

Health screening and selection: Evidence from biennial subsidies in South Korea*

Sihoh Park[†]

January 23, 2024

How can policies better target those most likely to benefit from health screening? 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 Screening Program. The biennial subsidies provided at even ages increase the take-up and produce spillover in take-up across different types of screenings and between spouses. Compliers to screening subsidies have lower income than always-takers and are hence more likely to find a disease through screening. Lastly, Health screenings induced by subsidies increase outpatient hospital visits for a new illness.

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

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

1 Introduction

The primary goal of health screening is to detect diseases early. Diagnosing diseases at an early stage can lead to treatments when they are more likely to be successful, reducing the risk of premature death and health care costs.¹ The effectiveness of health screening in enhancing population health depends on its ability to target individuals with underlying health issues who may be unaware of their conditions. However, empirical evidence often shows that participants in health screenings tend to be healthier, while those who are unhealthier and more likely to have undiagnosed diseases frequently do not participate.² (Pill et al., 1988; Waller et al., 1990; Strandberg et al., 1995; Bender et al., 2015; Jones et al., 2019). This mismatch between those who would benefit most from health screening and those who actually participate is also evident in other forms of preventive health care, such as vaccination, contraceptive use, and smoking cessation. (Thomas et al., 2021; Hungerford et al., 2016; Gafar et al., 2020; Wetter et al., 2005).

Policies can be designed to target specific group, thereby changing the composition of participants and some policies may perform better than others. Einav et al. (2020) and Kowalski (2023) examine mammogram recommendations in the US and Canada and find that marginal participants are healthier than people participating without recommendation. However, policies that directly appeal to less healthy individuals could be more effective in engaging those more likely to have a disease. One such policy is the provision of subsidies for health screenings. Low income individuals, who are often more responsive to financial incentives, would be more likely to respond to subsidies and participate. To the extent that lower income is correlated with poorer health conditions, this approach could lead to increased participation from those who are less healthy, aligning

¹5 year survival rates for various cancers significantly drop as the stage of cancer progresses (Howlader et al., 2021). Cutler (2008) argue that early detection and treatment of cancers through screenings made the most important and cost-effective contribution in the war on cancer since 1990 in the US.

²This pattern of selection is also often cited as the reason why many studies fail to find a significant effect of screening on mortality and morbidity. 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 RCTs concludes with high certainty that general screenings have little or no effect on future mortality and morbidity (Krogsbøll et al., 2012). The review suggests that this may be due to the frequent participation of healthier individuals. Another example is the decision by the United States Preventive Services Taskforce (USPSTF) to delay the recommended starting age for mammograms from 40 to 50 years, influenced by similar concerns.

more closely with the objectives of health screening programs.

This paper investigates three research questions. First, what is the effect of subsidies on health screening participation? Second, how does the selection pattern change in response to 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 and health care utilizations.

To answer these question, I use the research design that exploits biennial subsidies provided by the Korean National Health Insurance Service (NHIS).³ 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. Subsidies are provided from age 40, with the exception of cervical and colorectal screenings, for which subsidies start at ages 30 and 50, respectively. The biennial subsidies are provided in a year when one's age is even numbered.⁴ I show this subsidy rule is random conditional on age and creates large variation in screening take-up between the even age group, eligible for subsidies, and the odd age group, ineligible for subsidies.

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. This widens the take-up gap between the even and odd age group. To provide the effect of subsidies not affected by intertemporal substitution, I bin ages by 2 years and measure the discontinuous increase in 2-year average take-up at the subsidy starting age. It allows abstracting away from even-odd design and it can also be interpreted as a lower bound for the one year subsidy effect.

³I use the word ‘biennial’ in the sense that it is provided once every two years, now twice a year.

⁴The 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.

This study uses the Korean Health Panel Study datasets that include 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.⁵

This enables me to observe the exact date, type and the results of screenings.

This paper presents four 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 and they show large jump in 2-year average take-up at each subsidy starting age and at age 40. I examine 3 subsamples on different margins of biennial subsidy design to investigate intertemporal reallocation of screenings and they all suggest no strong sign of substitution.

Second, the biennial subsidies also generate two types of spillover. There is cross spillover across different types of screenings within an individual. 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 exhibit significantly higher take-up at even ages. Cervical and colorectal screening, despite their subsidy starting age being 30 and 50, respectively, show jump in take-up at age 40. This is due to the tendency to receive multiple screenings together on the same day. There is also positive spillover between spouses. One's subsidy eligibility not only increases one's own likelihood of screening, but also boosts the chances of spouse receiving screening. The spousal spillover is more pronounced from husband to wife than the opposite direction and it can also be explained by the couple's tendency to receive screening on the same day together.

Next, selection analysis reveals that compliers with subsidies have worse health conditions and are more likely to be diagnosed with a disease through screening compared to always-takers. This is in contrast to the finding of [Einav et al. \(2020\)](#) that US mammogram recommendation at age 40 induces participation of healthier compliers with the

⁵This was made possible through survey participants recording specifically designed health diaries and keeping receipts from every visit to hospitals. See Section 4 for detail.

lower rate of true positive breast cancer. One reason is that compliers are negatively selected on income. I show that providing direct monetary incentives attract compliers from lower socioeconomic background with lower income and less education. Due to strong income effect, compliers are less likely to smoke, drink and exercise (Cawley and Ruhm, 2011; Gallet and List, 2003; Gallet, 2007; Apouey and Clark, 2015; Armstrong et al., 2018; Thibaut et al., 2017). This suggests that providing subsidies for health screening better targets people with low income and worse health conditions. Compared to never-takers, I find compliers are positively selected in health behaviors, consistent with correlation in various positive health behaviors (Oster, 2020; Cutler and Lleras-Muney, 2010). Unlike selection in response to subsidies, the cross spillover compliers exhibit opposite pattern. The biennial screening participants who further go on to receive annual or no-subsidy screenings are healthier individuals less likely to find a disease and have higher income and education level than those who do not.

Lastly, I find screenings induced by subsidies lead to an increase in outpatient care usage, especially hospital visits for a new illness. Despite the information on screening results, it cannot be used to estimate the effect of screening on new diagnoses, since diagnoses are unobserved for nonparticipants. To overcome this econometric challenge, I infer diagnoses from outpatient care data that recorded hospital visits regardless of participation in screening. For each hospital visit, the dataset contain information on the reason of visit coded using the Korean version of the 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 estimate the impact on first hospital visits for various screenings. Using subsidy eligibility as an instrument, I find that screening leads to 9.2 percent significant increase in first hospital visits for a new illness.

The most directly related strand of literature is selection in health screening. Compared to Einav et al. (2020) that examine selection in response to mammogram age recommendation at 40, this study considers different treatment: subsidies for various screenings. While age-based recommendation attracts compliers healthier than always-

takers, the subsidies induce participation of less healthy compliers more likely to find a disease. While there are variations across different aspects, the different selection appears to be driven by the nature of the treatments themselves; subsidies tend to attract lower income individuals who, on average, are often in poorer health. This suggests selection patterns vary with treatments and a policy can be designed to better target those who are more likely to benefit from screening. Another related study is [Kim and Lee \(2017\)](#) that examine subsidies for cancer screenings at income cutoffs and find consistent selection pattern that compliers have worse health conditions than always-takers. Unlike [Kim and Lee \(2017\)](#), this study uses detailed socioeconomic information that allows richer characterization of compliers. It sheds light on the negative selection of income that provides a mechanism through which subsidies better target unhealthy individuals more likely to benefit from screening.

More broadly, it contributes to the study of various health screening policies and their impact on diagnoses and behaviors. Previous studies examined age recommendation, health insurance mandate, 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. Compared to these studies that normally consider one or limited set of screenings, I consider various screenings altogether which allows me to find cross spillover across different types of screenings.⁶ This indicates that screening policies could be strategically designed to utilize cross spillover effects, thereby maximizing participation. It also suggests that policies for various screenings should be developed in conjunction, rather than in isolation. Analyzing the impact of screening on diagnoses presents challenges not only in econometric analysis but also in the availability of diagnostic data. Since diagnoses are only recorded conditional on getting screened, [Einav et al. \(2020\)](#) and [Kowalski \(2023\)](#) rely on cancer registry data that contain all the cancer diagnoses. This approach, however, overlooks less severe symptoms that, while not as critical as cancers, are still

⁶[Bitler and Carpenter \(2016\)](#) also find increase in cervical screening take-up in response to mammogram mandates.

important for quality of life. Utilizing detailed survey data that provides diagnosis for each hospital visit allows the identification of milder symptoms and conditions that may act as a precursor to more serious illnesses, such as Helicobacter pylori infection which can lead to stomach cancer.⁷

Lastly, it relates to the literature on peer effects in health behaviors. People rarely make decisions in isolation. Instead, they often do so with others. There are numerous ways in 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 in health screening.

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 a nationwide program that encourages health screenings for all the Korean citizens through subsidies, recommendations and various policy tools. The program is composed of 3 types of screenings; 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 gen-

⁷Another 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 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.

eral 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 additionally considers lung and prostate cancer screenings that are not subsidized, but captured in the dataset.⁸ The subsidized medical test for stomach screening is gastroscopy, inserting flexible tubes with a camera through mouth to examine throat, oesophagus and stomach. Breast screening uses X-ray mammography for both breasts and cervical screening uses cervical cytology (also called a Pap test). Liver screening was subsidized only for the high risk group.⁹ Non-high risk group can pay the full cost to receive liver screening. The subsidized test is liver ultrasound or Maternal Serum Alpha-Fetoprotein Screening (MSAFP), which can be checked from the blood sample. The subsidized test for colorectal screening is fecal immunochemical test (FIT) which checks for hidden blood in the stool sample. If one is tested positive from FIT, then colonoscopy is subsidized, after which any colorectal screening test is not subsidized for the following 5 years. Receiving colonoscopy without going through FIT is not subsidized and one would need to pay the full cost. My data show the majority of colorectal screening participants actually receive colonoscopy directly which suggest not utilizing the subsidy for FIT. This is due to low perceived risk of colonoscopy procedures and low cost in Korea that makes it easily accessible ([Baik and Lee, 2023](#)).

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 one subsidized screening at any time during the even (odd)-numbered

⁸Lung cancer screening began to be subsidized right after the study period in 2019.

⁹Liver screening was subsidized only for high risk group and was subsidized up to twice a year. My dataset does not have enough information to accurately identify the high risk group. Throughout the analysis, I treat it as if everyone is subsidized, but one should keep in mind it is only for the small number of high risk group. The high risk group consists of (*i*) those with cirrhosis and chronic liver disease and (*ii*) those who were diagnosed positive in hepatitis B surface antigen test or hepatitis C virus HCV antibody test.

calendar years. For instance, a woman born in 1968, an even birth year, is entitled to get a subsidized stomach screening any time during 2020 and 2022, even years, but one must pay the full cost for the same stomach screening in 2023, an odd year. If I define age as the difference between the current year and the year of birth, that is, current year - year of birth, then one is entitled to subsidized screenings in 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.¹⁰ 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

¹⁰Liver screening was subsidized up to twice a year.

¹¹Colorectal screening used to be biennially subsidized until 2011. It became annually subsidized in 2012.

¹²Starting 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. I examine the effect of this policy change in Appendix Section A.

¹³There are 2 exceptions to the subsidy rules for the general health screening. First, one can be entitled to biennial subsidy for general screening before age 40 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 from age 40, not every other year. Note that these rules apply only to the general health screening, not any other cancer screenings. My data show that these two rules are not strictly followed. Therefore, this study abstracts away from these exceptions.

screening, but needs to pay 10% of the costs for cancer screenings.¹⁴ For those with low income, the 10% copays are also subsidized, making cancer screenings free.¹⁵ The rough copay amount is \$0 for fully subsidized general screening and \$7 for stomach screening.¹⁶

The program was designed to provide easy accessibility for the goal of inducing maximum participation. A public health screening program, initially targeting only public servants and private school staffs, started in 1980s and gradually expanded beneficiaries (Kang, 2022). In our study period, all the Korean citizens were covered by the program and was eligible for subsidies from certain ages without exception.¹⁷ People can receive subsidized screenings at either public health clinics or private clinics or hospitals designated by the NHIS.¹⁸ Appointment is normally not required for the simple general screening, but is needed for more involved screenings. Reminder paper mails, and more recently mobile notifications, are sent to people eligible for subsidies and they include the types of screenings to receive and available screening centers in the neighborhood.

3 Identification

This study exploits two sources of quasi-random variation in screening take-up generated by age cutoffs and biennial provision of subsidies. Focusing on age 40, I measure increase in participation at age cutoffs as screenings are recommended with subsidies. Next, I compare take-up between even ages who receive subsidies and odd ages who do not as people additionally participate due to subsidies or move screening from odd to even ages to benefit from subsidies.

Figure 1a motivates the identification strategy. It plots the average take-up of general screening at each age that is biennially subsidized from age 40. The take-up pattern

¹⁴The 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.

¹⁵Kim and Lee (2017) exploits the income cutoff for free cancer screenings to examine the effect of copays on take-up and cancer detection.

¹⁶Depending 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.

¹⁷My study period includes neither the introduction of the program nor expansion of the target. Hence, I use the variation in screening eligibility in the current program, not the change in eligibility over time.

¹⁸The list of private screening centers designated by the NHIS also gradually expanded. In December 2023, there were approximately 5,800 private clinics and hospitals designated by the NHIS to provide health screenings, which translates into 900 people with age 40 or above per center.

closely follows the subsidy schedule. Before age 40, there is negligible difference in screening rates between the even age group eligible for subsidies and the odd age group ineligible for subsidies. However, with the subsidies kicking in from 40, the screening rate jumps from 9 percent to 27 percent at even ages, while the screening rate at odd ages remains around 9 percent. The gap between the even and odd age groups persists until late 80s.

The observed differences in screening uptake between even and odd ages can be attributed to two main factors. First, the provision of subsidies at even ages encourages new individuals to participate in screenings. Second, there's a potential trend where individuals shift their screening schedules from odd to even ages to qualify for these subsidies. Aware of the financial benefits available at even ages, individuals might choose to either advance their screening by a year or postpone it to align with these subsidy periods. While this intertemporal substitution contributes to a larger discrepancy in uptake between even and odd ages, it's important to note that this is primarily a shift in the timing of screenings. It does not necessarily represent a net increase in overall screening participation.

To abstract away from intertemporal substitution and to focus on discontinuity at age 40, I bin ages by 2 years and plot the average take-up at each 2-year age bin as shown in Figure 1b. Since intertemporal substitution entails reallocating screening from odd to even years, taking average between the even and odd ages cancels out the substitution effect. The resulting jump in 2-year average take-up at age 40 reflects the effect of subsidies and the recommendation to start screening from age 40. The econometric specification is given as follows.

$$y_{it} = \alpha_0 + \alpha_1 \cdot a_{it} + \alpha_2 \cdot \mathbb{1}\{a_{it} > 0\} + \alpha_3 \cdot a_{it} \times \mathbb{1}\{a_{it} > 0\} + \varepsilon_{it} \quad (1)$$

The centered age variable is given as $a_{it} = (agebin_{it} - 39.5)$. Ages are binned in 2 years and we use samples 6 years before and after age 40. Hence, the age bins used are [34, 35], [36, 37], [38, 39], [40, 41], [42, 43], [44, 45] and the variable $agebin_{it}$ refers to the midpoint of each bin, hence, 34.5, 36.5, ..., 42.5, 44.5.¹⁹ The outcome variable y_{it} denotes

¹⁹The reason I center age bins around 39.5 instead of 40 is that the midpoint between the two bins

the screening take-up for individual i in year t . The estimate of interest, $\hat{\alpha}_2$, captures the jump in take-up at the cutoff and the constant term, $\hat{\alpha}_0$, captures the take-up right before the jump.

The second variation in take-up comes from biennial subsidies from age 40. While the distinction between even and odd ages seems plausibly exogenous, the even age group (treatment group) is younger than the odd age group (control group) by design. This is a mechanical effect due to the subsidy design that starts to provide subsidies at age 40, an even number.²⁰ I argue that age is the only difference between the treatment and the control group, so the two groups should be balanced after controlling for age. This motivates the following econometric specification.

$$y_{it} = \beta_0 + \beta_1 \cdot \text{age_even}_{it} + \mathbf{f}(\mathbf{age}_{it}) + \epsilon_{it} \quad (2)$$

The treatment group indicator variable is age_even_{it} that equals 1 when individual i has even age in year t . $\mathbf{f}(\mathbf{age}_{it})$ 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 G.

Table 2 shows the conditional difference ($\hat{\beta}_1$) between the treatment and the control group after including the linear splines of age with 5 years interval as control variables. The analytical sample is those with age from 40 to 89. First, note that the two groups are almost equally sized. Column 3 shows the differences between the two groups are small and statistically insignificant conditional on the function of age. Hence, we can attribute the difference in take-up between the treatment and the control group as the causal effect of screening subsidies.

While the research design only uses cross sectional variation, I use panel data. Hence, the standard errors in both specifications will be clustered at the individual level to ac-

[38, 39] and [40, 41] is 39.5.

²⁰Given 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.

count for the fact that same individuals appear multiple times in different years. 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 G, I present robustness checks where I include individual fixed effects when comparing even and odd age groups.

4 Data

This study utilizes the collection of Korean Health Panel Study datasets spanning the years 2008-2018.²¹ 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. 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. At every visit, they started with hospital visits from the last day of interview so that there is no missing period. For each visit to a hospital, the dataset contains the date of visit, hospital bills and drug expenditures incurred, diagnosis and whether the

²¹It is version 1.7.1 made jointly by Korean Institute for Health and Social Affairs (KIHSA) and National Health Insurance Services (NHIS).

hospital visit was a first visit for a new illness, second visit or a recurring visit for an illness one is already aware of.²²

Records on health screenings are also from the outpatient care dataset. Hospital visits for health screenings included additional information about the type of screening received (general or specific type of cancer screenings), medical tests performed, screening results and diagnosis 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 in addition to whether one engages in such activity. For smoking and drinking, I consider current drinker or current 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 interview.

5 Results

5.1 Effect of screening subsidies on take-up

Figure 1b, 1d, 1f and 1h present average 2-year screening rates for 4 types of biennial screenings. Binning ages by 2 years remove the variation between the even and odd ages and all figures show large increase in take-up at age 40 as subsidies kick in. Cervical screening is subsidized from age 30 and one can see a small increase at age 30, but it is considerably smaller than the one at age 40. I estimate the discontinuous jump at age 40 using Equation (1). Table 3 presents the estimation results. The 2-year average take-up in any of the biennial screenings jump by 10 percentage point (80 percent) from 12 percent at age 40. The increase in 2-year average take-up can also be interpreted as the lower bound of the one year subsidy effect locally around age 40. This is because we

²²Diagnoses were coded using Korean Classification of Diseases diagnosis code (Korean version of ICD-10).

are treating the entire variation between even and odd ages as the substitution effect by taking the average; all the difference in take-up between the two groups is due to people moving their screening timing. The resulting jump in take-up is the minimum subsidy effect after assuming maximum substitution effect.

The screening take-up for each age is shown in Figure 1a, 1c, 1e and 1g. There is a large divergence in screening rate between the treatment and the control group after the cutoff age. Using Equation (2), 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 4 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 indicate strong first stage ([Angrist and Pischke, 2009](#); [Bound et al., 1995](#)). The difference between the even and odd ages can be interpreted as the upper bound of the subsidy effect throughout the age range [40, 89]. We are assuming the lack of substitution and treating the entire variation between the two groups as subsidy effect. Hence, the gap in take-up is the maximum subsidy effect after assuming minimal substitution.

An important mechanism through which subsidies affect participation is intertemporal substitution. The knowledge of biennial subsidies from age 40 allows forward-looking individuals to temporally allocate screenings to take full advantage of the financial incentives. People can either delay or move screenings to an earlier date so as to receive them when they are entitled to subsidies. Disentangling the substitution effect is challenging due to the lack of valid counterfactuals which would be a recommendation for biennial screening from age 40 without subsidies. In Appendix section A, I present empirical evidence of substitution among subsamples on different margins. First, focusing on individuals whose age spans before and after 40, I investigate change in screening behavior as subsidies start to apply. Second, I examine monthly distribution of screening take-up, especially the beginning and end months of odd age, as these are the time when

substituting behaviors should be most pronounced, if they exist. Both pieces of evidence suggest there is no strong sign of substitution, implying the bigger impact of subsidies on inducing new participation.

An interesting age pattern of screening rates is that the take-up for all types of screenings 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.²³ In case of mammogram, the USPSTF guideline does not make any recommendation of (nor discourage) mammogram for women aged 75 or above due to insufficient evidence to weigh benefit against harm. 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.²⁴ Note that for our screening subsidies, there is no upper bound in age.

5.2 Cross spillover across different types of screenings

This section investigates spillover in screening take-up within an individual across different types of screenings. I show that annual and no-subsidy screenings display the same biennial pattern in take-up from age 40 following the recommended take-up pattern for general screening. This can be explained by the tendency to receive multiple screenings on the same day or to receive annual and no-subsidy screenings right after the general health screening.

Liver and colorectal screenings were subsidized every year. Annual subsidies imply there is no reason to expect systematic difference in take-up between even and odd ages. However, take-up patterns shown in Figure 2a and 2c clearly indicate much higher take-

²³Refer 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.

²⁴As people live longer and longer, the need to provide preventive screenings for older adults may become stronger. The 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).

up at even ages. Liver screening was subsidized twice a year from age 40. Figure 2b shows there is a large jump in take-up at age 40 and Figure 2a shows higher take-up at even ages than at odd ages. Colorectal screening was subsidized once a year from age 50. Figure 2d shows two discontinuous jump in take-up at age 40 and 50. Figure 2c also presents larger take-up at even ages than at odd ages starting from age 40. These are due to the fact that most biennial screenings have subsidies starting from age 40. The average rate of take-up of 4 biennial screenings is much larger than that of liver and colorectal screenings, which indicate that liver and colorectal screenings are secondary, less commonly received screenings. Hence, following biennial rules, liver and colorectal screening take-up also starts to increase at even ages from age 40, followed by another jump at 50 with its own subsidies in case of colorectal screening. Column 1 and 2 in Table 5 present estimates of increase in take-up at even ages. Liver screening shows 2.7 percentage point (96 percent) higher take-up at even ages. Colorectal screening shows 1.7 percentage point (63 percent) higher take-up at ages 40 to 49 and 4 percentage point (148 percent) higher take-up at ages 50 to 89.

Prostate and lung screenings were not subsidized at any year during the study period. This also implies there should be no discontinuous increase in take-up at age 40 and no systematic difference in take-up between even and odd ages. Figure 2f and 2h show there is no discernible jump in take-up at age 40, but Figure 2e and 2g show that take-up is generally higher at even ages and the divergence seems to start around age 40. Column 3 and 4 in Table 5 show the even group take-up is higher by 0.7 percentage point (78 percent) and 0.6 percentage point (67 percent) compared to the odd group take-up at 0.9 percent.

Whether the larger take-up at even ages represents positive or negative spillover depends on whether the screening concerned is subsidized or not. For high risk group in liver screening, subsidies are provided at every age from 40, so the reasonable reference point is permanent increase in take-up from age 40. However, the take-up at odd ages remain constant after subsidies start to apply suggesting under-utilization of subsidies at odd ages. This represents negative spillover from biennial screenings. For non-high risk

group, they are paying the full cost of liver screening and the reference point is no increase in take-up at age 40. The fact that take-up jumps at even ages from 40 represents positive spillover from biennial screenings.²⁵

Similarly for the colorectal screening, it is not subsidized at ages 40 to 49. The large increase in take-up at even ages represent positive spillover. From age 50, if the majority of testing used the fecal occult blood test (FOBT) which is subsidized every year, then the lower take-up at odd ages represent until-utilization of subsidies. However, my data shows the majority of colorectal screening participants were using colonoscopy. If these were colonoscopies without positive result from FOBT, then the tests are not subsidized and the take-up gap represents positive spillover.²⁶

Both the prostate and lung screenings were not subsidized. Hence, the larger take-up at even ages suggest there was positive spillover from biennial screenings. The spillover in take-up can also be found in the cervical screening. Its biennial subsidies began at age 30 which was lowered to 20 in 2016. Figure 1g shows that subsidies at even ages from 30 to 39 are under-utilized compared to after 40. This is because it is the only subsidized screening in this age range. Only a small number of people use the subsidies and the majority of people do not bother to come to hospital only for one type of screening. Table 5 shows the gap between even and odd ages increase by 12.8 percentage point after 40.

The overall spillover pattern suggests that all the screenings follow biennial take-up pattern from age 40. Bitler and Carpenter (2016) find similar cross spillover in mammogram mandate that led to increase in not only mammogram take-up, but also clinical breast exams and cervical screening. One mechanism of cross spillover is receiving multiple screenings at the same time in one hospital visit. Table A2 in Appendix section B presents the share of people who receive screenings on the same day with general screening, not just in the same year. Conditional on getting general screening, more than 85 percent of liver, colorectal, lung and prostate screening participants receive the screening on the same day with general screening. This could be due to some fixed costs in vis-

²⁵The dataset does not allow me to decompose the liver screening participants into high risk group for whom the screening is subsidized or non-high risk group who pays the full cost.

²⁶The dataset contains information on whether colonoscopy was used but not fecal occult blood test.

iting a hospital, such as psychological toll or travel costs. Or it could be that receiving general screening, the most basic screening, leads to take-up in other cancer screenings due to doctor's recommendation. Table A2 further shows that except colorectal screening, if liver, lung and prostate screenings are not received on the same day with general screening, they are mostly received in 30 days after the general screening.

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 in screening take-up

This section examines spillover in screening take-up between spouses. Using household members relationship information, I show that subsidy eligibility increases not only one's own screening participation rate, but also the spouse's participation. Instrumenting the spouse's participation with the spouse's subsidy eligibility, I estimate the peer effect in screening take-up.

The analytical sample is adjusted to currently married couples both of whose ages are 40 or above. 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 where the first even (odd) means one's own age is even (odd) and the second even (odd) means the spouse's age is even (odd). Equation (3) examines their participation rate.

$$y_{it} = \gamma_0 + \gamma_1 \cdot \text{age_even}_{it} + \gamma_2 \cdot \text{spouse_age_even}_{it} + \gamma_3 \cdot \text{age_even}_{it} \times \text{spouse_age_even}_{it} + \phi_{it} \quad (3)$$

The outcome variable, y_{it} , is screening take-up of individual i in year t . Our parameter of interest is γ_2 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, γ_3 , 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.

Figure 3 plots the participation rate in any type of screening for 4 types of couples. Red lines are when one's age is even and blue lines are when one's age is odd. Solid lines are when the spouse's age is even and the dashed lines are when the spouse's age is odd. Red lines exhibit significant increase at age 40, meaning one's participation rate rises as subsidies begin to apply. In both cases when one's age is even or odd (red or blue), the solid lines are above the dashed lines, meaning one is more likely to participate when the spouse has even age and hence is eligible for subsidies.

Table 6 estimates the positive spousal spillover in 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 (γ_1) and the spouse's subsidy eligibility (γ_2). The effect of subsidies measured in the married group is around 21 percentage point and is quite close to the one estimated using the whole sample in Table 4. The effect of spouse's subsidy eligibility is 1.6 percentage point and significant. However, there is no interaction effect (γ_3). Spouse's subsidy eligibility increases one's own take-up by the same magnitude regardless of whether one is eligible or ineligible for subsidies. To account for the possibility that 4 couple types by their age combinations might be correlated with other factors that can affect screening participation, I control for their characteristics. Column 2 presents estimates controlling for one's own and the spouse's demographic characteristics, and the results are robust.

To translate the subsidy effect into peer effect in screening, I instrument the spouse's screening participation with the spouse's subsidy eligibility. Column 3 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 7.8 percentage point. This corresponds to about one third of one's own subsidy effects. Column 4 show the estimates are robust to adding demographic control variables.

Appendix section C presents further analyses on potential mechanisms and heterogeneity in spousal spillover. Table A5 reveals the spousal spillover is pronounced from wives to husbands, but muted from husbands to wives. Husbands are more likely to

participate when wives are eligible for subsidies, but not vice versa. Table A6 presents evidence that taking screening together is one mechanism of spousal spillover. Conditional on both husband’s and wife’s participation, more than 40 percent receive screenings on the same day. The fact that spouses receive screenings on the same day can explain insignificant spousal spillover in breast, cervical and prostate screenings as shown in Table A7. These screenings are for women or men only. Compared to stomach or colorectal screening where both can participate together, they are less likely to be received during the same hospital visit.

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 and overdiagnoses (Rubin, 2019; Kowalski, 2023). What matters in terms of policy effectiveness is the marginal individuals who respond to screening subsidies and participate. They are the ones who would not have taken up screening without a 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 diseases 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 more likely to show positive health behaviors, such as getting flu shots, cervical screening (pap tests), breast self-examination and not smoking ([Einav et al., 2020](#); Kowalski, 2023). This is consistent with the empirical evidence on the correlations in various positive health behaviors ([Oster, 2020](#); [Cutler and Lleras-Muney, 2010](#)). [Kim and Lee \(2017\)](#) further show that compliers have lower stomach and breast cancer mortality and all-cause mortality in six years compared to never-takers.

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. On the other hand, [Kim and Lee \(2017\)](#) exploit income cutoff for stomach and breast screening subsidies in Korean national cancer screening program and find that both the cancer mortality and all-cause mortality are higher for compliers compared to always-takers.^{[27](#)}

To characterize compliers, I proceed in two steps.^{[28](#)} First, I restrict the sample to screening participants and compare treatment group with control group. After the sample restriction, the treatment group consists of always-takers and treated compliers, whereas the control group is only composed of always-takers.^{[29](#)} 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.^{[30](#)}

This provides graphic evidence that is intuitive and easy to understand, but it only provides quantitative comparison. This is because the screening participants in the treatment group consist of both compliers and always-takers. Therefore, in the second approach, I estimate the share of always-takers, compliers and never-takers from the first stage regression (Equation (1) and (2)) and use the shares to back out complier characteristics. This allows me to estimate the characteristics of three compliance groups and statistically test the difference in means. It also allows differentiating compliers in

²⁷[Kim and Lee \(2017\)](#) and my study both use health screenings in Korea, but they are different in 2 main aspects. First, [Kim and Lee \(2017\)](#) focus only on national cancer screening program, while my study considers all types of health screenings in Korea including the cancer screenings. Second, we use different sources of exogenous variation. In my setting, I exploit the biennial subsidy rule that, in case of cancer screenings, covers 90 percent of the cancer screening costs. For those with income below certain cutoffs were further subsidized and waived the 10 percent copay. [Kim and Lee \(2017\)](#) uses this income cutoff for the waiver of 10 percent copay.

²⁸It is straightforward in settings with only one-sided noncompliance. In a 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 untreated compliers. Similar comparison can be made in the treatment group when there are only never-takers but no always-takers.

²⁹Monotonicity assumption rules out defiers.

³⁰Similarly, once can restrict the sample to screening nonparticipants and compare treatment with control group. This allows comparing untreated compliers with never-takers.

the treatment group, which I call treated compliers, and compliers in the control group called untreated compliers. I compare always-takers with treated compliers and never-takers with untreated compliers to cancel out any causal effect of screening, thereby leaving only selection effect. I defer detailed estimation steps to Appendix section D. Both methods were used by [Einav et al. \(2020\)](#), [Kowalski \(2023\)](#) and [Kim and Lee \(2017\)](#) to characterize compliers in the health screening context.³¹

5.4.1 Compliers with screening subsidies

Figure 4a and 4b show that compliers with screening subsidies are more likely to find a disease through screening than always-takers. The screening participants were asked if they found any disease through screening. The figures plots the share of screening participants who reported they were diagnosed with or showed manifestations of a disease. The diagnosis rate jumps at age 40 in Figure 4a and the even age group consistently shows higher diagnosis rate through out the age range in Figure 4b. This implies compliers are less healthy than always-takers and are more likely to be diagnosed with a disease through screening. It reflects different underlying health conditions of the two groups prior to screening, not any treatment effect of screening. This selection pattern contrasts with the finding from [Einav et al. \(2020\)](#) where the true positive breast cancer rate drops immediately at age 40 cutoff in response to mammogram age recommendation.

Figure 5a and 5b provides formal comparison of treated compliers with always-takers. Table 7 and Table 8 provides the detailed estimation results. Figure 5a reports relative characteristics of treated compliers compared to always-takers using 2-year binned ages and age 40 cutoff. Consistent with Figure 4a, compliers are more likely to find a disease through screening and it is particularly pronounced in stomach screening with stomach related diagnoses. The differences were not significant for breast and colorectal screenings.³² The opposite selection pattern in case of cervical screening captures selec-

³¹Another way to characterize complier is to use Abadie's Kappa weighting ([Abadie, 2003](#)). However, I choose the above methods used in previous studies to ensure comparability of the results. The Abadie's Kappa weighting method produces similar characteristics of compliers as my method.

³²One reason for the lack of same pattern in breast 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 (18%) and also high enough diagnosis rate (23%). Colorectal screening also

tion in cross spillover effect due to the different subsidy starting age, not the selection in response to subsidies.³³ Figure 5b using even versus odd age group comparison to characterize compliers provides similar selection pattern with better precision. Compliers to biennial subsidies provided at even ages are more likely to find a disease, especially stomach related disease, than always-takers. Appendix section E presents some of the diagnoses from each screening. The most common diagnosis from stomach screening was gastritis and duodenitis which could be an early sign of a stomach cancer.

One reason compliers have worse health conditions and are more likely to find a disease through screening could be their lower socioeconomic background. Low income people who were financially constrained in screening participation could be more sensitive to subsidies and are more likely to participate given subsidies. To the extent that lower income is associated with worse health conditions, compliers from lower socioeconomic background could show worse health conditions (Lindahl, 2005). Table 5a and 5b show this is indeed the case. Compliers have lower individual and household income than always-takers. This could be because they are less likely to work. They are also less educated and are less likely to be college graduates. The overall selection pattern suggests compliers are from lower socioeconomic background. This is consistent with the finding that prohibiting deductibles for mammogram led to increase in take-up and the effect was concentrated among high school dropout with lower income (Bitler and Carpenter, 2016).

In terms of health behaviors, compliers are less likely to smoke, drink and exercise. The selection pattern is more precisely estimated using even versus odd age variation as shown in Figure 5b. Compared to always-takers, compliers are less likely to be current and everyday smoker, less likely to be current and everyday drinker, and less likely to do vigorous and moderate exercise.³⁴ The selection pattern is consistent with negative

shows similar level of diagnosis rate (20%) to stomach screening, but the average take-up is much lower (5%). Breast screening shows large take-up (16%) but lower diagnosis rate (2%).

³³As discussed in section 5.2, cervical screening subsidies begin at age 30, but there is a much larger increase in take-up at age 40 than at age 30. Therefore, a high risk group or those who are in more need of a cervical screening begin to participate from age 30 followed by lower risk group at age 40. This sorting can generate the drop in cervical disease diagnosis at age 40 as shown in Figure 5a.

³⁴Note that if there is any causal effect of screening on health behaviors, it cancels out in my analysis, since both the treated compliers and always-takers participate in screening.

selection on income. People with low income are less likely to afford cigarette, alcohol and exercise.³⁵ It also echoes previous studies on health behaviors that 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](#)).

One important finding different from previous studies is the worse health conditions of compliers compared to always-takers. This is in contrast with the finding from [Einav et al. \(2020\)](#) that compliers are less likely to have true positive breast cancer and have better health conditions than always-takers. One potential explanation is the different nature of treatment. The treatment studied in [Einav et al. \(2020\)](#) is mammogram age recommendation at age 40 provided by the U.S. Preventive Service Task Force (USPSTF). Compliers in this setting are the women who would not have get screened but respond 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. It is also consistent with findings from [Kim and Lee \(2017\)](#) that show compliers to screening fee subsidy have higher cancer mortality and all-cause mortality compared to always-takers.

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 impor-

³⁵My data show that the correlations between all the given health behaviors and household income are positive. However, correlation only measures linear association and may mask nonlinearity or non-monotonicity. Current drinker status and 3 types of exercise exhibit positive relationship throughout the income distribution. On the other hand, current smoker, everyday smoker and everyday drinker variable show inverted U-shape with income. People are more likely to be current and everyday smoker and everyday drinker as income rises, but at some point they start to show negative relationship. The margin captured in this study could be before the peak in the inverted U-shape relationship such that more income is correlated with higher probability of smoking and everyday drinking.

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

tant 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 of 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.

The comparison between compliers and never-takers shown in Figure 6a and 6b confirms positive selection in health behaviors, consistent with previous studies ([Kowalski, 2023](#); [Oster, 2020](#); [Cutler and Lleras-Muney, 2010](#)). Once again, comparison between even and age group presents more precise estimates than using age cutoff. 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 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.

5.4.2 Compliers in cross spillover

This section examines complier characteristics in cross spillover. Section 5.2 discusses the cross spillover from biennial screenings to annual and no-subsidy screenings. Despite annual subsidies for liver and colorectal screenings and no subsidy for prostate and lung screenings, they display biennial take-up pattern from age 40. The dataset shows that annual and no-subsidy participants are a subset of biennial screening participants. More than 96% of annual and no-subsidy screening participants also receive biennial screenings. This means that there is hardly anyone who gets colorectal screening but do not get general screening. This absence of always-takers makes selection analysis much easier. One only needs to examine who participate in annual or no-subsidy screening among biennial

screening participants. I run the following regression to estimate the characteristics of compliers in spillover.

$$y_{it} = \delta_0 + \delta_1 \cdot screen_{it} + \varepsilon_{it} \quad (4)$$

The sample is restricted to biennial screening participants.³⁷ The outcome variable, y_{it} , is diagnosis, socioeconomic status and health behaviors. The explanatory variable, $screen_{it}$, is an indicator variable for participating in any of the two annual or two no-subsidy screenings. The standard error is clustered at the individual level.

Table 9 presents the characteristics of annual and no-subsidy screening participants among biennial screening participants. The spillover compliers are less likely to have somtach, breast and cervical disease diagnoses, indicating better health conditions. This is explained by their better socioeconomic status. Compliers have more individual and household income, more education and more likely to working. Consistent with higher income, compliers are more likely to smoke, drink and exercise.

The overall selection indicates opposite pattern from the selection in response to subsidies. Those who further participate in annual and no-subsidy screenings after participating in biennial screenings are the ones from higher socioeconomics status. Hence, they show better health conditions and report fewer diseases. They are more likely to afford cigarettes and alcohol and more likely to exercise. The pattern is stronger for prostate and lung screenings that are not subsidized. While it is understandable that participants in no-subsidy screenings are the ones with higher income and more education, it is not immediately clear why there is a similar, albeit less strong, selection pattern for annual screenings that are subsidized every year. One potential reason could be that participants in liver and colorectal screenings are not actually using the subsidies. Since liver screening subsidies are for high risk group only, the non-high risk group pays the full cost of testing. Similarly, even though fecal occult blood tests are subsidized for the colorectal screening, data show many participants go straight to use colonoscopy and pay the full cost of around \$150. This could result in compliers with higher income and education for

³⁷When the outcome variable is stomach disease diagnosis, the sample is restricted to stomach screening participants, and similarly for breast and cervical screenings. For the other outcome variables, the sample is restricted to participants in any of the four biennial screenings.

both the annual and no-subsidy screenings.

5.5 Effect of health screening

This section estimates the causal effect of health screening on diagnoses and health care utilizations using the exogenous variation in screening take-up between even and odd ages.³⁸ 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.

$$y_{it} = \eta_0 + \eta_1 \cdot screen_{it} + \mathbf{f}(\mathbf{age}_{it}) + \varepsilon_{it} \quad (5)$$

The outcome variable, y_{it} , is the number of outpatient, inpatient or emergency hospital visits of individual i in year t . The number of hospital visits for screening is not included. For each care, I consider the total number of hospital visits and also the visits for specific organs that cancer screenings are designed to examine.³⁹ The main independent variable, $screen_{it}$, is the take-up of any biennial screening in case of the aggregate number of visits and the take-up of corresponding screening for each organ.⁴⁰ It is instrumented by the even age variable where the first stage is given in Table 4 and 5. I also include the flexible function of age as control variables. The standard errors, ε_{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 11 hypotheses in each domain with 10,000 bootstrap replications ([Jones et al., 2019](#)).

An econometric challenge in estimating the effect of screening on disease diagnosis is that diagnoses are defined conditional on getting screened. To overcome this problem, I infer diagnoses not from screening results data, but from outpatient hospital visit records that are collected regardless of screening take-up.⁴¹ Specifically, I use first hospital visits

³⁸Due to the coarseness of the running variable, I do not use the regression discontinuity design at age 40.

³⁹For cervical and prostate screenings, I consider more broadly the hospital visits for female and male genitals.

⁴⁰High blood pressure, hyperlipidemia and diabetes can be diagnosed from the general screening. So, the take-up of general screening is used as an independent variable in these cases.

⁴¹Another way to overcome this problem is to use cancer registry data that records all the cancer

for a new illness as a proxy for new diagnoses. For each outpatient visit to a hospital, the survey participants recorded whether this is (*i*) a first visit for a new illness, (*ii*) a second visit, or (*iii*) a recurring regular visit. They also recorded the diagnosis for each visit using Korean Classification of Diseases diagnosis code (Korean version of the ICD-10). I combine the diagnosis information with the first visit answer to count the number of first hospital visits for a certain diagnosis category and use it as a proxy for new diagnosis.

Panel A in Table 10 presents the effect of health screening on outpatient hospital visits. Local average treatment effects reported in column 3 shows the aggregate number of outpatient hospital visits increase by 0.478 (2.3 percent), but it was imprecisely estimated. The visits for other symptoms also seem to increase, but the estimates were mostly noisy. Stomach related visit alone showed positive increase that is also significant at the conventional level. It echoes the finding from the selection analysis that compliers are more likely to find stomach related disease through screening.

Panel B presents the effect on the number of first hospital visits, a proxy for new diagnosis. Health screening also leads to an increase in first outpatient visits and unlike in aggregate counts, the effects are statistically significant at the conventional level. The number of total first outpatient visits increase by 0.363 (9.2 percent) and it is precisely estimated. Across different types of diagnosis categories, the estimates are mostly positive and they are significant for diagnoses related to hyperlipidemia, stomach, breast, female genital and colon and rectum. This suggests people find new health conditions through health screenings and subsequently seek medical care.

One of the strongest risk factor for stomach cancer is infection with Helicobacter pylori (*H. pylori*), a type of bacteria that attacks stomach lining and cause inflammation and ulcers. Chronic infection and long lasting stomach inflammation due to *H. pylori* is known to lead to stomach cancer and potentially colorectal cancer as well (Polk and Peek Jr, 2010; Uemura et al., 2001; Butt and Epplein, 2019; Zuo et al., 2020; Boustany et al., 2023). Many studies show treating *H. pylori* with antibiotics leads to decrease

diagnoses in a country. Both Einav et al. (2020) and Kowalski (2023) use cancer registry data to detect cancer diagnoses regardless of screening take-up. This study overcome the challenge by using visit level outpatient data that are similar to claims data in that diagnoses for each hospital visit are recorded regardless of screening status.

in the risk of stomach cancer (Li et al., 2019; Ford et al., 2020; Chiang et al., 2021). Gastroscopy used for stomach screening can detect *H. pylori*, after which treatment can follow in case of positive result. Our stomach diagnosis category includes the visits for *H. pylori*. Separately examining the effect on *H. pylori* shows that stomach screening does lead to more people visiting hospitals to treat *H. pylori*. Screening leads to 0.014 more hospital visits (472 percent) and 0.004 more first hospital visits (526 percent) for *H. pylori* and both estimates are significant at 1% level. (Results not reported in the tables)

Screening leads to more inpatient care and less emergency care usage, but the results are estimated with noise. Panel A in Table 11 presents the effect of health screening on inpatient care usage. The aggregate inpatient care visit increases by 0.027 (11.7 percent) in response to health screening, but it is imprecisely estimated. Inpatient care utilizations across different diagnosis categories overall display increases in response to screening but most of the estimates are not statistically significant. Panel B presents the effect on emergency care usage. The aggregate ER visits decreases by 0.012 (9.3 percent), but it is also imprecisely estimated. Most of ER visits in different categories display decrease in response to screening, but they are also statistically insignificant.

The causal effect of health screening can also be studied using the variation from spousal spillover. Table A9 in Appendix section F presents the effect of one's screening and the spouse's screening on one's own first hospital visits. Consistent with evidence shown in Panel B Table 10, one's own screening take-up leads to a significant increase in first hospital visits. Spouse's screening take-up instrumented by spouse's subsidy eligibility also seems to lead to an increase in first hospital visits, but they are smaller in magnitude and also estimated with noise. This suggests despite statistical insignificance, spouse's screening leads to an increase one's own screening take-up and also an increase in new diagnoses.

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 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 and 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 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 suggests that screening fulfills its primary goal of inducing new diagnoses and people subsequently seeking medical 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.
- 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.
- Baik, S. M. and Lee, R.-A. (2023). National cancer screening program for colorectal cancer in korea. *Annals of Surgical Treatment and Research*, 105(6):333.
- 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.
- Boustany, A., Onwuzo, S., Almomani, A., and Asaad, I. (2023). Epidemiology and risk of colorectal cancer in patients with a history of helicobacter pylori infection: a population-based study. *Annals of Gastroenterology*, 36(2):203.
- Butt, J. and Epplein, M. (2019). Helicobacter pylori and colorectal cancer—a bacterium going abroad? *PLoS pathogens*, 15(8):e1007861.
- 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.
- Chiang, T.-H., Chang, W.-J., Chen, S. L.-S., Yen, A. M.-F., Fann, J. C.-Y., Chiu, S. Y.-H., Chen, Y.-R., Chuang, S.-L., Shieh, C.-F., Liu, C.-Y., et al. (2021). Mass eradication of helicobacter pylori to reduce gastric cancer incidence and mortality: a long-term cohort study on matsu islands. *Gut*, 70(2):243–250.
- 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.
- 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.
- Ford, A. C., Yuan, Y., Forman, D., Hunt, R., and Moayyedi, P. (2020). Helicobacter pylori eradication for the prevention of gastric neoplasia. *Cochrane database of systematic reviews*, (7).
- Gafar, A., Suza, D. E., Efendi, F., Has, E. M. M., Pramono, A. P., and Susanti, I. A. (2020). Determinants of contraceptive use among married women in indonesia. *F1000Research*, 9.
- 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.
- 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.
- Howlader, N., Noone, A., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D., Chen, H., Feuer, E., and Cronin, K. e. (2021). *SEER Cancer Statistics Review, 1975-2018, National Cancer Institute*. National Cancer Institute,

Bethesda, MD. Based on November 2020 SEER data submission, posted to the SEER web site, April 2021.

Hungerford, D., MacPherson, P., Farmer, S., Ghebrehewet, S., Seddon, D., Vivancos, R., and Keenan, A. (2016). Effect of socioeconomic deprivation on uptake of measles, mumps and rubella vaccination in liverpool, uk over 16 years: a longitudinal ecological study. *Epidemiology & Infection*, 144(6):1201–1211.

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.

Kang, H.-T. (2022). Current status of the national health screening programs in south korea. *Korean Journal of Family Medicine*, 43(3):168.

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.
- Li, W.-Q., Zhang, J.-Y., Ma, J.-L., Li, Z.-X., Zhang, L., Zhang, Y., Guo, Y., Zhou, T., Li, J.-Y., Shen, L., et al. (2019). Effects of helicobacter pylori treatment and vitamin and garlic supplementation on gastric cancer incidence and mortality: follow-up of a randomized intervention trial. *bmj*, 366.
- 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.

Polk, D. B. and Peek Jr, R. M. (2010). Helicobacter pylori: gastric cancer and beyond. *Nature reviews cancer*, 10(6):403–414.

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.

Strandberg, T. E., Salomaa, V. V., Vanhanen, H. T., Naukkarinen, V. A., Sarna, S., and Miettinen, T. A. (1995). Mortality in participants and non-participants of a multifactorial prevention study of cardiovascular diseases: a 28 year follow up of the helsinki businessmen study. *Heart*, 74(4):449–454.

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.

Thomas, K., Darling, J., and Cassil, A. (2021). Covid-19 vaccine hesitancy: Education divide widens. *USC Dornsife Center for Economic and Social Research Data Brief No, 2.*

- Uemura, N., Okamoto, S., Yamamoto, S., Matsumura, N., Yamaguchi, S., Yamakido, M., Taniyama, K., Sasaki, N., and Schlemper, R. J. (2001). Helicobacter pylori infection and the development of gastric cancer. *New England journal of medicine*, 345(11):784–789.
- 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.
- Wetter, D. W., Cofta-Gunn, L., Irvin, J. E., Fouladi, R. T., Wright, K., Daza, P., Mazas, C., Cinciripini, P. M., and Gritz, E. R. (2005). What accounts for the association of education and smoking cessation? *Preventive medicine*, 40(4):452–460.
- Zuo, Y., Jing, Z., Bie, M., Xu, C., Hao, X., and Wang, B. (2020). Association between helicobacter pylori infection and the risk of colorectal cancer: A systematic review and meta-analysis. *Medicine*, 99(37).

7 Tables

Table 1: Screening schedule

	Biennial subsidy				Annual subsidy		No subsidy	
	General	Stomach	Breast	Cervical	Liver	Colorectal	Lung	Prostate
Frequency	2 years	2 years	2 years	2 years	0.5 year	1 year		
Subsidy starting age	40	40	40	30	40	50		
Subsidy amount	100%	90%	90%	100%	90%	90%	0%	0%
Copay (\$)	0	7	3.5	0	10	5	110	20
Target		Female	Female	High risk group			Male	
Subsidized medical tests	Gastroscopy, biopsy	Mammogram	Pap smear	Ultrasound, MSAFP	Fecal immuno-chemical test, colonoscopy, biopsy			

Notes: This table summarizes the subsidy schedules 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 during the study period. Liver screening is subsidized up to twice a year. The subsidy starting age for cervical screening was lowered to 20 in 2016. The colorectal screening used to be biennially subsidized at even ages from age 50 until 2011. It became an annually subsidized screening from year 2012. The copay amounts in US dollars are rough estimates and can vary with hospitals and screening years.

Table 2: Balance table

	(1)	(2)	(3)
	Treatment group	Control group	Conditional difference
Age	58.697 (12.532)	59.240 (12.353)	- -
Female	0.530 (0.499)	0.532 (0.499)	-0.002* (0.001)
Currently married	0.799 (0.401)	0.798 (0.402)	-0.001 (0.001)
Years of education	10.320 (4.510)	10.227 (4.538)	-0.003 (0.008)
Working status	0.610 (0.488)	0.608 (0.488)	-0.003* (0.001)
Individual income	1446.3 (2081.6)	1425.7 (2068.1)	2.8 (5.2)
Household income	4104.4 (3708.6)	4086.7 (3737.9)	3.2 (14.3)
Own a house	0.734 (0.442)	0.737 (0.441)	-0.000 (0.001)
Number of household members	3.067 (1.317)	3.051 (1.317)	-0.004 (0.003)
N	54274	52909	
Share	(0.51)	(0.49)	
F(8, 15939)			1.65 (0.10)

Notes: This table reports the 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 differences between treatment and control group 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 subsidies on take-up using regression discontinuity

	(1)	(2)	(3)	(4)	(5)
	Any	General	Stomach	Breast	Cervical
Age ≥ 40	0.097*** (0.008)	0.086*** (0.007)	0.105*** (0.006)	0.112*** (0.009)	0.074*** (0.010)
Constant	0.121*** (0.005)	0.095*** (0.005)	0.061*** (0.004)	0.064*** (0.005)	0.093*** (0.006)
N	34713	34713	34713	17725	17725
Adj R^2	0.017	0.020	0.032	0.037	0.013
Sample age range	[34, 45]	[34, 45]	[34, 45]	[34, 45]	[34, 45]
Subsidy starting age		40	40	40	30

Notes: This table reports the effect of biennial subsidies on 4 types of screening take-up by using regression discontinuity design at age 40 after binning ages by 2 years. The econometric specification is given in Equation (1). The outcome variable is the average take-up in the 2 year age bins. Column 1 uses the take-up of any type of screening reported in column 2 to 5 as an outcome variable. Coefficients for $Age \geq 40$ measures the jump in take-up at age 40 and the constant measures the take-up right before the jump at age 40. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 4: Effect of biennial subsidies on take-up comparing even with odd age group

	(1)	(2)	(3)	(4)	(5)
	Any	General	Stomach	Breast	Cervical
Age even	0.204*** (0.003)	0.187*** (0.003)	0.190*** (0.003)	0.191*** (0.004)	0.164*** (0.003)
N	107183	107183	107183	56923	56923
Adj R^2	0.068	0.061	0.069	0.080	0.074
F-statistic	5012	4804	4830	2904	2520
Sample age range	[40, 89]	[40, 89]	[40, 89]	[40, 89]	[40, 89]
Subsidy starting age		40	40	40	30
Age controls	Y	Y	Y	Y	Y
Control group mean	0.122	0.102	0.083	0.067	0.056

Notes: This table reports the effect of biennial subsidies on 4 types of screening take-up by comparing even age group with odd age group from age 40. The econometric specification is given in Equation (2). The outcome variable is the average take-up at each age. Column 1 uses the take-up of any type of screening reported in column 2 to 5 as an outcome variable. The estimates measure the effect of screening subsidies 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 5: Cross spillover in take-up across different screenings

	(1)	(2)	(3)	(4)	(5)
	Annual subsidy		No subsidy		Biennial subsidy
	Liver	Colorectal	Prostate	Lung	Cervical
Age even	0.027*** (0.001)	0.017*** (0.002)	0.007*** (0.001)	0.006*** (0.001)	0.037*** (0.004)
Age even \times age ≥ 50		0.023*** (0.002)			
Age even \times age ≥ 40					0.128*** (0.005)
N	107183	107183	50260	107183	69236
Sample age range	[40, 89]	[40, 89]	[40, 89]	[40, 89]	[30, 89]
Subsidy starting age	40	50	40	40	30
Age controls	Y	Y	Y	Y	Y
Control group mean	0.028	0.027	0.009	0.009	0.056

Notes: This table reports the effect of biennial subsidies on the take-up of unsubsidized prostate and lung screenings and annually subsidized liver and colorectal screenings. It also compares the effect of biennial subsidies on colorectal screening take-up (subsidy starting age 50) before and after age 50 and on cervical screening take-up (subsidy starting age 30) before and after age 40. The estimates are measured conditional on the linear splines of age with 5 years interval. Control group mean reports the average take-up for the odd age group in the given sample age range. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 6: Spillover in take-up between spouses

	(1)	(2)	(3)	(4)
	Outcome var: Own screening take-up			
Age even	0.214*** (0.006)	0.213*** (0.006)	0.213*** (0.004)	0.213*** (0.004)
Spouse age even	0.016*** (0.005)	0.015*** (0.004)		
Age even \times Spouse age even	0.001 (0.009)	0.003 (0.009)		
Spouse screening			0.078*** (0.017)	0.079*** (0.017)
N	79962	79782	79962	79782
Odd/Odd group mean	0.128	0.128	0.128	0.128
Demographic controls		Y		Y
Estimator	OLS	OLS	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 both of whose age is 40 or above. Odd/Odd group mean refers to the average take-up if one's own and the spouse's ages are both odd. 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 3 to 4, spouse screening 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 7: Compliers with subsidies using age 40 cutoff

	(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	0.219 (0.019)	0.442 (0.045)	-	-	2.019*** (0.348)	-
Stomach disease diagnosis	0.125 (0.024)	0.300 (0.028)	-	-	2.397** (0.668)	-
Breast disease diagnosis	0.027 (0.014)	0.028 (0.016)	-	-	1.047 (7.810)	-
Cervical disease diagnosis	0.159 (0.028)	0.022 (0.053)	-	-	0.137** (0.391)	-
Colorectal disease diagnosis	0.120 (0.041)	0.144 (0.107)	-	-	1.204 (12.827)	-
Panel B. SES						
Individual income	2515 (108)	1880 (201)	1760 (366)	2181 (46)	0.748** (0.098)	0.807 (0.175)
Household income	5901 (136)	4751 (254)	4567 (704)	4748 (67)	0.805*** (0.054)	0.962 (0.158)
Years of education	14.175 (0.109)	13.797 (0.193)	13.975 (0.370)	13.660 (0.051)	0.973 (0.018)	1.023 (0.029)
College graduate	0.391 (0.025)	0.415 (0.041)	0.398 (0.078)	0.324 (0.010)	1.063 (0.151)	1.230 (0.257)
Working status	0.780 (0.021)	0.654 (0.039)	0.577 (0.085)	0.758 (0.009)	0.838** (0.064)	0.761** (0.116)
Panel C. Health behaviors						
Current smoker	0.195 (0.021)	0.202 (0.038)	0.388 (0.082)	0.288 (0.011)	1.033 (0.278)	1.349 (0.309)
Everyday smoker	0.182 (0.021)	0.194 (0.037)	0.375 (0.083)	0.272 (0.010)	1.064 (0.298)	1.379 (0.325)
Current drinker	0.868 (0.019)	0.775 (0.036)	0.831 (0.076)	0.819 (0.009)	0.893** (0.053)	1.015 (0.097)
Everyday drinker	0.038 (0.009)	0.040 (0.019)	0.071 (0.046)	0.047 (0.005)	1.036 (0.775)	1.499 (1.085)
Vigorous exercise	0.294 (0.023)	0.237 (0.046)	0.299 (0.089)	0.266 (0.009)	0.808 (0.207)	1.123 (0.355)
Moderate exercise	0.402 (0.026)	0.457 (0.050)	0.349 (0.104)	0.396 (0.010)	1.139 (0.185)	0.881 (0.277)
Walking	0.792 (0.022)	0.799 (0.042)	0.888 (0.092)	0.756 (0.009)	1.008 (0.072)	1.175 (0.131)

Notes: This table reports the average values of screening diagnoses, socioeconomic status, and health behaviors among always-takers, never-takers, treated compliers and untreated compliers. The compliance groups are defined using the age 40 cutoff. Ages are binned by two years to average out the substitution effect between the even and odd years after the age 40. 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 (7). 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. 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 8: Compliers with subsidies comparing even and odd groups after age 40

	(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	0.257 (0.006)	0.396 (0.007)	-	-	1.539*** (0.052)	-
Stomach disease diagnosis	0.172 (0.006)	0.272 (0.006)	-	-	1.586*** (0.073)	-
Breast disease diagnosis	0.023 (0.004)	0.026 (0.003)	-	-	1.158 (0.249)	-
Cervical disease diagnosis	0.069 (0.007)	0.066 (0.006)	-	-	0.951 (0.149)	-
Colorectal disease diagnosis	0.190 (0.011)	0.202 (0.011)	-	-	1.061 (0.105)	-
Panel B. SES						
Individual income	2610 (56)	1985 (50)	1946 (53)	2208 (40)	0.760*** (0.019)	0.882*** (0.018)
Household income	5571 (74)	5022 (68)	4990 (95)	4780 (54)	0.901*** (0.014)	1.044** (0.019)
Years of education	14.142 (0.066)	13.880 (0.068)	13.889 (0.080)	13.565 (0.046)	0.981*** (0.006)	1.024*** (0.005)
College graduate	0.386 (0.011)	0.333 (0.010)	0.336 (0.011)	0.333 (0.009)	0.864*** (0.021)	1.008 (0.021)
Working status	0.796 (0.010)	0.712 (0.010)	0.723 (0.012)	0.754 (0.008)	0.894*** (0.013)	0.958*** (0.014)
Panel C. Health behaviors						
Current smoker	0.233 (0.010)	0.192 (0.010)	0.172 (0.011)	0.298 (0.009)	0.822*** (0.033)	0.576*** (0.028)
Everyday smoker	0.218 (0.010)	0.182 (0.009)	0.158 (0.011)	0.284 (0.009)	0.834*** (0.034)	0.558*** (0.028)
Current drinker	0.848 (0.009)	0.813 (0.009)	0.798 (0.012)	0.816 (0.007)	0.958*** (0.012)	0.978* (0.013)
Everyday drinker	0.044 (0.005)	0.027 (0.005)	0.035 (0.006)	0.051 (0.004)	0.625*** (0.114)	0.683*** (0.118)
Vigorous exercise	0.309 (0.009)	0.270 (0.010)	0.284 (0.012)	0.264 (0.007)	0.871*** (0.030)	1.077** (0.039)
Moderate exercise	0.456 (0.010)	0.426 (0.010)	0.427 (0.015)	0.389 (0.008)	0.935*** (0.023)	1.098*** (0.034)
Walking	0.791 (0.008)	0.786 (0.008)	0.778 (0.012)	0.754 (0.007)	0.994 (0.011)	1.032** (0.016)

Notes: This table reports the average values of screening diagnoses, socioeconomic status, and health behaviors among always-takers, never-takers, treated compliers and untreated compliers. 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 (7). 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 individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 9: Compliers in cross spillover

	(1)	(2)	(3)
	Annual	No-subsidy	Sample mean
Panel A. Diagnoses			
Stomach disease diagnosis	-0.026*** (0.006)	-0.086*** (0.010)	0.228
Breast disease diagnosis	-0.006* (0.003)	-0.019*** (0.004)	0.022
Cervical disease diagnosis	-0.018*** (0.006)	-0.022 (0.016)	0.062
Panel B. SES			
Individual income	874*** (49)	1499*** (110)	1592
Household income	1012*** (66)	1393*** (145)	4564
Years of education	0.975*** (0.073)	1.342*** (0.129)	10.769
College graduate	0.074*** (0.007)	0.131*** (0.014)	0.196
Working status	0.063*** (0.008)	0.141*** (0.012)	0.635
Panel C. Health behaviors			
Current smoker	0.044*** (0.006)	0.116*** (0.013)	0.146
Everyday smoker	0.041*** (0.006)	0.100*** (0.013)	0.138
Current drinker	0.066*** (0.008)	0.149*** (0.012)	0.655
Everyday drinker	0.018*** (0.004)	0.035*** (0.009)	0.060
Vigorous exercise	0.050*** (0.007)	0.104*** (0.014)	0.235
Moderate exercise	0.050*** (0.008)	0.107*** (0.014)	0.409
Walking	0.012** (0.006)	0.031*** (0.010)	0.812

Notes: This table reports the characteristics of annual- and no-subsidy screening participants among biennial screening participants. The sample is stomach or breast or cervical screening participants when outcome variable is stomach or breast or cervical disease diagnosis, respectively. For the other outcomes, the sample is participants in any of the 4 biennial screenings. Column 1 reports the coefficient of annual screening take-up and column 2 reports the coefficient of no-subsidy screening take-up. Column 3 reports the sample mean of the outcome variables. All the coefficients are from separate regressions. Standard errors are clustered at the individual level. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 10: Effect of health screening on outpatient care utilizations

	(1)	(2)	(3)	(4)	(5)
	Control group mean	ITT	LATE	Adjusted p-values	N
Panel A. Outpatient visits					
Outpatient visit	20.8088	0.0977 (0.0757)	0.4784 (0.3709)	0.866	107183
High blood pressure	2.7100	0.0001 (0.0115)	0.0007 (0.0618)	0.993	107183
Hyperlipidemia	0.9847	0.0073 (0.0073)	0.0390 (0.0391)	0.888	107183
Diabetes	1.1378	0.0115 (0.0097)	0.0616 (0.0521)	0.866	107183
Stomach	0.9716	0.0685*** (0.0117)	0.3613*** (0.0616)	0.000	107183
Breast	0.1141	0.0013 (0.0116)	0.0066 (0.0608)	0.993	56923
Female genital	0.3440	0.0053 (0.0094)	0.0321 (0.0571)	0.932	56923
Liver	0.1114	0.0030 (0.0043)	0.1118 (0.1604)	0.932	107183
Colorectal	0.3069	0.0086 (0.0064)	0.2603 (0.1955)	0.866	107183
Male genital	1.2352	-0.0343 (0.0259)	-4.6905 (3.6077)	0.866	50260
Lung	0.1435	0.0066 (0.0064)	1.0699 (1.0402)	0.888	107183
Panel B. First outpatient visits					
First outpatient visit	3.9335	0.0742*** (0.0153)	0.3632*** (0.0749)	0.000	107183
High blood pressure	0.0509	0.0015 (0.0015)	0.0082 (0.0080)	0.767	107183
Hyperlipidemia	0.0239	0.0034*** (0.0010)	0.0184*** (0.0054)	0.005	107183
Diabetes	0.0255	0.0009 (0.0011)	0.0048 (0.0057)	0.771	107183
Stomach	0.1863	0.0246*** (0.0031)	0.1300*** (0.0161)	0.000	107183
Breast	0.0085	0.0023** (0.0011)	0.0121** (0.0055)	0.164	56923
Female genital	0.0891	0.0062** (0.0027)	0.0380** (0.0166)	0.150	56923
Liver	0.0097	0.0009 (0.0007)	0.0320 (0.0272)	0.737	107183
Colorectal	0.0786	0.0035* (0.0019)	0.1054* (0.0590)	0.351	107183
Male genital	0.0681	-0.0024 (0.0027)	-0.3222 (0.3749)	0.771	50260
Lung	0.0197	0.0003 (0.0010)	0.0551 (0.1689)	0.771	107183

Notes: This table reports the effect of health screening on outpatient care utilizations. Panel A reports the effect on the total number of outpatient hospital visits and the ones by diagnosis categories. Panel B reports the effect on the first outpatient hospital visits for a new illness. First visit is inferred from the visit-level outpatient dataset question asking whether the hospital visit was a first visit for a new illness, second visit or a recurring visit for an illness one is already aware of. The diagnosis category is inferred from Korean Classification of Diseases diagnosis code (Korean version of the ICD-10) recorded for each visit to a hospital. Column 2 ITT estimates report the effect of biennial subsidy on outcome variables by comparing even and odd age groups controlling for ages. Column 3 LATE estimates report the effect of screening take-up on outcome variables using biennial subsidy eligibility at even ages as an instrument. The sample consists of those with age 40 to 89. Westfall-Young adjusted p-values from second stage regressions are reported in column 4. They account for all 11 hypotheses in each domain 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.

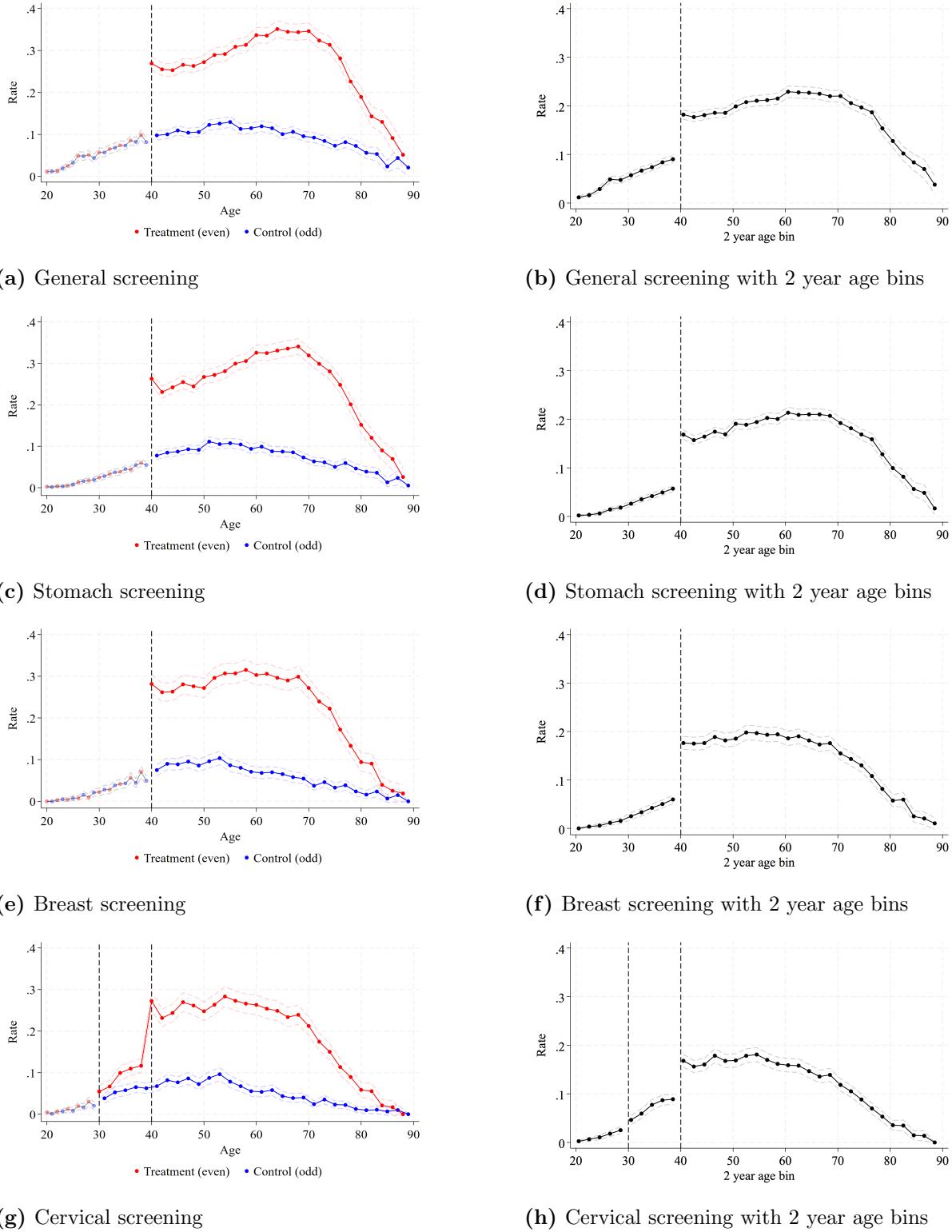
Table 11: Effect of health screening on inpatient and emergency care utilizations

	(1)	(2)	(3)	(4)	(5)
	Control group mean	ITT	LATE	Adjusted p-values	N
Panel A. Inpatient visits					
Inpatient visit	0.23291	0.00555 (0.00391)	0.02719 (0.01915)	0.804	107183
High blood pressure	0.00571	0.00021 (0.00056)	0.00115 (0.00299)	0.997	107183
Hyperlipidemia	0.00043	0.00018 (0.00018)	0.00095 (0.00097)	0.939	107183
Diabetes	0.00692	-0.00025 (0.00067)	-0.00133 (0.00357)	0.997	107183
Stomach	0.01181	0.00133 (0.00101)	0.00703 (0.00534)	0.843	107183
Breast	0.00721	0.00041 (0.00146)	0.00216 (0.00762)	0.997	56923
Female genital	0.00412	-0.00016 (0.00083)	-0.00094 (0.00507)	0.997	56923
Liver	0.00524	0.00082 (0.00069)	0.03084 (0.02593)	0.882	107183
Colorectal	0.01633	-0.00026 (0.00135)	-0.00799 (0.04094)	0.997	107183
Male genital	0.01249	0.00059 (0.00127)	0.08096 (0.17323)	0.997	50260
Lung	0.01183	0.00207* (0.00116)	0.33331* (0.19063)	0.588	107183
Panel B. Emergency visits					
ER visit	0.12520	-0.00238 (0.00259)	-0.01165 (0.01268)	0.956	107183
High blood pressure	0.00113	-0.00004 (0.00025)	-0.00021 (0.00136)	0.993	107183
Hyperlipidemia	0.00008	-0.00001 (0.00007)	-0.00004 (0.00035)	-	107183
Diabetes	0.00261	-0.00032 (0.00045)	-0.00174 (0.00239)	0.956	107183
Stomach	0.00758	-0.00015 (0.00061)	-0.00077 (0.00324)	0.993	107183
Breast	0.00043	-0.00020 (0.00024)	-0.00106 (0.00126)	0.956	56923
Female genital	0.00060	-0.00031 (0.00023)	-0.00190 (0.00139)	0.858	56923
Liver	0.00157	0.00054 (0.00046)	0.02024 (0.01742)	0.928	107183
Colorectal	0.00720	0.00072 (0.00062)	0.02189 (0.01886)	0.928	107183
Male genital	0.00586	-0.00055 (0.00072)	-0.07544 (0.09920)	0.956	50260
Lung	0.00414	0.00009 (0.00049)	0.01407 (0.07856)	0.993	107183

Notes: This table reports the effect of health screening on inpatient and emergency care utilizations. Panel A reports the effect on the total number of inpatient hospital visits and the ones by diagnosis categories. Panel B reports the effect on ER visits. The diagnosis category is inferred from Korean Classification of Diseases diagnosis code (Korean version of the ICD-10) recorded for each visit to a hospital. Column 2 ITT estimates report the effect of biennial subsidy on outcome variables by comparing even and odd age groups controlling for ages. Column 3 LATE estimates report the effect of screening take-up on outcome variables using biennial subsidy eligibility at even ages as an instrument. The sample consists of those with age 40 to 89. Westfall-Young adjusted p-values from second stage regressions are reported in column 4. They account for all 11 hypotheses in inpatient domain with 10000 bootstrap replications. They account for 10 hypotheses in emergency domain due to small number of (less than 10) positive values for ER for hyperlipidemia outcome. Standard errors are clustered at the individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

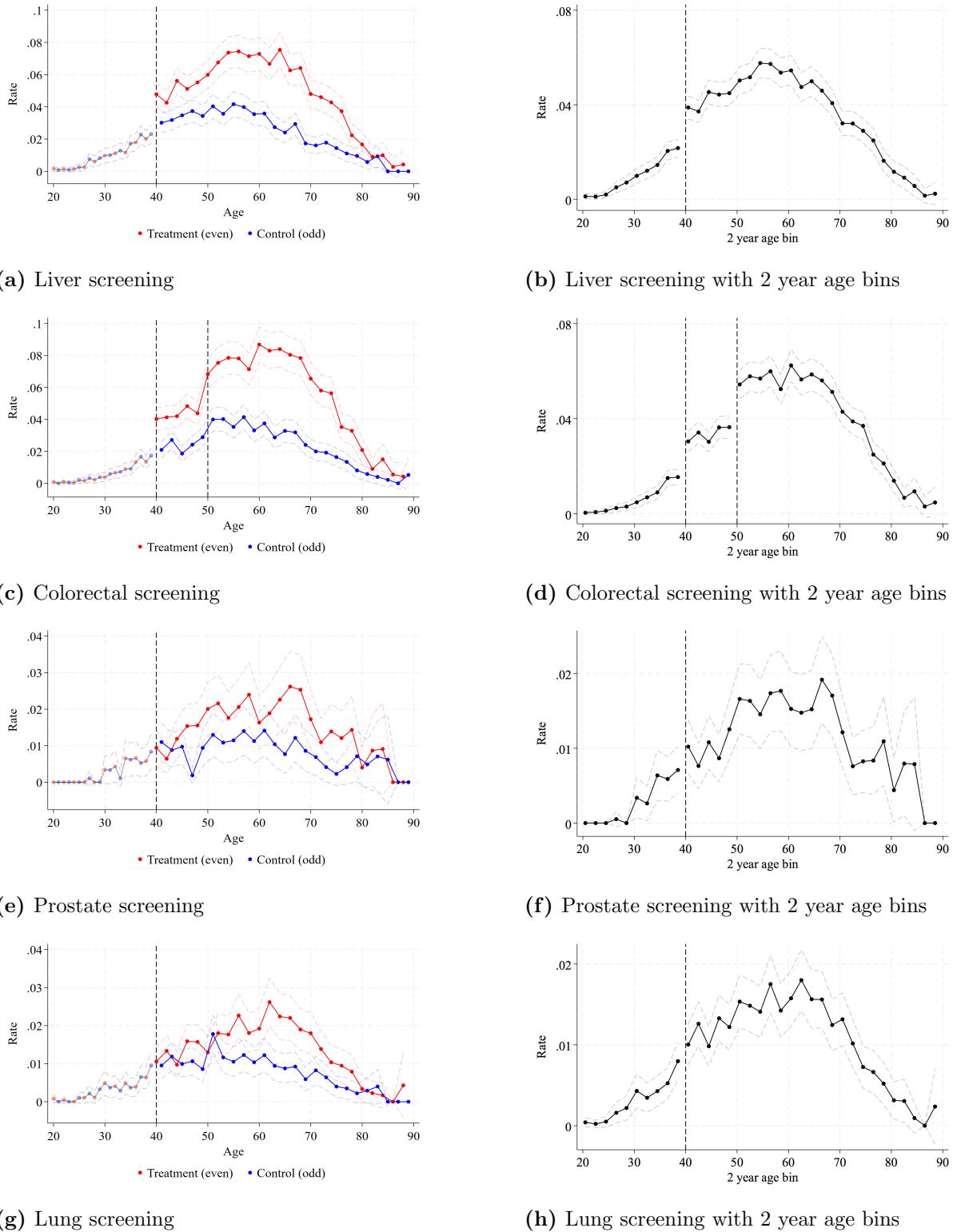
8 Figures

Figure 1: Screening rates for biennial screenings



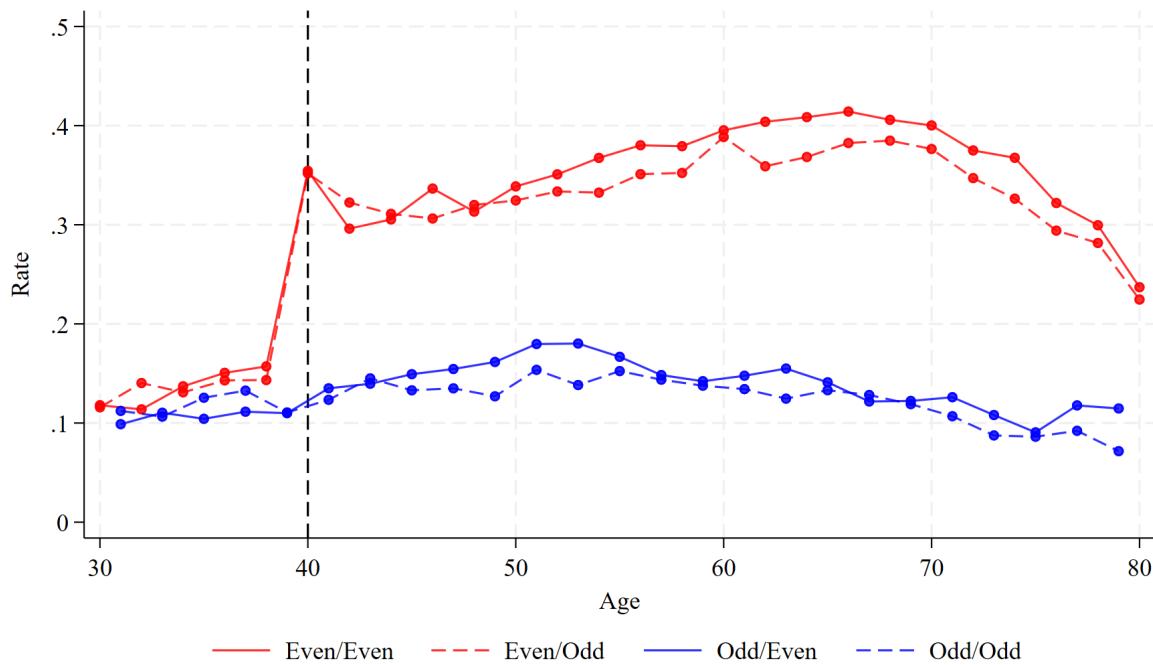
Notes: Figures show the take-up rate for 4 types of biennially subsidized screenings. Figures on the left side plot the average take-up for each age, while the ones on the right side plot the average take-up for each 2-year age bin. In the left figures, even ages are colored in red and odd ages are colored in blue. Dashed vertical lines show 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 black dashed line.

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



Notes: Figures show the average screening rates by age for liver and colorectal screening that are subsidized annually and prostate and lung screenings that are not subsidized. Even ages are colored in red and odd ages are colored in blue. Dashed 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.

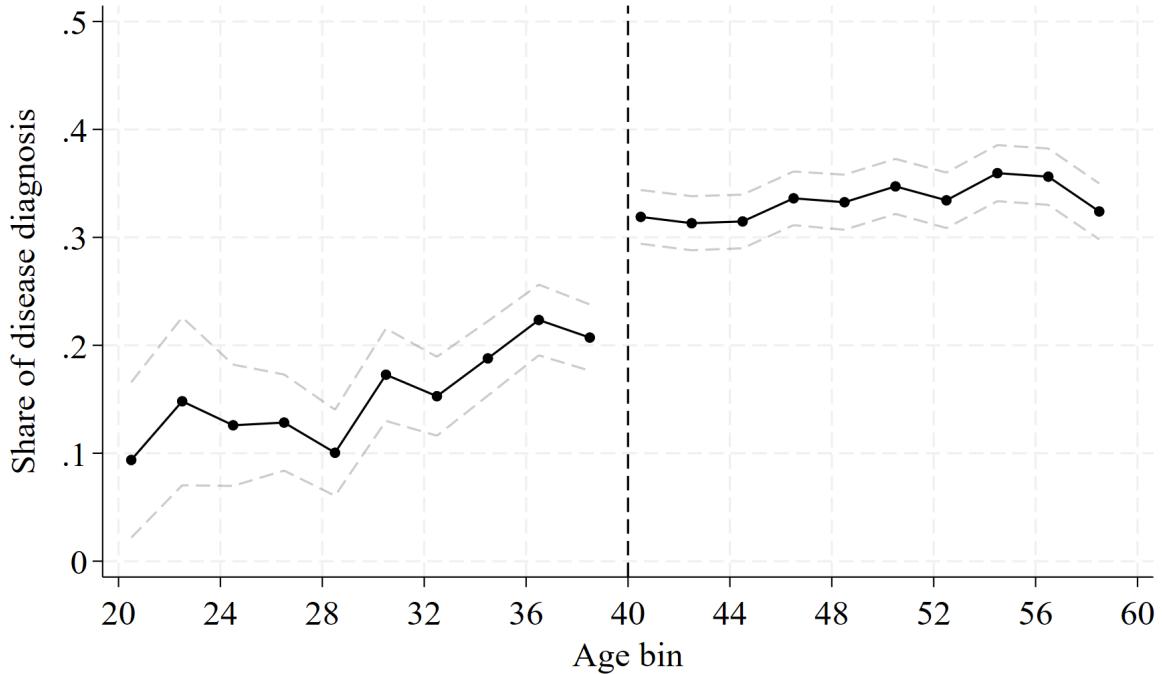
Figure 3: Spillover in screening take-up between spouses



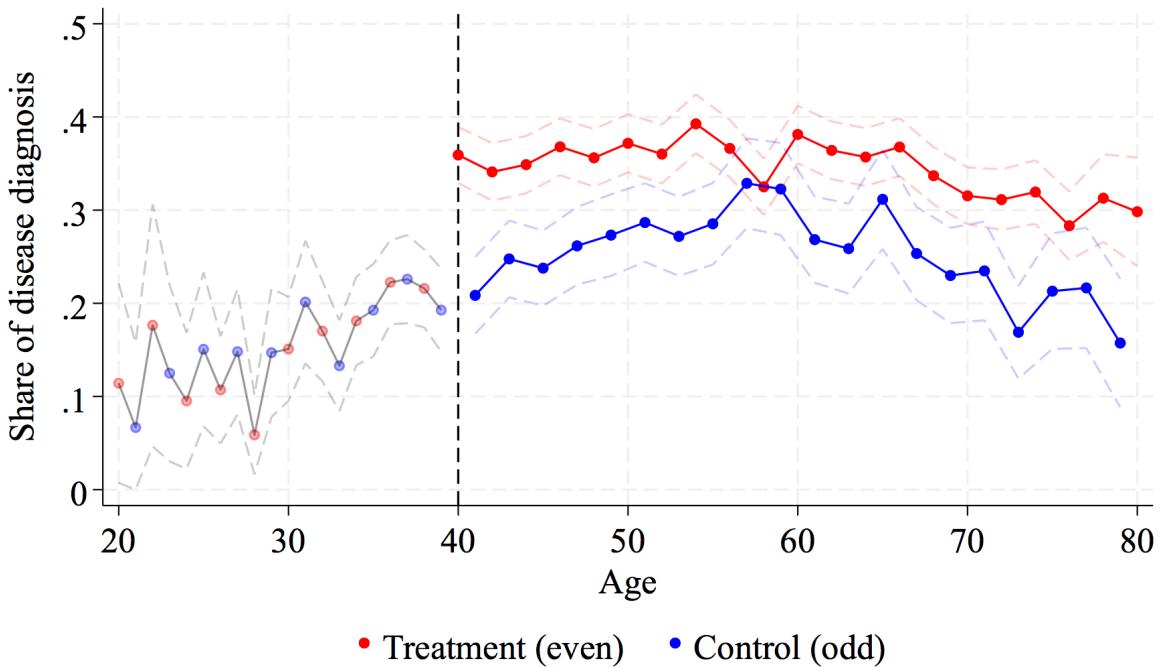
Notes: The figure plots the take-up of screening of any kind for 4 groups of the currently married people. The sample consists of individuals who are currently married and whose age is in the range [30, 80]. The legend Even/Odd refers to a group where one's own age is even, hence eligible for subsidies, and the spouse's age is odd, not eligible for subsidies. The other 3 groups are similarly defined. 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

(a) Comparing before and after age 40 using 2 year age bins



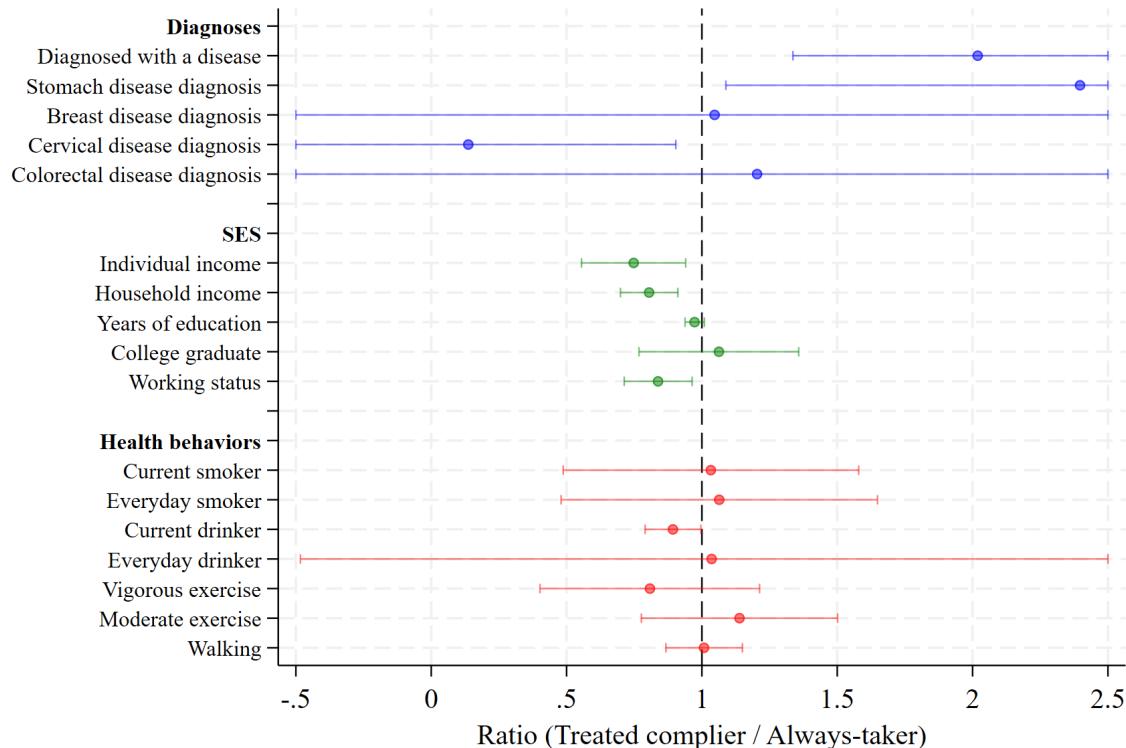
(b) Comparing even with odd age group using each age



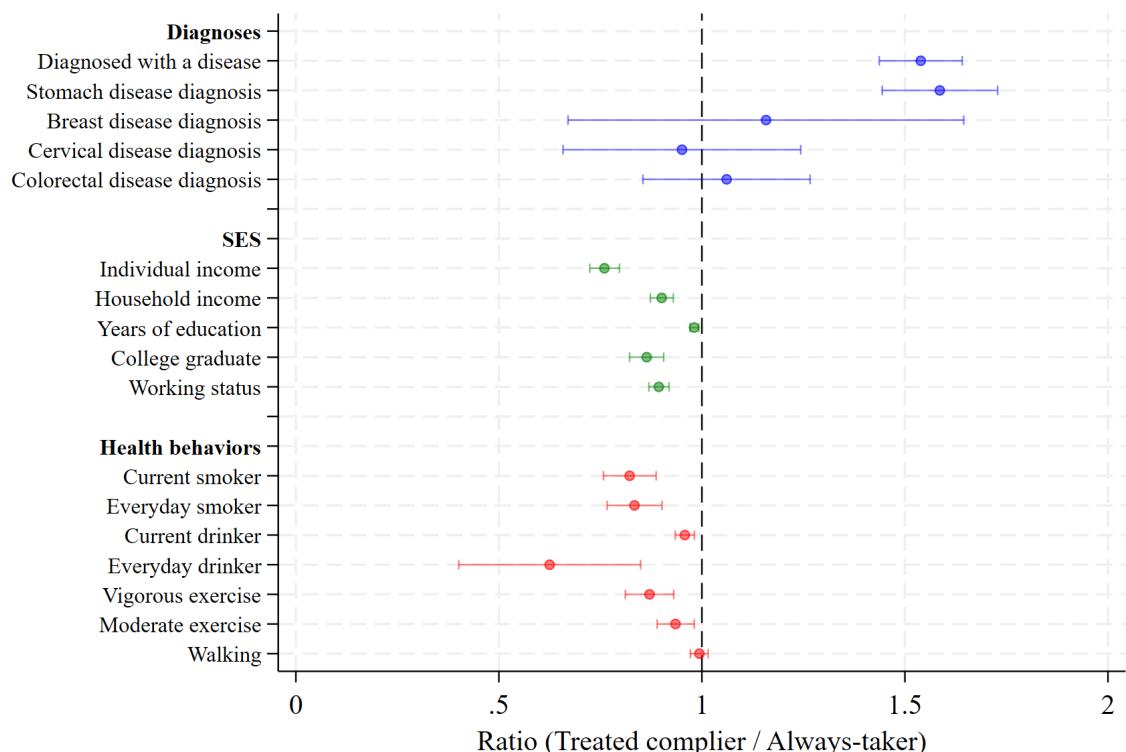
Notes: Figures plot the share of screenings where participants were diagnosed with or showed manifestations of a disease. The sample is restricted to screening participants. For each age or 2-year age bin, the figures plot the share of screening participants who reported they were diagnosed with a disease or had manifestations of a disease that require further examinations. The types of diseases found were coded using Korean Classification of Diseases (Korean version of ICD-10) and some examples are reported in Appendix section E. In the bottom figure, even ages are colored in red and odd ages are colored in blue. Dashed vertical line shows the subsidy starting age at 40. Confidence intervals at 95 percent are shown in dashed line

Figure 5: Comparison between treated compliers and always-takers

(a) Comparing before and after age 40 using 2 year age bins



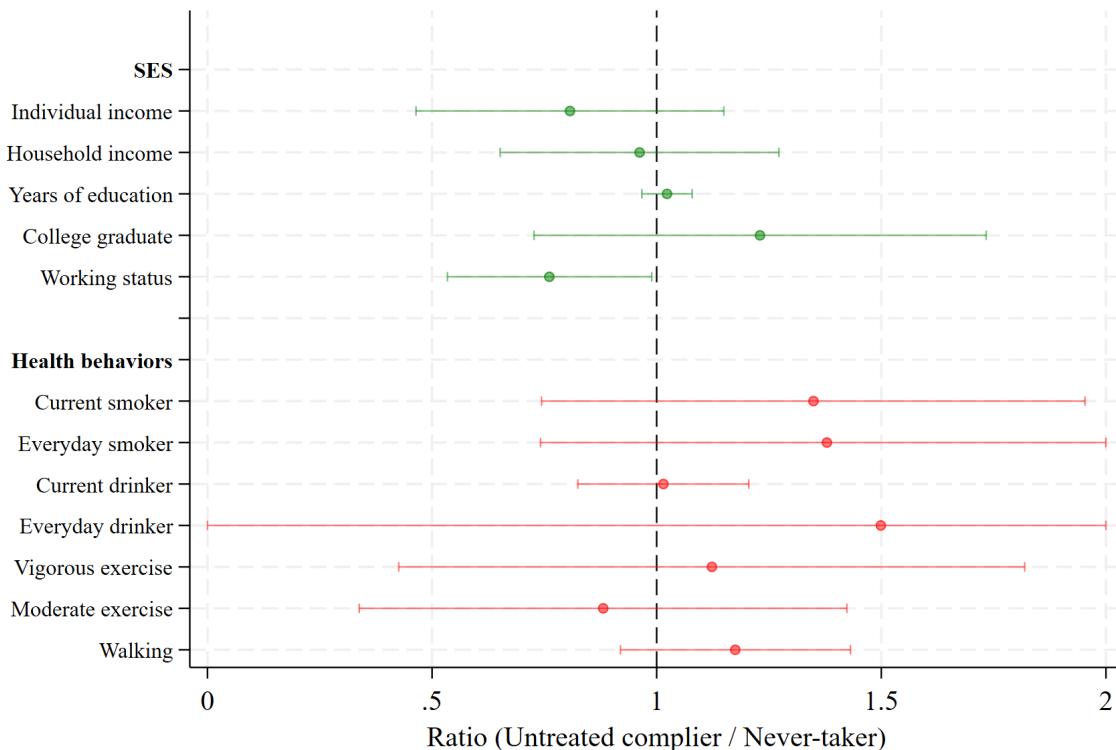
(b) Comparing even with odd age group using each age



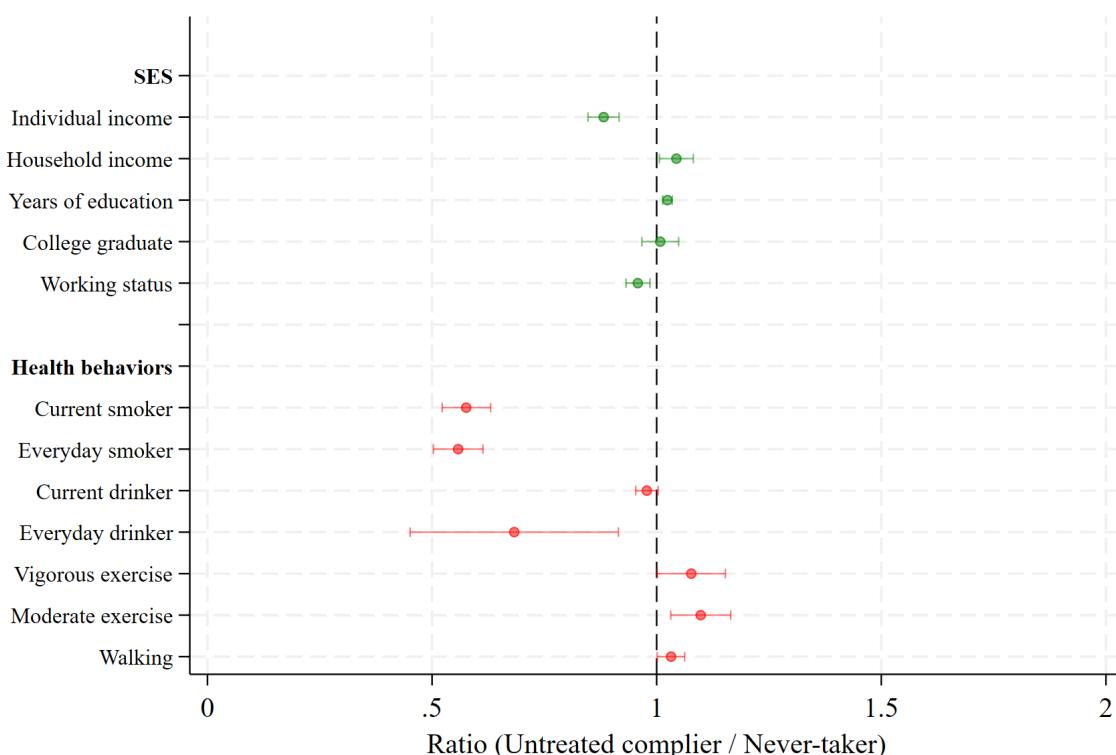
Notes: Figures plot the ratio of average characteristics of treated compliers to always-takers with 95 percent confidence intervals. Treated compliers are compliers in the treatment group who participate in screening. The average values and ratios are calculated from the estimation of equation (7) and are reported in Table ???. 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

(a) Comparing before and after age 40 using 2 year age bins



(b) Comparing even with odd age group using each age



Notes: Figures plot 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 (7) and are reported in Table ???. 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 analysis on intertemporal substitution. The screening schedule that provides subsidies at even ages give individuals an incentive to receive screening at even ages when one is eligible for subsidies. Section 5.1 provides a lower and an upper bound of subsidy effect by binning the ages differently. The goal of this section is to find evidence of whether substitution exists or not by examining people on the margin of subsidies. I present 3 pieces of evidence by examining following subsamples: (i) cervical screening participants with ages 20 to 29 before and after the expansion of subsidy starting age, (ii) stomach and breast screening participants with ages 36 to 43 before and after the subsidies, and (iii) people receiving stomach and breast screening in January and December of odd ages. They all represent people who are on different margins of subsidies. As subsidies start to apply, I examine their change in screening behavior to find evidence of substitution. All three pieces of evidence point limited substitution in screening take-up.

First evidence comes from the expansion of cervical screening subsidies. Until 2016, it was biennially subsidized at even ages from age 30. Starting from 2016, the subsidy starting age was lowered to 20. Hence, women with even age between 20 and 29 start to become eligible for subsidies from 2016. Table A1 examines if these women exhibit drop in take-up at odd ages. The sample is women with age from 20 to 29 and year 2013 to 2018. Column 1 examines take-up between even and odd ages before and after 2016. The insignificant coefficient of Age even variable means there was no significant difference in take-up between even and odd ages before the expansion. After the policy change, even age group is 1 percentage point more likely to participate in screening as shown in the sum of Age even and Age even \times After 2016 coefficients and it is significant at 5% level. Our coefficient of interest is After 2016, that is, the change in screening take-up for the odd age group. Positive coefficient estimate suggests rules out substitution. Column 2 further controls for the linear trend in screening take-up. Although the coefficient of After 2016 turns negative, it is still small and insignificant. This suggests that expansion of subsidies increase the take-up in even age group, but hardly has any negative impact

on odd age group.

Second, I focus on cohorts around age 40 to examine how screening take-up changes as one passes age 40 threshold and becomes eligible for subsidies. A sufficient sign of intertemporal substitution would be drop in take-up at odd ages after age 40. However, as new people start to participate from age 40, this opposing force makes it hard to detect a drop in screening rate. Therefore, I focus on those who were already participating before 40 to examine if they exhibit any drop at odd ages after 40. I track 4 age cohorts at ages 36 to 43. Panel A and B in Figure A2 present stomach and breast screening take-up pattern for participants at age 36, 37, 38 and 39 separately. By definition, age 36 participants show take-up rate of one at age 36.⁴² Comparing before and after 40, one can see that take-up at odd ages after 40 is not any lower than pre-40 take-up level, but the take-up at even ages are clearly much higher. This suggests that as one passes age 40, take-up at even ages rise due to subsidies and it does not come at the cost of drop in take-up at odd ages.

Lastly, I examine monthly distribution of take-up in stomach and breast screening. If there were intertemporal substitution, then it would be most pronounced in January or December of the odd age. One can get screening a couple of weeks early and receive it in December of even age instead of January of odd age. Similarly, one can delay it a couple of weeks and receive it in January of even age instead of December of odd age. Hence, we expect lower number of screenings in January and December of odd age compared to months in the middle.

To make it more rigorous, I employ Difference-in-differences design and compare the monthly change in take-up before and after age 40 to take into account the underlying distribution of monthly take-ups in the absence of subsidies. Using the exact day of screening information, I transform the individual-year data into individual-month-year

⁴²One reason I do not examine those who participated at least once before age 40 is due to age 39 participants. By definition, they show average take-up of one at age 39. Therefore, the take-up pattern for those who participated at least once before 40 shows abnormally large take-up at age 39 and makes it hard to compare before and after 40.

data and run the following econometric specification.

$$\begin{aligned}
screen_{imt} = & \theta_0 + \theta_1 \cdot after40_{imt} + \theta_2 \cdot age_even_{imt} + \sum_{m=2}^{12} month_m \\
& + \theta_3 \cdot after40_{imt} \cdot age_even_{imt} + \sum_{m=2}^{12} month_m \cdot after40_{imt} \\
& + \sum_{m=2}^{12} month_m \cdot age_even_{imt} + \sum_{m=2}^{12} month_m \cdot after40_{imt} \cdot age_even_{imt} + \varepsilon_{imt}
\end{aligned} \tag{6}$$

It is a fully saturated model of the following variables: (i) $after40_{imt}$, a dummy variable that equals one if the age of individual i is 40 or above in month m and year t , (ii) age_even_{imt} , a dummy variable that equals one if the age of individual i is even numbered in month m and year t , and (iii) $\{month_m\}_{m=2}^{12}$, dummy variables for months using January as the reference category. The outcome variable, $screen_{imt}$, is a dummy variable that equals one if individual i received the stomach or breast screening in month m year t . The standard errors are clustered at the individual level. Our first variable of interest is $age_even_{imt} + \sum_{m=2}^{12} month_m \cdot age_even_{imt}$ that provides comparison in monthly take-up between even and odd ages before 40 and there should be no systematic difference, since it is before subsidies start to apply. The second variable of interest is $\sum_{m=2}^{12} month_m \cdot above40_{imt}$ that provides comparison in monthly take-up of odd ages before and after 40. If intertemporal substitution were most pronounced in the beginning and end of odd ages, we should observe an inverted U-shape. The increase in screening take-up should be the smallest in January and December.

Figure A2a and A2e present the average monthly take-up of screening for even and odd ages. To compare easily, I present even ages again after the end of odd ages. The take-up in every month of even age is clearly larger than that in months of odd age. While there seems to be not much monthly variation in odd age, the take-up in even age seems to grow from January to December with another local spike in March. This is because many regional offices of National Health Insurance Service send reminder mails for health screenings in March and April. The spike in December seems to be due to procrastination rather than substitution as will be shown below.⁴³

⁴³Some firms enforce employees to receive general screening when one is eligible, since workers failure

The second figures in each panel, Figure A2b and A2f, plot the monthly difference in screening take-up between even and odd ages before 40. There seems to be negligible differences as expected, since this is period before subsidies. The third figures, Figure A2c and A2g, plot the monthly difference in screening take-up in odd ages before and after 40. We do not find inverted U-shape pattern. The take-up in other months, compared to January, is not statistically different. Hence, we fail to find significant intertemporal substitution across year thresholds. We may expect the substitution to be stronger for low income households, since they are more responsive to subsidies. Therefore, I present the same figures, the odd age take-up before and after 40, for households with income less than median. The fourth figures in each panel, Figure A2d and A2h, show there is no discernible difference across months. This suggests limited substitution even among low income households.

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 explain procrastination and the hike in take-up in December.

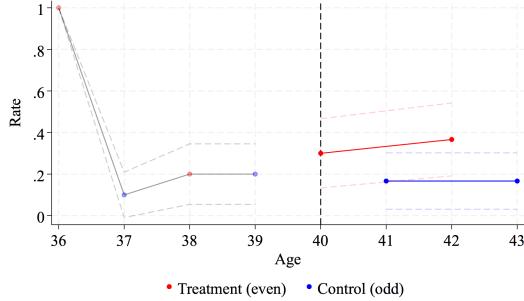
Table A1: Intertemporal substitution using the expansion in cervical screening subsidies

	(1)	(2)
Outcome var: cervical screening		
Age even	0.002 (0.003)	0.002 (0.003)
After 2016	0.006* (0.004)	-0.003 (0.006)
Age even × After 2016	0.008 (0.006)	0.008 (0.006)
Year		0.003* (0.002)
Age even + Age even × After 2016	0.010** (0.005)	0.010** (0.005)
N	5221	5221
Adj R^2	0.003	0.003

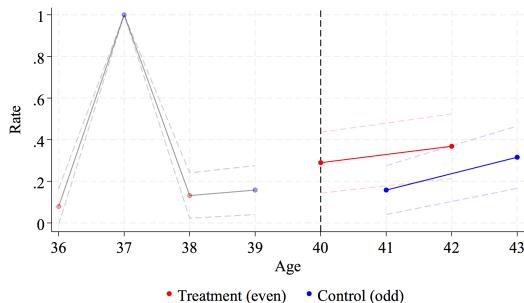
Notes: This table reports the change in cervical screening take-up between even and odd ages before and after 2016. The sample is women with age from 20 to 29 and year from 2013 to 2018. Cervical screening subsidies were provided to age 30 and above, but it was lowered to 20 and above from year 2016. After 2016 is a variable capturing year 2016 to 2018, the period after the policy change. Year is a continuous year variable to capture the linear trend in cervical screening take-up. Standard errors are clustered at the individual level and reported in parentheses. A * / ** / *** indicates significance at the 10/5/1% levels.

Figure A1: Screening take-up for participants before 40

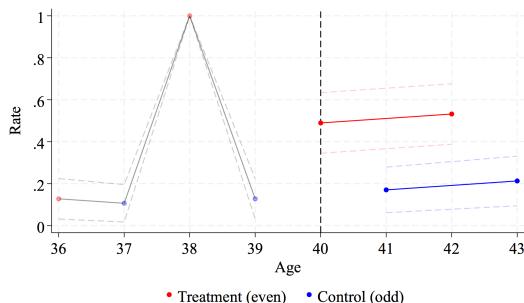
Panel A. Stomach screening participants



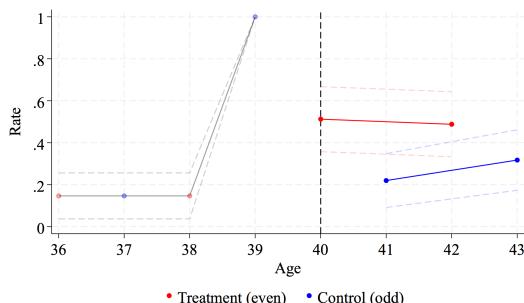
(a) Age 36 participants



(b) Age 37 participants

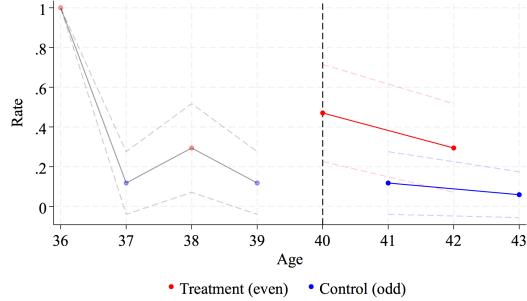


(c) Age 38 participants

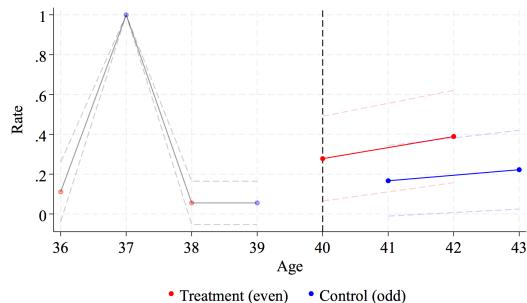


(d) Age 39 participants

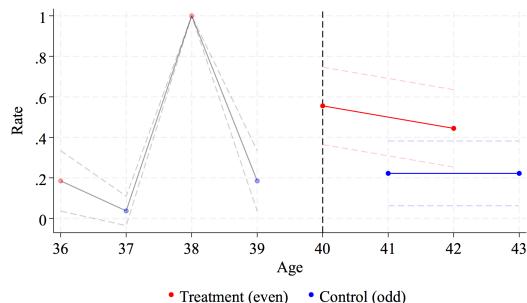
Panel B. Breast screening participants



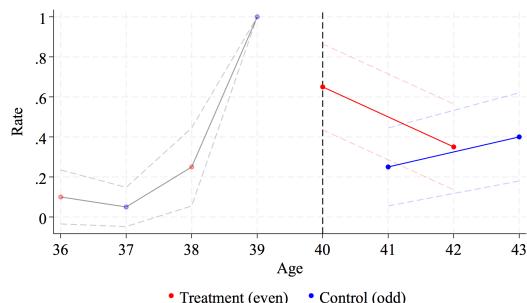
(e) Age 36 participants



(f) Age 37 participants



(g) Age 38 participants

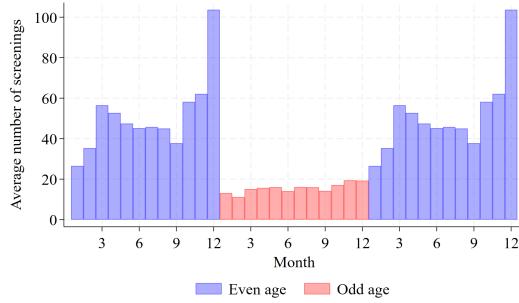


(h) Age 39 participants

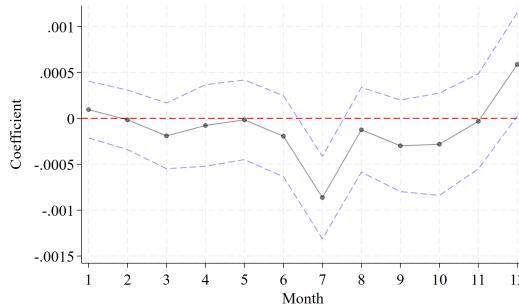
Notes: Figures plot the stomach and breast screening take-up for those who already participate in screening before age 40. The sample is restricted to four age cohorts around age 40. Each figure plots the take-up for either the stomach or breast cancer participants at age 36, 37, 38 or 39. Even ages are colored in red and odd ages are colored in blue. 95 percent confidence intervals are shown in dashed line.

Figure A2: Monthly screening take-up

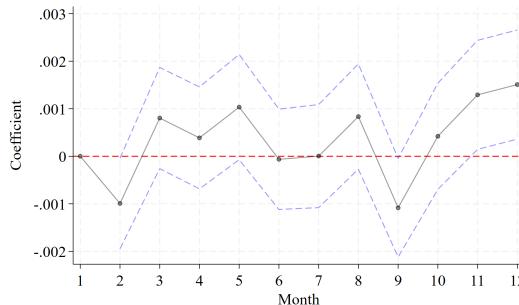
Panel A. Stomach screening



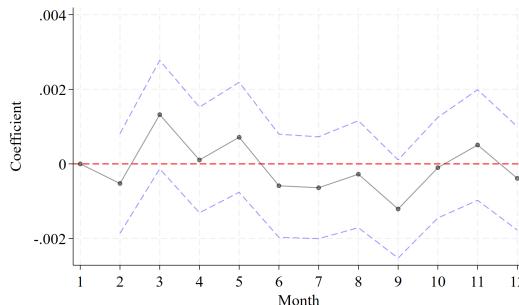
(a) Average monthly take-up



(b) Even vs odd age take-up before 40

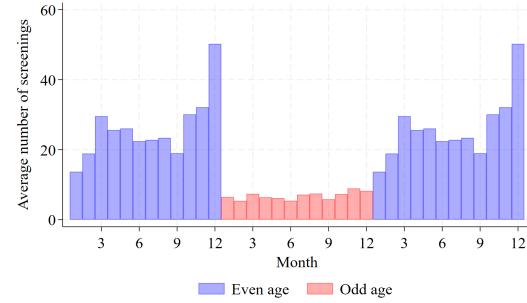


(c) Odd age take-up before and after 40

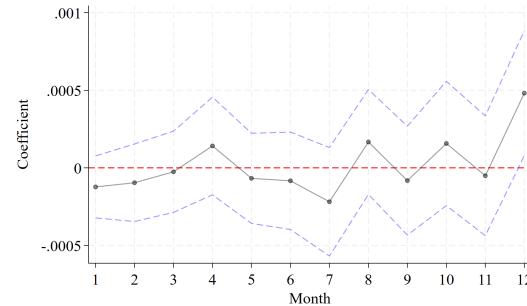


(d) Low income odd age take-up before and after 40

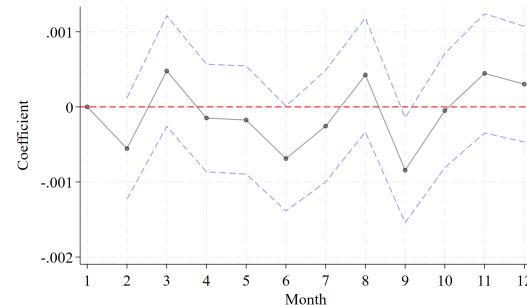
Panel B. Breast screening



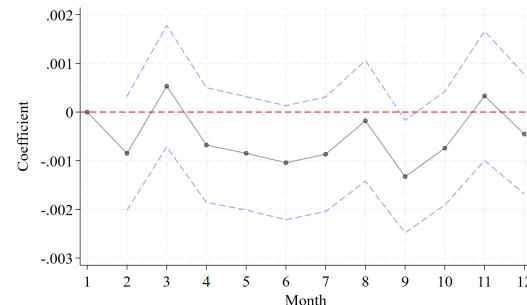
(e) Average monthly take-up



(f) Even vs odd age take-up before 40



(g) Odd age take-up before and after 40



Notes: Figures plot the analysis results using the individual-month-year level data. First figures in each panel plots the average monthly take-up at even and odd ages. Even age group is repeated after odd age to provide an easy comparison. Second figures plot the monthly differences in take-up between even and odd age group before age 40. Third figures plot the monthly differences in take-up in odd age group before and after age 40 using January as the reference month. Four figures plot the same monthly differences in take-up in odd age group before and after age 40 using only the households with below median income.

Appendix B Cross spillover

This section provides additional analysis on the mechanism of cross 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. Another mechanism is that receiving a general health screening, the one with the highest take-up, can lead to other screenings, since general screening includes consultation with a doctor.

Table [A2](#) presents the share of annual and no-subsidy screenings that happen on the same day with general screening. Given the take-up of general screening, 85-96 percent of annual and no-subsidy screenings are received on the same day. If they are not received on the same day, most of them are received after the general screening consistent with the hypothesis that general screening is the one generating spillover to other screenings. Table [A3](#) also show that even age group is more likely to receive annual and no-subsidy screenings on the same day with general screening.

To examine which of the 4 biennial screenings is generating spillover effect, I examine gender differences 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 [A4](#) does not support this hypothesis. If anything, they seem to be slightly smaller for the colorectal screening. 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.

Table A2: Cross spillover: share of screenings by screening date

	(1)	(2)	(3)	(4)
	Liver	Colorectal	Prostate	Lung
Pr(general = 1 screen = 1)	0.844	0.799	0.786	0.699
Pr(same day screen = 1, general = 1)	0.948	0.856	0.960	0.937
Pr(general first screen = 1, general = 1)	0.036	0.120	0.024	0.047
Pr(general later screen = 1, general = 1)	0.008	0.178	0.004	0.002

Notes: This table examines if people receive screenings on the same day with the general health screening. The sample is those with age 40 to 89. *screen* = 1 refers to the take-up of liver, colorectal, lung and prostate screening in each column. General first (later) means the screening concerned is received after (before) the general screening within 30 days.

Table A3: Cross spillover by screening day relative to general screening

	(1)	(2)	(3)	(4)
	Annual subsidy		No subsidy	
	Liver	Colorectal	Prostate	Lung
Panel A. Outcome: conducted on the same day with general screening				
Age even	0.023*** (0.001)	0.024*** (0.001)	0.005*** (0.001)	0.004*** (0.001)
N	107183	107183	50260	107183
Control group mean	0.022	0.017	0.007	0.006
Panel B. Outcome: conducted after general screening				
Age even	0.0012*** (0.0002)	0.0040*** (0.0004)	0.0001 (0.0001)	0.0005*** (0.0001)
N	107183	107183	50260	107183
Control group mean	0.0007	0.0022	0.0002	0.0002
Panel C. Outcome: conducted before general screening				
Age even	0.0003*** (0.0001)	0.0064*** (0.0005)	0.0001 (0.0001)	0.0000 (0.0000)
N	107183	107183	50260	107183
Control group mean	0.0001	0.0029	0.0000	0.0000
Sample age range	[40, 89]	[40, 89]	[40, 89]	[40, 89]
Subsidy starting age	40	50		
Age controls	Y	Y	Y	Y

Notes: This table examines if people receive screenings on the same day with general health screening. Take-up before or after general screening considers 30 days window before and after the screening day. Panel A, B and C present the difference in the take-up of each screening that happened on the same day, after, and before general screening between even and odd age groups. The sample is those with age 40 to 89. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A4: Gender differences in cross spillover

	(1)	(2)	(3)
	Liver	Colorectal	Lung
Age even	0.025*** (0.002)	0.036*** (0.002)	0.007*** (0.001)
Age even × Female	0.003 (0.003)	-0.005* (0.003)	-0.002 (0.001)
Female	-0.017*** (0.002)	-0.012*** (0.002)	-0.008*** (0.001)
N	107183	107183	107183
Control group mean	0.028	0.027	0.009

Notes: This table reports estimates of cross 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, share of couples receiving screening on the same day 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 A5 shows that the spillover effects run from wives to husbands. I split the sample into husbands and wives and estimate the equation (3) separately. First two columns show that 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 A6 presents the probability of couples receiving screening on the same day. Conditional on both wife's and husband's participation, more than 40 percent receive screening on the same day. This is more likely to be the case when both the husband and the wife are eligible for subsidies or when both are ineligible for subsidies. If the two do not receive it on the same day, they are equally likely to receive it earlier or later than the other.

Table A7 presents the estimates of spillover effect for each screening type. It shows that spousal spillover does not exist for breast, cervical and prostate screenings. It could be because these are women or men only screenings and are less likely to be received together.

Table A5: Spousal spillover directions

	(1)	(2)	(3)	(4)
	Among wives (husband \Rightarrow wife)		Among husbands (wife \Rightarrow husband)	
Age even	0.220*** (0.004)	0.219*** (0.004)	0.142*** (0.004)	0.141*** (0.004)
Spouse age even	0.006 (0.004)		0.017*** (0.004)	
Spouse screening		0.046 (0.030)		0.079*** (0.017)
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 males. In column 2 and 4, spouse screening variable is instrumented by spouse even age variable. No control variables are included in the regressions. Standard errors are clustered at the couple level and reported in parentheses. A */**/** indicates significance at the 10/5/1% levels.

Table A6: Spousal spillover: share of same day screenings

	(1)	(2)	(3)	(4)	(5)
	Total	Even/Even	Even/Odd	Odd/Even	Odd/Odd
Pr(same day both participate)	0.423	0.494	0.303	0.362	0.462
Pr(Spouse first both participate)	0.114	0.132	0.095	0.105	0.096
Pr(Spouse later both participate)	0.114	0.134	0.088	0.113	0.091

Notes: This table reports the share of spouses getting screening on the same day given that both participate in screening in a year. Take-up before or after spouse considers 30 day window before and after the screening day. Column 1 reports the share of the couples who participate in a same year. Columns 2 to 5 reports the same share of couples by each even and odd combination.

Table A7: Spousal spillover by screening types

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	General	Stomach	Liver	Colorectal	Lung	Breast	Cervical	Prostate
Age even	0.163*** (0.003)	0.165*** (0.003)	0.023*** (0.001)	0.030*** (0.001)	0.005*** (0.001)	0.168*** (0.004)	0.156*** (0.003)	0.007*** (0.001)
Spouse age even	0.010*** (0.003)	0.012*** (0.003)	0.002 (0.001)	0.004*** (0.001)	0.001 (0.001)	0.001 (0.004)	0.000 (0.003)	-0.001 (0.001)
N	101726	101726	101726	101726	101726	50863	50863	50863

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. Standard errors are clustered at the couple level and reported in parentheses. A */**/** indicates significance at the 10/5/1% levels.

Appendix D Complier selection

D.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 [Kim and Lee \(2017\)](#); [Einav et al. \(2020\)](#) and [Kowalski \(2023\)](#).⁴⁴ The key idea is to infer always-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. 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 present detailed steps to infer complier characteristics in the even vs odd design. A similar design can be used to characterize compliers in the regression discontinuity setting.⁴⁵ I estimate the following equation to estimate group characteristics.

$$y_{it} = \lambda_1 age_even_{it} + \lambda_2 screen_{it} + \lambda_3 age_even_{it} \times screen_{it} + \boldsymbol{\lambda}_4' \mathbf{f}(\mathbf{age}_{it}) + \nu_{it} \quad (7)$$

Given a characteristic variable, y_{it} , the above equation can be used to estimate the average characteristic of always-takers by imposing the condition $age_even_{it} = 0$ 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 $g_{AT}(y) = \hat{\lambda}_2 + \hat{\boldsymbol{\lambda}}_4' \mathbf{f}(\mathbf{age}_{it})$. 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 resulting estimates are given by $g_{NT}(y) = \hat{\lambda}_1 + \hat{\boldsymbol{\lambda}}_4' \mathbf{f}(\mathbf{age}_{it})$.

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

⁴⁴[Marbach and Hangartner \(2020\)](#) gives a nice summary of the methodology.

⁴⁵Appendix in [Einav et al. \(2020\)](#) present the detailed steps to characterize compliers at age 40 in the regression discontinuity setting.

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 characteristics 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, respectively. Imposing $age_even_{it} = 1$ and $screen_{it} = 1$, we get $g_T(y) = \hat{\lambda}_1 + \hat{\lambda}_2 + \hat{\lambda}_3 + \hat{\lambda}_4' \mathbf{f}(\text{age}_{it})$. The share of always-takers and compliers can be calculated from the first stage regression given in (2) as $\pi_C = \hat{\beta}_1$, $\pi_{AT} = \hat{\beta}_0$ adjusting for age.⁴⁶ Inserting all the estimated terms to the equation $g_C^1(y) = [(\pi_{AT} + \pi_C)g_T(y) - \pi_{AT}g_{AT}(y)]/\pi_C$, the complier characteristic can be backed out. The average characteristics for untreated compliers can be calculated in a similar way.⁴⁷

To characterize compliers in reference to always-takers and never-takers, I take ratios between treated compliers and always-takers, $\frac{g_C^1(y)}{g_{AT}(y)}$, and between untreated compliers and never-takers, $\frac{g_C^0(y)}{g_{NT}(y)}$.⁴⁸ Standard errors are calculated using bootstrap with 500 replications clustering at the individual level. The null hypothesis used for ratios is that the ratio is equal to one.

⁴⁶Under monotonicity, the share for never-takers is $\pi_{NT} = 1 - \pi_{AT} - \pi_C = 1 - \hat{\beta}_0 - \hat{\beta}_1$

⁴⁷I examine selection pattern at age 40 by imposing $age_{it} = 40$. The results are robust to the different choices of age.

⁴⁸The 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.

D.2 Selection using panel information in even-odd design

This section provides an alternative way of characterizing compliers using the repeated experiments. Biennial subsidy rule provides screening subsidies when one's age is even and this affects screening take-up pattern of not only the biennial screenings, but also annual and no-subsidy screenings. Compliers are the ones who respond to subsidies and participate in screening when subsidies are provided but do not participate when there is no subsidy. Hence, another way, or perhaps a more intuitive way, to define compliers is to use the history of 11 years information altogether and find those who participated in even ages and not in odd ages. Other compliance groups can be defined similarly. Always(Never)-takers are the ones who (do not) participate both at even and odd ages. Defiers can be defined as ones who participate at odd ages but do not participate at even ages.

An advantage of the alternative definition is that it allows checking if the compliance behavior holds over time. The compliance group defined originally in [Imbens and Angrist \(1994\)](#) and [Angrist et al. \(1996\)](#) is based on potential outcome framework and does not take into account any panel structure. In our setting, we have the same experiments happening over time and panel data to capture the evolution of compliance behaviors. An individual receives subsidies at age 42, one experiment between even and odd ages. The same experiment happens next year at age 43, this time in the control group, and again at 44, back in the treatment group. If one participated at age 42 implying that she is either an always-taker or a complier, does she also participate at age 44? It is absurd to treat experiment at age 42 and 44 as identical, but it is also not unrealistic to expect the behaviors to be correlated over time. Whether repeated experiments happening at different times can serve as a valid counterfactual is by no means theoretically grounded, but it is an interesting exercise and provides a nice robustness check on cross sectional analysis.

To avoid attrition issue, I restrict the sample to unattrited survey participants that show up in the dataset for the entire 11 years. I also impose the age condition such that the age in the first year is 40 or above. This leaves unique 5,514 individuals. For

classification, I devise two scores based on the following rule.

$$\begin{aligned} even_score_i &= \sum_t [\mathbb{1}\{screen_{ia} = 1\} - \mathbb{1}\{screen_{ia} = 0\}], & a \text{ even} \\ odd_score_i &= \sum_t [\mathbb{1}\{screen_{ia} = 1\} - \mathbb{1}\{screen_{ia} = 0\}], & a \text{ odd} \end{aligned} \tag{8}$$

The idea is to score screening take-up at even and odd years separately. For those with an even age in the first year, their *even_score* will range from -6 to 6 with only even numbers and their *odd_score* will range from -5 to 5 with only odd numbers. Perfect compliance means participating at all the even ages and not participating at all the odd ages, which gives us maximum even score and minimum odd score. Defiers will have the opposite pattern, that is, minimum even score and maximum odd score. Always-takers will have both the maximum even and odd scores and never-takers will have both the minimum even and odd scores. For those with even age in the first year, the bivariate distribution with the shares is given in Figure A3. For exhaustive definition, I define those in the first quadrant ($even_score > 0, odd_score > 0$) as always-takers, second quadrant ($even_score < 0, odd_score > 0$) as defiers, third quadrant ($even_score < 0, odd_score < 0$) as never-takers and fourth quadrant ($even_score > 0, odd_score < 0$) as compliers. Those with zero even or odd scores are randomly assigned to the two adjacent groups.

Table A8 presents the characteristics of four compliance groups and the ratios compared to compliers. The final shares of the compliance groups are given at the bottom of the table. The group shares are quite similar to the ones calculated from the first stage equation (2). One difference is the presence of defiers which is assumed away with monotonicity assumption in Section 5.1. Given that they are the ones who participate in odd years and do not participate in even years, they would be classified as either always-takers or never-takers in the cross sectional analysis. The share of defiers is small suggesting validity of the monotonicity assumption. One explanation for defiers comes from spousal spillover discussed in Section 5.3. Despite odd age, if a spouse has even age, it increase one's probability of screening giving rise to defiance behavior. I further

restrict the sample to 4,612 married individuals check the share of people whose age is even when the spouse's age is odd or whose age is odd when the spouse's age is even. Consistent with our expectation, the share of off-age combination is highest in the defier group and lowest in the complier group.

The selection pattern shown in Table A8 is highly consistent with the one from cross sectional analysis. Column 5 provides comparison between compliers and always-takers. Panel A shows compliers are more likely to find a disease than always-takers. This pattern is shown for all the screenings, while it was only the case for stomach screening in cross sectional analysis. This is consistent with negative selection in income and education as shown in Panel B. Due to negative selection in income, Panel C shows compliers are less likely to smoke, drink and exercise. One difference was that they are more likely to be everyday drinker.

Comparison with never-takers also shows consistent pattern. Compliers show better health behaviors than never-takers. They are less likely to smoke, less likely to drink everyday and more likely to exercise. In terms of socioeconomic status, we only again get mixed finding and unclear selection pattern. It is interesting that compliers seem to be slightly more likely to find a stomach and breast disease than never-takers. However, this would be clearly sensitive to classification. Never-takers, by definition, do not participate in screening. Their diagnosis effect is identified off of those who are on the margin with other compliance groups. With stricter classification, the number of screenings participated by never-takers would be small making the comparison difficult.

Overall, the selection pattern from using panel information is similar to the one reported in Table 7 and 8 using cross sectional variation. This gives confidence to my results and also suggest that compliance behavior may hold over time.

Figure A3: Bivariate distribution of the even and odd score

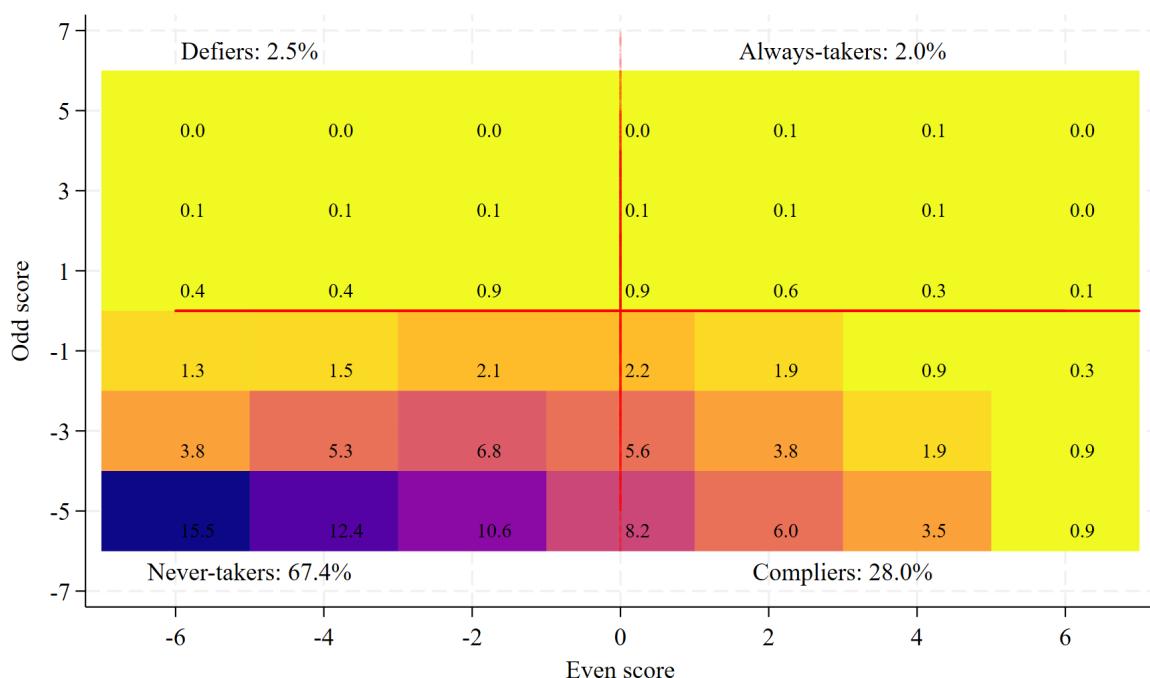


Table A8: Compliers with subsidies using panel information in even-odd design

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Always	Compliers	Defiers	Never	Compliers / Always	Compliers / Defiers	Compliers / Never
Panel A. Diagnoses							
Diagnosed with a disease	0.277 (0.018)	0.350 (0.006)	0.306 (0.019)	0.321 (0.007)	1.264*** (0.084)	1.146*** (0.076)	1.093*** (0.031)
Stomach disease diagnosis	0.141 (0.014)	0.203 (0.006)	0.154 (0.016)	0.182 (0.006)	1.435*** (0.145)	1.314*** (0.139)	1.114*** (0.047)
Breast disease diagnosis	0.011 (0.004)	0.011 (0.002)	0.012 (0.006)	0.009 (0.002)	1.005** (0.422)	0.933** (0.471)	1.241*** (0.297)
Cervical disease diagnosis	0.030 (0.007)	0.031 (0.003)	0.026 (0.010)	0.031 (0.004)	1.032*** (0.264)	1.168*** (0.445)	0.999*** (0.146)
Colorectal disease diagnosis	0.042 (0.008)	0.042 (0.003)	0.048 (0.009)	0.042 (0.003)	1.001*** (0.207)	0.886*** (0.180)	0.993*** (0.095)
Panel B. SES							
Individual income	2456 (244)	990 (37)	2391 (209)	1225 (27)	0.403*** (0.043)	0.414*** (0.039)	0.808*** (0.035)
Household income	5817 (313)	3862 (67)	5443 (271)	3634 (44)	0.664*** (0.038)	0.710*** (0.037)	1.063*** (0.023)
Years of education	11.949 (0.359)	9.789 (0.099)	11.476 (0.343)	9.445 (0.075)	0.819*** (0.026)	0.853*** (0.027)	1.036*** (0.013)
College graduate	0.279 (0.040)	0.115 (0.008)	0.271 (0.037)	0.137 (0.006)	0.411*** (0.066)	0.424*** (0.065)	0.840*** (0.068)
Working status	0.742 (0.031)	0.555 (0.010)	0.729 (0.031)	0.598 (0.007)	0.748*** (0.034)	0.761*** (0.036)	0.928*** (0.020)
Panel C. Health behaviors							
Current smoker	0.128 (0.024)	0.099 (0.007)	0.222 (0.031)	0.207 (0.006)	0.779*** (0.158)	0.447*** (0.068)	0.479*** (0.035)
Everyday smoker	0.118 (0.023)	0.096 (0.007)	0.208 (0.030)	0.200 (0.006)	0.811*** (0.171)	0.460*** (0.073)	0.478*** (0.036)
Current drinker	0.701 (0.034)	0.601 (0.010)	0.705 (0.032)	0.593 (0.007)	0.858*** (0.044)	0.853*** (0.041)	1.013*** (0.020)
Everyday drinker	0.038 (0.011)	0.057 (0.004)	0.096 (0.018)	0.080 (0.003)	1.479*** (0.443)	0.589*** (0.120)	0.702*** (0.062)
Vigorous exercise	0.298 (0.022)	0.200 (0.005)	0.272 (0.019)	0.191 (0.004)	0.671*** (0.053)	0.735*** (0.056)	1.050*** (0.035)
Moderate exercise	0.498 (0.020)	0.385 (0.006)	0.462 (0.023)	0.338 (0.004)	0.773*** (0.034)	0.833*** (0.043)	1.141*** (0.023)
Walking	0.843 (0.016)	0.816 (0.005)	0.817 (0.013)	0.771 (0.003)	0.969*** (0.019)	0.999*** (0.017)	1.059*** (0.008)
Panel D. Married subsample							
Pr(even/odd or odd/even)	0.505	0.477	0.616	0.500			
Share	0.022	0.283	0.026	0.669			

Notes: This table reports the average characteristics of compliance groups defined using the history of health screenings in 11 years. Compliance groups are defined using the rule shown in Equation (8) based on how many times one participated in screening at even and odd ages. Compliers are the ones who participate more than half at even ages and less than half at odd ages. Always-takers are the ones who participate more than half at both the even and odd ages. Never-takers are the ones who participate less than half at both the even and odd ages. Defiers are the ones who participate more than half at odd ages and less than half at even ages. The sample is restricted to those who do not show any attrition and participated in the survey for all 11 years. It is also restricted to those whose age in the first year is 40 or above. After classification, the sample consists of 5,514 individuals and all the characteristic variables are averages of 11 year values. Pr(even/odd or odd/even) refers to the probability that one's age and the spouse's age is even-odd or odd-even. For this share calculation, the sample is further restricted to 4,612 individuals who are married. Share reports the share of each compliance group using the whole sample. Robust standard errors are used. They are reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

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 the Korean Classification of Diseases diagnosis code (Korean version of the ICD-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 Effect of health screening

Table A9: Effect of screening on first outpatient visits using spousal spillover

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
First outpatient visit	High blood pressure	Hyperlipidemia	Diabetes	Stomach	Breast	Female genital	Liver	Colorectal	Male genital	Lung	
Panel A. Reduced form regressions											
Age even	0.0014 (0.0179)	0.0032*** (0.0012)	-0.0005 (0.0012)	0.0245*** (0.0035)	0.0014 (0.0014)	0.0075** (0.0035)	0.0001 (0.0008)	0.0047** (0.0022)	-0.0020 (0.0031)	0.0008 (0.0012)	
Spouse age even	0.0198 (0.0179)	0.0017 (0.0017)	0.0010 (0.0012)	0.0015 (0.0012)	0.0046 (0.0035)	0.0008 (0.0014)	0.0003 (0.0035)	0.0001 (0.0008)	0.0009 (0.0022)	0.0008 (0.0031)	(0.0012)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
Panel B. Second stage regressions											
Screening	0.0062 (0.0079)	0.0146*** (0.0054)	-0.0029 (0.0055)	0.1135*** (0.0164)	0.0049 (0.0054)	0.0290** (0.0138)	0.0005 (0.0040)	0.0022** (0.0103)	-0.0122 (0.0186)	0.0037 (0.0057)	
Spouse screening	0.0076 (0.0079)	0.0037 (0.0054)	0.0073 (0.0055)	0.0126 (0.0164)	0.0043 (0.0081)	0.0004 (0.0209)	0.0002 (0.0040)	-0.0017 (0.0103)	0.0045 (0.0120)	0.0035 (0.0057)	
Estimator	LATE	LATE	LATE	LATE	LATE	LATE	LATE	LATE	LATE	LATE	LATE
N	79782	79782	79782	79782	79782	39890	39890	79782	79782	39892	79782
Odd/Odd group mean	4.0076	0.0466	0.0244	0.0252	0.1947	0.0104	0.1063	0.0089	0.0819	0.0738	0.0193
Demographic controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table reports the effect of health screening on first outpatient visits using spousal spillover in screening take-up. Panel A reports the effect of one's own and spouse's subsidy eligibilities on the total number of first outpatient hospital visits and the ones by diagnosis categories. Panel B reports the effect of one's own and the spouse's screening take-ups instrumented by respective subsidy eligibilities. The diagnosis category is inferred from the Korean Classification of the Diseases diagnosis code (Korean version of the ICD-10) recorded for each visit to a hospital. All the regressions include demographic control variables that include age, gender, years of schooling, health insurance type, handicap status, working status, household income decile and survey year of oneself and the spouse. The sample consists of couples both of whose ages are between 40 and 89. Standard errors are clustered at the couple level and reported in parentheses.

A * / ** / *** indicates significance at the 10/5/1% levels.

Appendix G 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 ??, ??, ??, ??, and ??.

Lastly, I present bounding estimates using three different samples: (i) $\text{age} \in [39, 89]$, (ii) $\text{age} \in [40, 89]$, and (iii) $\text{age} \in [41, 89]$. As argued in section 3, the sample starting age is creating mechanical imbalance between the treatment and the control group. Table ?? 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 ??, ??, ??, and ??.