)	(0.003)
Stomach	0.191***	0.189***	0.190***
	(0.003)	(0.003)	(0.003)
Breast	0.192***	0.191***	0.190***
	(0.003)	(0.004)	(0.004)
Cervical	0.165***	0.165***	0.162***
	(0.003)	(0.003)	(0.003)
N	110121	107183	104153

Notes: This table reports first stage using samples with different starting age (39, 40, 41). It does not include any control variable. The coefficients report the effect of subsidy on screening take-up. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A10: Robustness check for inter-screening spillover

	(1)	(2)	(3)	(4)
	Annual subsidy		No subsidy	
	Liver	Colorectal	Lung	Prostate
Panel A. Linear splines of age				
Interval 3	0.027***	0.034***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Interval 5	0.027***	0.033***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Interval 7	0.027***	0.033***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Panel B. Linear splines with 5 years interval plus additional covariates				
Full controls	0.027***	0.033***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Individual FE	0.028***	0.034***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)

Notes: This table reports ITT estimates of inter-screening spillover results with different control variables. Panel A uses linear splines of age with 3, 5, and 7 years intervals as controls. Panel B uses linear splines of age with 5 years interval plus additional covariates. Full controls specification includes all the variables reported in balance table (Table 2) as controls. Individual FE specification includes individual fixed effects as controls. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A11: Bounding estimates for inter-screening spillover

	(1)	(2)	(3)
	Age $\in [39, 89]$	Age $\in [40, 89]$	Age $\in [41, 89]$
Liver	0.027***	0.027***	0.027***
	(0.001)	(0.001)	(0.001)
Colorectal	0.033***	0.033***	0.034***
	(0.001)	(0.001)	(0.001)
Lung	0.006***	0.006***	0.007***
	(0.001)	(0.001)	(0.001)
Prostate	0.007***	0.007***	0.008***
	(0.001)	(0.001)	(0.001)
N	110121	107183	104153

Notes: This table reports the ITT estimates of inter-screening spillover results using samples with different starting age (39, 40, 41). It does not include any control variable. The coefficients report the effect of biennial subsidy on the take-up of annual-subsidy and no-subsidy screenings. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A12: Robustness check for complier selection

	(1)	(2)	(3)	(4)	(5)	(6)
	Interval 3		Interval 5		Interval 7	
	CP/AT	CP/NT	CP/AT	CP/NT	CP/AT	CP/NT
Panel A. SES						
Individual income	0.754***	0.895***	0.758***	0.899***	0.763***	0.903***
	(0.019)	(0.016)	(0.019)	(0.015)	(0.018)	(0.015)
Household income	0.898***	1.049***	0.900***	1.051***	0.900***	1.050***
	(0.015)	(0.011)	(0.014)	(0.011)	(0.014)	(0.011)
Years of education	0.980***	1.023***	0.981***	1.023***	0.981***	1.023***
	(0.006)	(0.004)	(0.006)	(0.004)	(0.006)	(0.004)
College graduate	0.858***	0.998	0.861***	1.001	0.863***	1.001
	(0.022)	(0.018)	(0.021)	(0.018)	(0.021)	(0.018)
Working status	0.892***	0.942***	0.894***	0.944***	0.896***	0.945***
	(0.013)	(0.010)	(0.013)	(0.010)	(0.013)	(0.010)
Panel B. Diagnoses						
Diagnosed with a disease	1.554***	-	1.548***	-	1.544***	-
	(0.054)	-	(0.053)	-	(0.052)	-
Stomach disease diagnosis	1.591***	-	1.594***	-	1.597***	-
	(0.074)	-	(0.074)	-	(0.074)	-
Breast disease diagnosis	1.210	-	1.180	-	1.170	-
	(0.297)	-	(0.263)	-	(0.254)	-
Cervical disease diagnosis	0.951	-	0.950	-	0.952	-
	(0.153)	-	(0.149)	-	(0.152)	-
Colorectal disease diagnosis	1.057	-	1.061	-	1.059	-
	(0.095)	-	(0.104)	-	(0.101)	-
Panel C. Health behaviors						
Current smoker	0.820***	0.643***	0.822***	0.644***	0.823***	0.644***
	(0.034)	(0.021)	(0.033)	(0.021)	(0.033)	(0.020)
Everyday smoker	0.832***	0.641***	0.834***	0.641***	0.835***	0.642***
	(0.035)	(0.021)	(0.034)	(0.021)	(0.034)	(0.021)
Current drinker	0.957***	0.995	0.957***	0.995	0.957***	0.996
	(0.012)	(0.009)	(0.012)	(0.009)	(0.012)	(0.009)
Everyday drinker	0.625***	0.534***	0.623***	0.534***	0.612***	0.520***
	(0.118)	(0.076)	(0.115)	(0.074)	(0.118)	(0.076)
Vigorous exercise	0.866***	1.017	0.871***	1.021	0.872***	1.021
	(0.031)	(0.024)	(0.030)	(0.023)	(0.030)	(0.023)
Moderate exercise	0.933***	1.093***	0.934***	1.094***	0.934***	1.094***
	(0.024)	(0.019)	(0.024)	(0.019)	(0.023)	(0.018)
Walking	0.994	1.043***	0.994	1.042***	0.994	1.042***
	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)	(0.008)

Notes: This table reports the ratio of socioeconomic status, screening diagnoses, and health behaviors between always-takers, never-takers and compliers using different interval length for age linear splines. Diagnoses are not reported for never-takers, since by definition, they do not receive screening. The null hypotheses used for ratios are $H_0 : \frac{y_{cp}}{y_{at}} = 1$ and $H_0 : \frac{y_{cp}}{y_{nt}} = 1$ for comparison with always-takers and never-takers, respectively, where at = Always-takers, nt = Never-takers and cp = Compliers. All the ratios are calculated at age 40. Standard errors are calculated using bootstrap with 500 replications and are clustered at individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A13: Bounding estimates for complier selection

	(1)	(2)	(3)	(4)	(5)	(6)
	Age ∈ [39,89]		Age ∈ [40, 89]		Age ∈ [41, 89]	
	CP/AT	CP/NT	CP/AT	CP/NT	CP/AT	CP/NT
Panel A. SES						
Individual income	0.552***	0.774***	0.568***	0.784***	0.536***	0.765***
	(0.020)	(0.023)	(0.023)	(0.025)	(0.023)	(0.025)
Household income	0.810***	1.030**	0.820***	1.034**	0.804***	1.026*
	(0.016)	(0.015)	(0.017)	(0.015)	(0.016)	(0.015)
Years of education	0.915***	1.014*	0.934***	1.022***	0.906***	1.012
	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)	(0.008)
College graduate	0.635***	0.898***	0.679***	0.926*	0.594***	0.853***
	(0.033)	(0.039)	(0.035)	(0.038)	(0.033)	(0.039)
Working status	0.810***	0.929***	0.819***	0.933***	0.803***	0.929***
	(0.014)	(0.013)	(0.015)	(0.013)	(0.014)	(0.013)
Panel B. Diagnoses						
Diagnosed with a disease	1.578***	-	1.556***	-	1.552***	-
	(0.052)	-	(0.052)	-	(0.053)	-
Stomach disease diagnosis	1.649***	-	1.629***	-	1.619***	-
	(0.075)	-	(0.077)	-	(0.073)	-
Breast disease diagnosis	1.182	-	1.165	-	1.163	-
	(0.283)	-	(0.276)	-	(0.306)	-
Cervical disease diagnosis	0.852	-	0.942	-	0.877	-
	(0.127)	-	(0.142)	-	(0.133)	-
Colorectal disease diagnosis	1.102	-	1.077	-	1.087	-
	(0.119)	-	(0.131)	-	(0.129)	-
Panel C. Health behaviors						
Current smoker	0.679***	0.517***	0.687***	0.519***	0.666***	0.515***
	(0.041)	(0.023)	(0.042)	(0.023)	(0.042)	(0.025)
Everyday smoker	0.696***	0.519***	0.704***	0.522***	0.683***	0.517***
	(0.043)	(0.024)	(0.043)	(0.023)	(0.044)	(0.025)
Current drinker	0.892***	0.985	0.910***	0.992	0.890***	0.990
	(0.015)	(0.013)	(0.015)	(0.013)	(0.016)	(0.013)
Everyday drinker	0.780***	0.710***	0.749***	0.697***	0.772***	0.702***
	(0.077)	(0.050)	(0.077)	(0.051)	(0.083)	(0.053)
Vigorous exercise	0.767***	1.012	0.775***	1.016	0.756***	1.012
	(0.035)	(0.034)	(0.035)	(0.033)	(0.034)	(0.033)
Moderate exercise	0.894***	1.122***	0.896***	1.123***	0.890***	1.125***
	(0.023)	(0.022)	(0.025)	(0.023)	(0.026)	(0.023)
Walking	0.996	1.055***	0.992	1.053***	0.995	1.055***
	(0.011)	(0.008)	(0.011)	(0.008)	(0.011)	(0.008)
N	110121	110121	107183	107183	104153	104153

Notes: This table reports the ratio of socioeconomic status, screening diagnoses, and health behaviors between always-takers, never-takers and compliers using samples with different starting age (39, 40, 41). It does not include any control variable of age. The null hypotheses used for ratios are $H_0 : \frac{y_{cp}}{y_{at}} = 1$ and $H_0 : \frac{y_{cp}}{y_{nt}} = 1$ for comparison with always-takers and never-takers, respectively, where at = Always-takers, nt = Never-takers and cp = Compliers. Standard errors are calculated using bootstrap with 500 replications and are clustered at individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A14: Robustness check for the effect of health screening

	(1)	(2)	(3)	(4)	(5)
	Interval 3	Interval 5	Interval 7	Full controls	Individual FE
Panel A. Health care utilizations					
Outpatient visit	0.5737	0.3494	0.3828	0.2874	0.0776
	(0.3858)	(0.3713)	(0.3701)	(0.3702)	(0.3410)
Inpatient visit	0.0300	0.0272	0.0268	0.0259	0.0236
	(0.0194)	(0.0192)	(0.0190)	(0.0192)	(0.0191)
Emergency visit	–0.0061	–0.0117	–0.0122	–0.0127	–0.0069
	(0.0127)	(0.0127)	(0.0128)	(0.0127)	(0.0124)
Panel B. First outpatient visits					
First outpatient visit	0.3006***	0.2976***	0.2906***	0.3055***	0.2845***
	(0.0764)	(0.0747)	(0.0746)	(0.0745)	(0.0718)
First outpatient visit for a cancer	0.0136**	0.0143***	0.0142***	0.0140**	0.0165***
	(0.0057)	(0.0055)	(0.0055)	(0.0055)	(0.0054)
Stomach cancer	0.0036	0.0037	0.0036	0.0037	0.0035
	(0.0023)	(0.0024)	(0.0023)	(0.0024)	(0.0023)
Breast cancer	0.0089**	0.0086**	0.0087**	0.0085**	0.0096**
	(0.0043)	(0.0042)	(0.0042)	(0.0041)	(0.0042)
Cervical cancer	–0.0005	–0.0007	–0.0008	–0.0007	–0.0006
	(0.0008)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Liver cancer	0.0047	0.0071	0.0073	0.0071	0.0068
	(0.0105)	(0.0104)	(0.0103)	(0.0105)	(0.0093)
Colorectal cancer	0.0159	0.0150	0.0156	0.0148	0.0166
	(0.0128)	(0.0126)	(0.0127)	(0.0126)	(0.0123)
Lung cancer	–0.0101	0.0011	0.0012	0.0002	0.0121
	(0.0710)	(0.0666)	(0.0655)	(0.0669)	(0.0628)
Prostate cancer	–0.1791*	–0.1525	–0.1581	–0.1533	–0.1258
	(0.1032)	(0.0994)	(0.1010)	(0.0999)	(0.0937)

Notes: This table reports the effect of health screening on health care utilizations with different control variables. Column 1 to 3 reports LATE estimates with the linear splines of age with 3, 5, and 7 years interval. Column 4 reports LATE estimates controlling for 5 year linear splines of age plus all the variables reported in the balance table (Table 2). Column 5 reports LATE estimates controlling for 5 year linear splines of age plus individual fixed effects. Panel A uses the number of outpatient, inpatient and emergency hospital visits except for screening purpose as outcome variables. Panel B uses the total number of first outpatient hospital visits and for specific cancers as outcome variables. First visit is inferred from the visit-level outpatient dataset asking whether the hospital visit was a first visit for a new illness or a recurring visit for an illness one is already aware of. The type of cancer is inferred from ICD-10 diagnosis code recorded for each visit to a hospital. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A15: Bounding estimates for the effect of health screening

	(1)	(2)	(3)
	Age ∈ [39, 89]	Age ∈ [40, 89]	Age ∈ [41, 89]
Panel A. Health care utilizations			
Outpatient visit	2.2332***	–1.2255***	2.5360***
	(0.3685)	(0.3760)	(0.3880)
Inpatient visit	0.0442**	0.0081	0.0477**
	(0.0184)	(0.0189)	(0.0194)
Emergency visit	–0.0032	–0.0168	–0.0039
	(0.0124)	(0.0128)	(0.0130)
Panel B. First outpatient visits			
First outpatient visit	0.4677***	0.1485**	0.5256***
	(0.0731)	(0.0746)	(0.0762)
First outpatient visit for a cancer	0.0155***	0.0122**	0.0165***
	(0.0053)	(0.0055)	(0.0056)
Stomach cancer	0.0038*	0.0032	0.0040*
	(0.0022)	(0.0023)	(0.0024)
Breast cancer	0.0092**	0.0087**	0.0099**
	(0.0041)	(0.0041)	(0.0043)
Cervical cancer	–0.0006	–0.0007	–0.0008
	(0.0009)	(0.0009)	(0.0009)
Liver cancer	0.0078	0.0061	0.0084
	(0.0101)	(0.0104)	(0.0106)
Colorectal cancer	0.0165	0.0137	0.0173
	(0.0122)	(0.0126)	(0.0127)
Lung cancer	0.0057	–0.0109	0.0068
	(0.0646)	(0.0663)	(0.0651)
Prostate cancer	–0.1519	–0.1772*	–0.1515
	(0.0975)	(0.1028)	(0.0979)
N	110121	107183	104153

Notes: This table reports the effect of health screening using samples with different starting age (39, 40, 41). It does not include any control variable. The coefficients report the LATE estimates of health screening instrumented by the biennial subsidy at even ages. Panel A uses the number of outpatient, inpatient and emergency hospital visits except for screening purpose as outcome variables. Panel B uses the total number of first outpatient hospital visits and for specific cancers as outcome variables. First visit is inferred from the visit-level outpatient dataset asking whether the hospital visit was a first visit for a new illness or a recurring visit for an illness one is already aware of. The type of cancer is inferred from ICD-10 diagnosis code recorded for each visit to a hospital. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Appendix G Additional analysis on selection into screening

This section presents additional analyses for the selection results. First, I present selection patterns at different ages. As discussed in section 5.4, I have estimated the characteristics at age 40. Table A16 additionally presents the selection pattern at ages 50 abnd 60. While the point estimates slightly vary, the results are qualitatively the same.

Next, I present selection pattern in terms of health care utilizations and expenditures. Panel A in Table A17 provides comparisons of hospital visits for outpatient, inpatient and emergency care. Since screening itself can affect health care utilizations, I compare always-takers with treated compliers and never-takers with untreated compliers. The untreated compliers have more outpatient visits than never-takers, but less hospitalizations. Treated compliers seem to have slightly more outpatient visits than always-takers. When I compare health care spending as shown in Panel B, untreated compliers spend more on outpatient care than never-takers, consistent with more outpatient visits.

Lastly, I compare the medical tests received during screenings between the always-takers and treated compliers. Note that since untreated compliers and never-takers do not participate, we have no information on their tests. Table A18 Panel A compares three kinds of tests that are covered by the National Health Insurance Service (NHIS). Since these are covered by the NHIS, the costs are either fully or heavily subsidized. Compared to always-takers, compliers are more likely to take them. However, tests not covered by the NHIS exhibit the opposite pattern. Panel B shows that compliers are less likely to use the tests that are not covered by the NHIS. This is consistent with the selection pattern that compliers have lower income and hence less likely to use expensive tests that are not covered by the insurance.

Table A16: Complier selection at different ages

	(1)	(2)	(3)	(4)	(5)	(6)
	Age 40		Age 50		Age 60	
	CP/AT	CP/NT	CP/AT	CP/NT	CP/AT	CP/NT
Panel A. SES						
Individual income	0.758***	0.899***	0.759***	0.890***	0.587***	0.767***
	(0.019)	(0.015)	(0.018)	(0.015)	(0.029)	(0.029)
Household income	0.900***	1.051***	0.900***	1.040***	0.873***	1.039***
	(0.014)	(0.011)	(0.014)	(0.011)	(0.018)	(0.015)
Years of education	0.981***	1.023***	0.977***	1.024***	0.970***	1.028***
	(0.006)	(0.004)	(0.007)	(0.005)	(0.009)	(0.007)
College graduate	0.861***	1.001	0.816***	0.987	0.593***	0.923
	(0.021)	(0.018)	(0.028)	(0.025)	(0.057)	(0.072)
Working status	0.894***	0.944***	0.893***	0.940***	0.865***	0.919***
	(0.013)	(0.010)	(0.013)	(0.010)	(0.016)	(0.013)
Panel B. Diagnoses						
Diagnosed with a disease	1.548***	-	1.570***	-	1.610***	-
	(0.053)	-	(0.053)	-	(0.057)	-
Stomach disease diagnosis	1.594***	-	1.628***	-	1.694***	-
	(0.074)	-	(0.078)	-	(0.085)	-
Breast disease diagnosis	1.180	-	1.200	-	1.240	-
	(0.263)	-	(0.299)	-	(0.380)	-
Cervical disease diagnosis	0.950	-	0.948	-	0.945	-
	(0.149)	-	(0.156)	-	(0.165)	-
Colorectal disease diagnosis	1.061	-	1.087	-	1.096	-
	(0.104)	-	(0.148)	-	(0.166)	-
Panel C. Health behaviors						
Current smoker	0.822***	0.644***	0.805***	0.627***	0.672***	0.465***
	(0.033)	(0.021)	(0.036)	(0.022)	(0.057)	(0.033)
Everyday smoker	0.834***	0.641***	0.822***	0.632***	0.695***	0.469***
	(0.034)	(0.021)	(0.036)	(0.022)	(0.060)	(0.033)
Current drinker	0.957***	0.995	0.951***	0.992	0.937***	0.986
	(0.012)	(0.009)	(0.013)	(0.010)	(0.017)	(0.013)
Everyday drinker	0.623***	0.534***	0.754***	0.684***	0.760***	0.695***
	(0.115)	(0.074)	(0.078)	(0.051)	(0.077)	(0.052)
Vigorous exercise	0.871***	1.021	0.874***	1.010	0.823***	0.999
	(0.030)	(0.023)	(0.029)	(0.022)	(0.040)	(0.034)
Moderate exercise	0.934***	1.094***	0.935***	1.080***	0.921***	1.086***
	(0.024)	(0.019)	(0.023)	(0.018)	(0.027)	(0.023)
Walking	0.994	1.042***	0.994	1.040***	0.993	1.039***
	(0.011)	(0.008)	(0.011)	(0.008)	(0.012)	(0.009)

Notes: This table reports the ratio of socioeconomic status, screening diagnoses, and health behaviors between always-takers, never-takers and compliers at different ages (40, 50, 60). The null hypotheses used for ratios are $H_0 : \frac{y_{cp}}{y_{at}} = 1$ and $H_0 : \frac{y_{cp}}{y_{nt}} = 1$ for comparison with always-takers and never-takers, respectively, where at = Always-takers, nt = Never-takers and cp = Compliers. Standard errors are calculated using bootstrap with 500 replications and are clustered at individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A17: Comparing health care utilizations and expenditures

	(1)	(2)	(3)	(4)	(5)	(6)
	Average value				Ratio	
	Always-takers	Treated compliers	Untreated compliers	Never-takers	CP_1/AT	CP_0/NT
Panel A. Health care utilizations						
Outpatient visit	9.098	10.263	10.007	6.532	1.128**	1.532***
	(0.360)	(0.350)	(0.482)	(0.213)	(0.058)	(0.087)
Inpatient visit	0.089	0.085	0.057	0.117	0.954	0.481***
	(0.012)	(0.011)	(0.020)	(0.009)	(0.167)	(0.176)
Emergency visit	1.221	1.211	1.356	1.332	0.992	1.018
	(0.055)	(0.064)	(0.096)	(0.050)	(0.043)	(0.068)
Panel B. Health care expenditures						
Outpatient expenditure	294931	280914	301873	161079	0.952	1.874***
	(13122)	(13001)	(19773)	(7445)	(0.064)	(0.147)
Inpatient expenditure	67598	75791	97104	79261	1.121	1.225
	(15311)	(13440)	(28395)	(9612)	(0.347)	(0.400)
Emergency expenditure	3467	4475	4642	3811	1.291	1.218
	(713)	(821)	(1742)	(584)	(0.365)	(0.539)

Notes: This table reports health care utilizations and expenditures among always-takers, treated compliers, untreated compliers and never-takers. Treated compliers are compliers in the treatment group who participate in screening. Untreated compliers are compliers in the control group who do not participate. Panel A presents the number of hospital visits for outpatient, inpatient and emergency care. Panel B presents the expenditures incurred for outpatient, inpatient and emergency care. The unit of currency is Korean Won. The average value is calculated using Equation (3). The null hypotheses used for ratios are $H_0 : CP_1/AT = 1$ and $H_0 : CP_0/NT = 1$ for comparison with always-takers and never-takers, respectively, where AT = Always-takers, NT = Never-takers, CP_1 = Treated compliers and CP_0 = Untreated compliers. All the average values and ratios are calculated at age 40. Standard errors are calculated using bootstrap with 500 replications and are clustered at individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A18: Comparing medical tests received

	(1)	(2)	(3)
	Average value		Ratio
	Always-takers	Treated Compliers	CP_1/AT
Panel A. Tests covered by NHIS			
Blood/urine/stool/X-ray	0.801	0.886	1.106***
	(0.005)	(0.006)	(0.012)
Endoscopy	0.574	0.776	1.352***
	(0.007)	(0.008)	(0.024)
Biopsy	0.026	0.031	1.190
	(0.002)	(0.002)	(0.153)
Panel B. Tests not covered by NHIS			
Sonogram	0.319	0.275	0.863***
	(0.007)	(0.007)	(0.031)
CT	0.042	0.016	0.374***
	(0.003)	(0.002)	(0.071)
MRI	0.010	0.006	0.593**
	(0.001)	(0.001)	(0.180)
PET	0.002	–0.000	–0.229***
	(0.001)	(0.000)	(0.266)
EEG	0.002	0.002	0.743
	(0.001)	(0.001)	(0.618)
EKG	0.159	0.127	0.794***
	(0.005)	(0.005)	(0.048)
Bone density	0.033	0.029	0.857
	(0.003)	(0.003)	(0.126)

Notes: This table reports the receipt of medical tests among always-takers and treated compliers. Panel A presents tests that are covered by the National Health Insurance Service (NHIS). Blood/urine/stool/X-ray tests are main tests of general screening. Endoscopy is used in stomach and colorectal screenings. Biopsy are used in various screenings, normally after simpler diagnostic tests, for closer examination of tissues. Panel B presents tests that are not covered by the NHIS, and hence, screening participants should bear the full cost of the tests. The average value is calculated using Equation (3). It is not reported for never-takers and untreated compliers, since by definition, they do not receive screening. The null hypotheses used for ratios are $H_0 : CP_1/AT = 1$ where AT = Always-takers and CP_1 = Treated compliers. Standard errors are calculated using bootstrap with 500 replications and are clustered at individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.