Effect of health screening on health care utilization and health behavior: Evidence from Korean screening policy

Siho Park

March 22, 2022

Abstract

This study provides evidence that regular health screening induces short run changes in health care usage and health behaviors. I use the eligibility for free biannual screening provided by the National Health Screening Program in Korea based on the year of birth being odd or even as IV for identification. Screening leads to a large increase in outpatient care for a new illness and substitution of visits to local clinics with visits to large hospitals offering specialized services. It also significantly decreases the visits to emergency room, but hardly affects inpatient care. The event study analyses reveal outpatient and inpatient care exhibit anticipatory downward trend, large spike in the month of screening, and gradual drop, suggesting intertemporal reallocation. Screening also induces positive behavioral modification in smoking, drinking and exercise. Lastly, I examine complier characteristics and document spillover effect in take-up between spouses.

1 Introduction

Regular health screening is advocated as one of the most important keys to healthy living and lower health care costs. The early diagnosis of a disease increases the likelihood of successful treatment, prevents premature death, and lowers health care costs. This is important for insurers and has prompted governments with national health insurance system to use public policies to encourage participation in screening. With growing interest in health, the public screening programs tend to grow offering screening to more people, more frequently, and from earlier age. Korea started offering public health screening program in the 1980s focusing only on civil servants and over time, the target expanded to include all the citizens 20 years or older in 2019. With larger coverage, the budget for health screening program grew from \$2.3 million in 2011 to \$640 million in 2021. Understanding the impact of large-scale regular health screening program matters as the program gets more sophisticated with various cancer screenings and covers longer life span of more people.

However, the current evidence provides neither comprehensive nor accurate picture of the causal impact of public screening program. Most of the empirical evidence comes from clinical studies. While the randomized controlled trials (RCT) guarantee causal estimates, the studies are usually conducted using a small sample in a controlled setting that hinders generalization to mass public. They also mask selection into screening and suffer from near perfect compliance in screening participation, which is far from real. Many other studies use observational data, but most of them overlook selection into screening. There is a strong pattern of selection in the decision to get screening and the failure to take it into consideration produces evidence that suffers from endogeneity. Lastly, there are studies that takes advantage of artificial thresholds in the design of screening programs to estimate the causal impact. However, they typically focus on narrow, specific components of screening, such as the impact of being classified as diabetic using cutoff in blood sugar level, rather than the impact of getting screening itself. Random variation in screening is scarce and analyses conditional on getting screened could result in low power if most of the impact comes from the act of getting screened itself. Additionally, they generally do not delve into temporal variation in outcome, which is important given recurring aspect of health screening.

To overcome the limitations of prior literature, this paper investigates two research questions. First, which types of people participate in national health screening program and how do economic incentives and peer affect participation? Understanding the pattern of selection into screening and what can affect this decision matters for the design of the program that is expanding to cover more people. The program would not be considered successful if it does not induce additional participation and simply ends up subsidizing those who would have gotten screening anyway. Second, what are the causal impacts of regular screening on health care utilization, medical expenditures, and health behaviors across time? Estimating the static causal effect of getting health screening per se is important and incorporating the dynamic aspect of how the effect evolves over time would provide a more complete picture of the screening program and behavioral responses.

To answer these research questions, I use the instrumental variable (IV) strategy to identify the causal effect of screening on health care utilization and health behaviors. The research design exploits exogenous variation in screening uptake from free screening provided biannually by the Korean National Health Insurance Service (NHIS) based on the year of birth.[1] The NHIS provides free health screening in an alternating fashion based on the year of birth being odd or even. People born in odd (even) - numbered years can get free screening in odd (even) - numbered years.[2] Using the eligibility for free screening as an IV, this paper estimates the local average treatment effect (LATE) of getting health screening on health care usage and health behaviors.[3] Note that those who are eligible for free screening in a year (treatment group) and those who are not (control group) switch every year. The assignment into treatment or control group is not individual specific, but is conditional on year, and one takes turn belonging to treatment

[1] I use the word 'biannual' in the sense that it is occurring once every two years. A similar and perhaps more precise word would be 'biennial'.

[2] One concern is if there is any age difference between those who are eligible for free screening and those who are not. I show there is a small age difference between the two groups in the balance table provided in Table 2. For every analysis, I also provide robustness checks where I control for various covariates including age. The results are robust to the inclusion of covariates.

[3] Note that the IV is equivalent to having an even-numbered age, since the difference between odd numbers or even numbers are always even.

and control group each year. What IV analysis internally does is compare the level of health care usage and behaviors for each calendar year and attribute the difference scaled by the first stage coefficient to the causal effect of screening. Hence, all the estimates represent short run effects that happen within a year. Since screening happens in the middle of a calendar year, not necessarily in the beginning, the estimates capture the sum of anticipatory effect and ex-post effect that show up before and after the screening, respectively.

The rich dataset on individual / household characteristics and health care usage allows me to perform additional analyses to gain deeper understanding of the regular health screening program. The granular data on every hospital visit for outpatient, inpatient, and emergency care allow conducting event study analyses at month level and examine how the demand for health care varies around the screening date. This complements the IV analysis. While IV analysis captures the aggregate behavioral changes that arise before and after the screening, event study analysis sheds light on how people react and adjust health care across time. I also use the detailed individual and household characteristics including the relationship between household members to examine complier characteristics and spillover effect in screening take-up between spouses.

This paper uses Korean Health Panel Survey data, individual level panel survey data of nationally representative households from 2009 to 2017. Datasets on basic demographic information, health care usage including health screening records, and survey of health behaviors are used to answer the research question. Compared to another Korean health screening dataset based on administrative data, my dataset has several appealing features of survey datasets that I take advantage of in this paper.[4] First, the survey data include health behavior information when individuals do not participate in screening. This provides counterfactuals and enables comparing behaviors between years with and without screening. Second, it provides information that cannot be captured in administrative data, such as the purpose of hospital visits. This enables checking if there is an increase in hospital visits for a new illness that one was previously not aware of, which concerns

[4]The administrative dataset is much larger than the survey dataset and was used by Kim et al. (2019) and Kim and Lee (2017).

the primary objective of health screening. Lastly, the survey data contain information on the relationship between household members. It is used to check if there exist spillover effects in screening take-up between spouses.

The analytical sample consists of those who belong to the target population of free biannual screening, which is determined by age and the type of job. Detailed demographic information is used to identify the target population.[5] The baseline analysis shows providing free screening significantly raises the screening rate from 10 to 30 percent. The increment is stable over the years and is robust to the inclusion of various covariates and individual fixed effects. To claim that the IV plausibly satisfies exclusion restriction, individual and household characteristics are compared between those who are entitled to free screening (treatment group) and those who are not (control group). I also show that IV does not affect the screening rate among non-target population of the free screening policy. In the complier analysis section, I investigate which demographic group benefits from the policy and show that female, non-working, old and married people display larger increase in screening participation thanks to the free provision of biannual screening. Furthermore, people with low education and low household income have lower base rate without any incentive but exhibit much larger increase in screening participation in response to public screening program. This suggest the National Health Screening Program is successfully inducing screening participation among disadvantaged people.

Comparison of outpatient care usage during a calendar year between groups with and without free screening suggests screening leads to higher demand for more specialized medical services and more outpatient hospital visits and medical expenditures for a new illness. Screening induces people to substitute easily accessible outpatient care services at local clinics with more specialized services at large hospitals. I estimate the annual number of hospital visits to local clinics and medical expenditures incurred there decrease by 7.6% and 11.8%, while large general hospitals experience 9.6% and 8.3% increase, respectively. Using information on the purpose of visit that allows me to distinguish first

[5] Details on the target and non-target population is provided in section 2. As a robustness check on the limitation of job characteristic variable, analysis with adjusted sample is provided in the Appendix and the results are highly robust.

outpatient visits for new illnesses, I document large increase in the number of hospital visits and medical expenditures due to new illness across different types of hospitals. The number of aggregate hospital visits for a new illness increases by 9.4% in response to screening. This proves screening fulfills its primary objective and induces the discovery of new health conditions one is not previously aware of.

The screening has negligible effect on inpatient care, but significantly decreases emergency room (ER) visits. All the estimated coefficients on the impact of screening on the number of hospital visits for inpatient care and medical expenditures are small and insignificant. On the other hand, the number of ER visits, especially in general hospitals, significantly decreases due to screening.

Next, the event study analyses of health care utilization reveal three patterns emerge consistently in outpatient and inpatient care. First, there is a decreasing trend in the demand for outpatient and inpatient care prior to screening. Then, the visits to hospitals and medical expenditures increase substantially at the month of screening. They subside gradually in the following months to return to their original level. Note that these visits and expenditures exclude the ones incurred from screening. This suggests there is an intertemporal reallocation of health care usage. Before upcoming health screening, people delay visiting hospitals until after getting screening results. This pattern is more pronounced for inpatient care usage, plausibly because inpatient care treats more serious health conditions than outpatient care. However, emergency care does not exhibit such pattern. It shows an erratic pattern before and after the screening date, implying it is not subject to intertemporal reallocation.

This study also documents positive modification in health behaviors in response to screening. Imposing mild assumptions on the structure of behavioral responses to health screening, I show smoking, drinking, and exercising behaviors change in directions perceived to be a healthy lifestyle. Smoking mildly decreases on the extensive margin, in frequency and amount, but the estimates are loosely estimated. On the other hand, drinking shows clear and tightly estimated responses. There is a pronounced decrease in drinking and in binge drinking, and the effect mainly stems from decrease in drinking frequency rather than drinking amount. I also show the distributional effect that drinking frequency declines more among frequent drinkers. However, on the extensive margin, the share of current drinkers increases by 2.5 percentage point reflecting the fact that moderate alcohol consumption has health benefits. Using detailed question on drinking frequency, I pinpoint that screening-induced new drinkers drink no more than 4 to 5 times a year. Lastly, for exercise, there is a significant increase in walking, but exercise of higher intensity shows negligible response. This may be attributable to large measurement error in exercise variables and pre-screening recommendations of no intensive exercise that limit the scope of behavioral change.

The household level survey design makes it possible to map out household structure and identify the relationships between members. Using this information and the composition of free screening eligibilities, this study documents positive spillover in screening take-up between spouses. In addition to 20 percentage point increase in screening in response to one's own eligibility for free screening, spouse's screening additionally increases take-up rate by 8 percentage point. This arises irrespective of whether husband and wife are entitled to screening in the same or different years.

This paper makes contribution to 3 strands of literature. First, it adds to the literature on health screening. While screening is highly encouraged from the public health perspective, the participation rate is low. Kim and Lee (2017) finds similar rate using much larger administrative data of Korean NHIS data, and Jones et al. (2019) also documents small participation rate in the workplace wellness programs that include health screening. This study shows economic incentive in the form of screening fee waiver is successful in increasing the take-up significantly. This is also true of cost-sharing in cancer screening (Kim and Lee (2017)), and the demand to learn HIV status after testing in Malawi (Thornton (2005)). Despite large increase in screening take-up, the incentive is still not enough to cover the majority of population. In my study sample, 70 percentage of people still do not get screened even when screening is provided for free. This implies there are other binding constraints other than monetary concerns. Oster et al. (2013) provides an explanation where individuals advertently and rationally do not choose to get screening to avoid anticipatory disutility from the diagnosis of an illness.[6]

This paper complements the studies using artificial thresholds in health indicators to examine the impact of marginal information and treatment. Following the work of Almond et al. (2010), Kim et al. (2019) and Iizuka et al. (2021) focus on arbitrary thresholds in defining health conditions, such as 125 mg/dL level of fasting blood sugar for the diagnosis of diabetes, and use regression discontinuity design (RDD) to estimate the effect of marginal information on future health outcome, health care utilization and behavior. While both studies find limited behavioral responses, my studies document significant effects that may be attributable to two factors. First, I examine the effect of getting health screening per se, while the RDD studies are conditional on getting screening. If much of the effect on health care and behavior come from the act of getting screened, the RDD analysis might lack power to detect the effects. Additionally, while this study also does not find significant effect on aggregate outpatient care, examining heterogeneity by the size of medical institutions reveals substitution effect that cancels each other.

Second, it adds another evidence of behavioral responses to screening or diagnoses of certain diseases. There are many RCTs that examine behavioral response to various screenings.[7] (Wood et al. (1994), Group (1995), Larsen et al. (2007), Strychar et al. (1998), Jones et al. (2019)) They normally find limited change in behavior, such as smoking, drinking, diet, and gym usage. Unlike these studies that generally focus on specific group people in a controlled environment, this paper studies the behavioral effect of screening in a real setting where public screening policy is implemented en masse to the general population. Furthermore, this study documents significant behavioral changes that provide evidence that screening induces people to act in health-conscious ways.

Some studies focus on the behavioral response that follows the diagnosis of certain diseases. (Oster (2015), Oster (2012), Thornton (2005), Slade (2012)) Oster (2015) uses scanner data to examine the change in diet after the purchase of glucose testing product,

[6] This is more relevant if a disease involved is untreatable like Huntington disease in Oster et al. (2013). General health screening is more focused on finding any health abnormalities that are likely to be treatable.

[7] Deutekom et al. (2011) provides a nice review of clinical studies on behavioral response to screening.

and find significant, but limited reduction in calorie purchased. Oster (2012) uses IV strategy to find significant reduction in risky sexual behavior that differs by life expectancy in response to HIV in Africa. Thornton (2005) documents significant but limited change in condom purchase after testing positive for HIV in Malawi. Contracting diabetes or HIV is an endogenous behavior, and hence, the analysis of behavioral response requires a source of exogenous variation. Unlike these studies, this paper examines the behavioral response to testing for diseases, not necessarily the diagnosis of a certain illness, and uses the random variation in free screening provision to claim causal effect on health behavior.

Lastly, it contributes to the study of risky health behavior in general, and to the study of peer effect in particular. (Grossman (1972), Becker and Murphy (1988), Thaler and Shefrin (1981), Cawley and Ruhm (2011), Kenkel and Sindelar (2011)) It relates to literature that examines the relationship between different types of health behavior. Health behavior can be understood as investments and the determinant of this investment decision will drive correlation between behaviors. (Grossman (1972)) Education, income, macroeconomic cycle and other factors can play such a role. (Kenkel (1991), Cutler and Lleras-Muney (2010), Ruhm (2000)) My study shows that screening can also contribute to the short run comovement of health behaviors by inducing people to act in healthy ways.

People rarely make decisions in isolation, but with each other, and there can be many ways through which one's behavior can be affected by peers. (Leibenstein (1950), Manski (1993), Manski (2000)) Using exogenous variation in peers, many empirical studies document spillover effects in health behavior, 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 that screening behavior is also subject to peer effect that can be used as a tool to maximize take-up.

The rest of the paper is organized as follows. Section 2 introduces the institutional detail of Korean health screening policy along with IV. Section 3 explains the dataset, the construction of variables and the analytical sample. Next, section 4 lays out two-stage least square regression specifications used in the analyses. Assumptions of the IV and main empirical findings are discussed in section 5. Lastly, section 6 concludes. The Appendix includes supplementary regression tables and figures.

2 Korean health screening system

The National health screening program in Korea is composed of 3 main programs: general health screening, cancer screening and infant/children health screening. General health screening is the backbone of the screening program that examines the most 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 screening participants, those who are diagnosed to be at the high risk of hypertension, diabetes or cognitive dysfunction are asked to participate in additional tests with consultation. This study focuses on this general health screening to measure the responses to screening.

During the period of this study, the National Health Insurance Service (NHIS) provided fully covered biannual general health screening to people who satisfy one of the 3 criteria: (i) those 40 years old or above or (ii) all formal workers or (iii) heads of households. We will refer to this group as target population. See Figure 2. Among the target population, non-office workers are entitled to annual free screening, while the rest are entitled to biannual free screening. The analytical sample used in this study is those who are entitled to biannual free screening, hence (B) in Figure 2.[8]

The eligibility for free screening among biannual screening group, (B), is determined by the year of birth. Those born in odd-numbered years get free screening on odd-numbered years and those born in even-numbered years get free screening on even-numbered years. Participating in an off year when one is not eligible has to be paid by out-of-pocket cost. This eligibility condition based on the year of birth roughly divides the target population into half of eligible and ineligible group for each year and

[8] I use the job characteristics to distinguish office and non-office workers in the dataset following guidelines provided by the NHIS. However, this might not be a perfect way to distinguish the two. Hence, a robustness check of using the entire target population, (A) + (B), as an analytical sample is provided in the Appendix.

they will be referred to as treatment and control group. Note that treatment and control group assignment is switched every year and this is why this study focuses on short term behavioral response to screening.

Besides medical tests, the screening program provides health risk evaluation based on the questionnaire filled out by the examinee regarding medical history, smoking, drinking, and exercising behavior. Examinees are asked to provide the detailed frequency of these activities and based on the answers, have a short interview with the doctor where they are recommended to change their lifestyle or, in some serious cases, to join a stop-smoking or alcohol abstinence program provided by the NHIS. After the screening, the medical institution that carried out tests is legally obligated to notify the checkup results to the examinee in fifteen days. In addition to detailed medical test results, a checkup report includes cardiovascular disease risk assessment section where current and target health behaviors are provided. An example of the form is provided in Figure 1. Combining all the health-related factors, the estimated probability of developing cardiovascular disease within 10 years is also provided to alert the danger of developing a disease and to induce positive behavioral modification.

In addition to general health screening, the NHIS also provides 5 cancer screening programs: stomach, liver, colorectal, breast and cervical cancer, covering 90% of the cost. Each cancer screening has its own eligibility conditions and recommended frequency. However, many of them follow the same schedule as general screening and are usually carried out together biannually on the same day on top of general screening. This is even true of cancer screenings provided annually or screenings that are not covered by the NHIS. Empirical evidence suggests people receive many types of cancer screenings biannually even though the NHIS subsidizes them annually or not at all. The eligibility conditions and the first stage regression results for different types of cancer screenings are summarized in Table A1. For this reason, while this study specifically uses eligibility condition for general health screening and examine its behavioral impacts, it should be understood as the impact of health screening broadly encompassing both the general and cancer screening.[9] For robustness check, the analysis using cancer screening instead of general screening is also provided in the Appendix and the results are highly robust.

3 Data

This study utilizes the collection of Korean Health Panel Study datasets spanning the years 2009-2017.[10] It is a collection of yearly panel datasets with nationally representative sample that started with about 6,000 households and 19,100 individuals in 2009. With gradual attrition in the sample, to guarantee national representativeness and statistical precision of analyses, the second cohort of 1,800 households and 5,000 individuals was added to the sample in 2014. The size and the attrition of two survey cohorts are shown in Table 1.

The collection of datasets include vast information on basic demographic and socioeconomic characteristics, health care usage, health-related lifestyle and so forth. I use the datasets on outpatient care, inpatient care, and emergency care to construct variables on health care usage. Detailed records on health screening also come from the outpatient care dataset. Health-related lifestyle dataset contains information on smoking, drinking and exercise. The demographic and socioeconomic information is used to construct various control variables and the instrumental variable that depends on the year of birth. Data collection was done by 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.

Three types of health care are considered, outpatient care, inpatient care, and emergency care. For each type, I examine the number of visits to hospitals, hospital bills and drug expenditures.[11] Understanding the unit of observation is a key to the study design.

[9] One might suggest separating the effect of general screening and cancer screening by including them simultaneously in one regression. While theoretically possible, this is empirically difficult due to high correlation (0.83).

[10] It is version 1.6 made jointly by Korean Institute for Health and Social Affairs (KIHSA) and National Health Insurance Services (NHIS). While the dataset starts at 2008, since it lacks health behavior information in this year, we use the sample from 2009 onwards.

[11] Hospital bills and drug expenditures are amount net of insurance payment. The national health insurance payment is deducted automatically from the total bill at the moment of payment. Hence, these are the remaining amount that consumers end up paying.

Outpatient, inpatient, and emergency care datasets consist of every visit to a hospital. The unit of observations is each visit and this allows constructing outcome variables on health care usage at the yearly, monthly or any desired frequency. For every visit, it also includes information on the type of medical institution visited, purpose of visit, hospital bills incurred, and drug expenditures incurred if one is prescribed any. Hence, I consider the aforementioned outcome variables differentiated by 3 types of hospital sizes.[12] For outpatient care, there is information on whether the visit was the first-time visit for a new illness as opposed to subsequent or recurring regular visits. To test the hypothesis that screening may lead to more discovery of new illness and subsequent increase in hospital visits for a new illness, the first visit variables in outpatient care will be separately considered as outcome variables.

The health screening records also come from the outpatient care dataset with information on the type of screening, the medical tests performed, whether one found any disease, and the exact date of screening. The type of screening refers to distinction between general screening and various cancer screenings. This study uses general screening as baseline analysis and provides robustness check using cancer screening in the Appendix.[13] The detailed date of health care usage and health screening allows an event study analysis of examining the temporal change in health care usage in response to screening which I turn to in section 5.2 and 5.3.[14]

Unlike health care usage, health behavior variables are yearly observations. They are answers to questions on smoking, drinking, and exercising behavior in the recent past relative to the survey date. Hence, while both the screening and survey dates are known, it is not possible to examine the temporal change in health behavior in response to health

[12] A hospital is classified into 3 types by size approximated by the number of beds to accommodate inpatients. First, local clinics are the smallest with less than 30 beds. Second, a local hospital has 30 to 100 beds. Lastly, a general hospital has more than 100 beds.

[13] For reasons why cancer screenings are considered separately, see section 2.

[14] The health care usage information on a case-by-case basis is collected by survey participants and enumerators going through every visit to medical institutions from the last date of survey, which is approximately a year ago. Participants in the Korean Health Panel Study are required to keep detailed record of their health care usage by using a specifically designed health diary and storing receipts from every visit to hospitals and pharmacies. The enumerators collect the health diary and all the receipts at annual visit, compare each entry in the diary with receipts, and record their data.

screening due to infrequent observations.[15] For smoking and drinking, respondents were asked the frequency of smoking and drinking in the past one month before the survey. Conditional on smoking and drinking, they were also asked the number of cigarettes smoked and the cups of alcohol consumed on the days of smoking and drinking.[16] Exercise was categorized into 3 different types depending on the intensity; vigorous exercise, moderate exercise, and walking. For each type, respondents were similarly asked the frequency and amount of the activities in minutes in the last one week. Since the health behavior questions were asked in intervals, rather than in continuous numbers, threshold crossing model is used where dummy outcome variables are used to indicate a frequency or amount exceeding a certain threshold.

As briefly explained in section 2, the analytical sample is the target population for biannual general health screening. See Figure 2 for ease of understanding. One of the three conditions should be satisfied to be the target population: (i) age 40 years or above, or (ii) formally employed, or (iii) head of household. Among this group, non-office workers (A) get annual free health screening, and the rest (B) gets biannual free health screening.[17] Therefore, my analytical sample is (B), the ones who satisfy one of the three conditions, dropping non-office workers. This subset of population is entitled to NHIS-provided free health screening once every two years. This effectively leaves me with 73,535 individual-year pair observations during 9 years.

It is worth comparing the current dataset to another frequently used data source on Korean health screening program. Kim and Lee (2017) and Kim et al. (2019) uses administrative data of Korean health screening policy to examine behavioral responses to disease classification and stomach/breast cancer screening by taking advantage of artificial cutoff in health indicators and income. While administrative data are much larger and include similar information, my dataset has appealing features specific to

[15] The discrepancy between the screening date and the survey date has implications for the interpretation of the coefficients of causal impact. I turn to this issue in section 5.4.

[16] Consumption of different types of alcohol were considered by using different size cups for different alcohols. Smoking behavior was not differentiated by smoking materials.

[17] The distinction between office and non-office workers is ambiguous. A variable on job description is used to distinguish them. The robustness check including non-office workers are presented in the Appendix.

survey datasets that I take advantage of. First, my survey data contain yearly records of health behavior regardless of screening take-up, while the administrative data only have behavior information conditional on getting screened. Since health behavior information exists as answers to health risk assessment questions during screening, the administrative data do not contain behavior information for those who are not screened. The health behavior for people that are not screened provides an important counterfactual, and therefore administrative data cannot be used for answering the current research question. Second, the survey dataset offers subjective information that cannot be measured in administrative data. One example is whether a visit to a hospital is for a new illness. Checking whether screening leads to an increase in the diagnoses of new illnesses that people were unaware of and subsequent increase in hospital visit for a new illness is directly linked to the fundamental goal of health screening. The information on the purpose of hospital visit provided in the survey data makes it possible to answer this question. Lastly, while both datasets contain individual level information, the survey dataset is based on household level sampling, and it includes information on all the household members. This allows identifying relationship among household members, which administrative data do not allow. I take advantage of this information to examine spillover effect in screening take-up and health behavior in section 5.6.

4 Empirical strategy

I perform two types of baseline analysis. First, using eligibility for free biannual screening as an instrument, I investigate the effect of screening uptake on health care usage and health behaviors. Given that free screening eligibility is determined at yearly basis, this analysis also uses outcome variables defined at yearly level. Next, I take advantage of the fact that health care usage data have more frequent observations and perform event study analysis in which I examine the temporal variation in health care usage around the screening date at monthly level. Note that event study analysis is not possible for health behavior outcome variables due to infrequent observations.

The specification for IV analysis is given in the following equations.

$$Screening_{it} = \alpha_0 + \alpha_1 Eligible_{it} + \mathbf{X_{it}} + \eta_{it} \tag{1}$$

$$y_{it} = \beta_0 + \beta_1 Screening_{it} + \mathbf{X_{it}} + \varepsilon_{it} \tag{2}$$

First stage regression given in Equation 1 examines the effect of IV, eligibility for free health screening, on the screening participation for individual i in year t. Tightly estimated positive coefficient α_1 signifies a strong instrumental variable that meets relevance condition. The second stage regression shown in Equation 2 uses two-stage least square regression to explain variations in the outcome variable, y_{it}, with the predicted value of screening from the first stage regression. The coefficient of interest is β_1 and it captures the local average treatment effect (LATE) of health screening on the outcome variable.

$\mathbf{X_{it}}$ is a vector of control variables. The preferred specification used in baseline analyses in section 5 do not use any control variable. Hence, $\mathbf{X_{it}} = 0$. For robustness, I try two other alternative specifications and present the estimation results in the Appendix. First, I include various demographic and socioeconomic variables as covariates that include age, gender, marital status, years of schooling, income decile, insurance type, handicapped status, working status, odd year of birth, month of survey and year fixed effects. Second, taking advantage of the panel structure of the dataset, I include individual and year fixed effects as covariates in which case $\mathbf{X_{it}} = \gamma_i + \delta_t$. The results are robust to the inclusion of various control variables. η_{it} and ε_{it} are error terms and the standard errors are clustered at the individual level.

In each domain of health care or health behavior, there are multiple outcome variables. Multiple hypotheses testing raises the chances of making Type I error beyond the desired size of the test. Hence, this study uses Westfall-Young stepdown adjusted p-values, following the procedures of Westfall and Young (1993) developed in Jones et al. (2019), to control the familywise error rate, the probability of committing any type I error. For this study, I use 1,000 bootstrap replications. Furthermore, following Kling et al. (2007), an index variable is calculated for each domain of behavioral outcome variables.

In the section on heterogeneity, the same regression specification shown in Equation 1 and 2 will be estimated for different demographic groups by splitting the sample. The subsample analysis with the first stage equation constitutes complier analysis where I examine which group participated in screening in response to free screening that otherwise would not. Following Angrist and Pischke (2008), the relative likelihood that a complier belongs to a demographic group can be learned from comparing the first stage coefficient for the group to that of the entire sample. I also estimate the second stage IV regression for separate groups to examine heterogeneity in behavioral response.

Next, I employ event study framework to investigate the evolution of health care usage around the screening date at monthly level. Note that this is possible only for outpatient, inpatient and emergency care datasets thanks to frequent observations. The specification is given as follows.

$$y_{imt} = \gamma_i + \delta_{mt} + \sum_{\substack{k=-13 \\ k\neq-1}}^{12} \tau_k D_{imt}^k + \mu_{imt} \tag{3}$$

The outcome variable y_{imt} is health care usage variable for individual i in month m, year t. Note that the time dimension is switched to monthly level. The γ_i refers to individual fixed effects and δ_{mt} represents month-year fixed effects. The event time variable D_{imt}^k is defined as follows.

$$\begin{aligned} D_{imt}^{-13} &= \mathbf{1}[mt \leqslant e_i - 13] \\ D_{imt}^{12} &= \mathbf{1}[mt \geqslant e_i + 12] \\ D_{imt}^{k} &= \mathbf{1}[mt = e_i + k] \quad for \quad -12 \leqslant k \leqslant 11 \end{aligned}$$

where e_i is the date of screening for individual i

This is a standard event study design with 12 months of pre-periods and post-periods. Dummies for the time period -13 and 12 capture the time outside the event window. One month before the screening date is used as the reference category.

One issue in using event study design is that individuals can get screening multiple times during the study period. This creates an issue in defining the event time variable, D^k_{imt}. To deal with the multiple events for an individual, I use the stacked dataset approach. After making a balanced panel dataset at individual-month-year level, observations for each individual are copied by the number of screenings during the study period. The event time variable is coded such that each copy of an individual's observations treats different screening as the event. This allows avoiding the issue of multiple treatments. The never-treated, or those who never participated in screening, do not affect the estimates and, hence, are dropped from the sample. μ_{imt} represents the error term, and the standard errors are clustered at the individual level.

5 Results

5.1 Validity of IV

This study uses instrumental variable (IV) to estimate the causal effect of health screening on health care utilization and health behaviors. The instrument is the eligibility for free health screening provided every other year. The eligibility criterion is based on even-odd design; those born on odd (even)-numbered years get free health screening on odd (even)-numbered years.[18] A valid IV should satisfy two criteria. First, it should have a strong effect on the explanatory variable, which is evidenced in strong first stage regression. Next, it should satisfy exclusion restriction, that is, it should be plausibly exogenous and the only channel through which it affects outcome variable should be the explanatory variable.

First, I provide evidence that the instrument has a strong impact on screening take-up. Figure 3a shows the average screening rate for the two groups, those born in odd and even years. One can clearly see an alternating pattern in screening rate. Those born in odd-numbered years have higher screening rate on odd years and lower rate on even years, and the pattern is reversed for those born in even-numbered years. The baseline screening rate when there is no incentive is about 10 percent. Given free screening, the take-up increases by slightly less than 20 percentage point and the increment is stable over

[18]This is equivalent to one's age being even-numbered, since the difference between even numbers or odd numbers is an even number.

the years. Regression results shown in Table 3 confirms the estimate. Column 1 shows the effect of eligibility for free screening on take-up, and it increases by 18 percentage point. The coefficient is positive and tightly estimated. As I add demographic and socioeconomic covariates and year fixed effects in column 2, the estimate hardly changes. As shown in column 3, the estimate is also robust to adding individual fixed effects in place of covariates. This suggests the IV is not highly correlated with individual or household characteristics, as desired.

Next, to argue randomness of the instrument, I check the balance on covariates between treatment and control group. In a given year, the IV roughly divides the population into two equally sized group of those who are eligible for free screening (treatment group) and those who are not (control group). Table 2 shows comparison of the two groups on various individual and household characteristics. They are well balanced on individual and household covariates and the F-test of joint significance fails to reject the null hypothesis of no difference.[19] This provides evidence that the IV is plausibly random.

It still remains to show that the IV does not affect the health behavior through other channels. To the best of my knowledge, there is no any other policy that relies on the even-odd design of year of birth. Furthermore, I can also conduct a falsification test using a sample that is not a target population for screening policy. This corresponds to group (C) in the Figure 2. The screening rate should not respond to even-odd design if one is not entitled to free health screening. The Figure 3b plots the screening rate for the non-target population. It shows that there is no alternating take-up pattern for both groups as previously seen in Figure 3a.[20] This boosts confidence that our even-odd

[19] There was a marginally significant difference in age. Treatment group is slightly younger. Given that young people are more likely to smoke, drink and exercise on average, it will bias our estimates upwards. This results in more conservative estimates for the coefficients on smoking and drinking behavior, and less conservative coefficients for exercise.

[20] There can still be ways through which those not subject to general screening get screened following even-odd design. First, they can be subject to cancer screenings that follow even-odd design, such as breast and cervical cancer screenings. See Table A1 for detail. Second, getting health screening once every two years is a common rule that many people follow, and to the extent that it is understood as socially recommended screening schedule, people can follow it without any economic incentive. Third, if there is spillover effect in screening take-up, then non-target population can also show an alternating screening pattern. In section 5.6, I show this is indeed the case between spouses. The first stage regression coefficient for the non-target population is also positive and statistically significant, but it is much smaller compared to that of the target population. (coef: 0.016, p-value < 0.01)

design has an impact, if any, through screening take-up.[21]

5.2 Outpatient care

5.2.1 IV analysis

This section presents the baseline result of the effect of screening on outpatient care usage using the specification outlined in Equation 2. The screening participation is instrumented by the eligibility for free biannual screening to exploit the random variation in screening uptake, and I do not include any control variables.[22] Outcome variables are the total number of hospital visits, hospital bills and drug expenditures for outpatient care incurred in a calendar year. They are further decomposed into 3 levels by the size of the medical institution visited.[23] Furthermore, I specifically examine the first visit to a medical institution for an illness based on the answer to a question asking if the hospital visit was either the first visit for a new illness, second visit or regular recurring visit. This sheds light on the effectiveness of health screening in achieving its fundamental goal of leading to a discovery and subsequent treatment of illnesses.

Before diving into main regression result, it is worth examining why naïve regression of health care usage on screening is misleading and how IV can remedy the problem. The naïve regression results presented in Table **??**show that screening is positively associated with health care usage and medical expenditures. This could be due to omitted variables, such as underlying health conditions that require more medical attention. There is a

[21]Non-office workers who are subject to annual screening, in principle, should show the same behavior as those not subject to screening policy. However, they also show a clear alternating screening pattern very similar to those subject to biannual screening. (First stage coefficient: 0.15, p-value < 0.01) This could be because the job characteristics variable used to define non-office workers is not precise enough. Another reason stems from administrative processes to be registered as non-office workers. There are many mismatches between one's actual job type and the one in the NHIS database. The default setting is that an employee is categorized as office worker and is subject to biannual screening until the employer goes through administrative processes in the NHIS to register them as non-office workers. This additional process and the default choice result in many non-office workers categorized as office workers, hence being entitled to biannual, not annual, free screening. For robustness, I enlarge the analytical sample to include the non-office workers and perform the same analysis. Main findings are robust to this sample adjustment and are shown in the Appendix.

[22]The estimation results including control variables or individual fixed effects are presented in the Appendix. The results are robust to the inclusion of various controls.

[23]The number of beds to accommodate inpatients is used as a proxy for the size of medical institutions. General hospitals have the largest capacity with more than 100 beds. Local hospitals can accommodate 30-100 patients. Local clinics are smallest with 0-30 beds.

negative selection in terms of health conditions and this biases the coefficients upwards. Including controls slightly diminishes the coefficients, but is not enough to drive down the selection effect. Hence, we use the random eligibility for free screening to estimate the causal effect of screening on outpatient care usage and expenditures.

The reduced form figures that demonstrate the effect of free screening eligibility on the number of visits and first visits to hospitals are shown in Figure 3c, 3d. Figure 3c plots the detrended number of visits to hospitals for those born in odd and even years. Since the dataset is panel data, it is the same people who are included in different years. As years pass by, they age, and the number of visits to hospital shows increasing trend. I center each year's observations by subtracting the yearly mean values to accentuate the difference between those born in even and odd years./footnoteThe reduced form regression compares the average outcome values between those born in even versus odd years. Subtracting the same number from the two average values does not affect the difference. As shown in the first stage regression in Figure **??**, we expect to see the opposite alternating pattern between the two groups if there is any effect of screening. In Figure 3c, there is an alternating pattern from year 2013 with those born in odd year displaying lower value in odd year, but the pattern is less clear in years prior to 2013. However, the pattern is much clearer in Figure 3d where first visits to hospitals are plotted after detrending. Those born in odd years exhibit larger values in odd years and lower values in even years, and the pattern is reversed for those born in even years. Combined with first stage, this implies health screening leads to increase in first visits to hospitals.

The baseline results on outpatient care usage and medical expenditures using IV is presented in Table 4. Panel A contains the results for the total number of visits to hospitals, hospital bills and drug expenditures, and panel B contains the corresponding outcomes for the first visit. The first two columns present the means in the control and treatment group. The third column presents the intention-to-treat (ITT) effect, which is the difference in the means of the two groups. The fourth column presents the local average treatment effect (LATE), that is, ITT scaled by the first-stage coefficient presented in Table 3. This summarizes the causal effect of screening on outcome variables.

Screening leads to an overall decrease in the number of outpatient hospital visits, hospital bills and drug expenditures. The number of visits decreases by 0.847 (4.5%), hospital expenditures decrease by 21,553 KRW (6.3%), and drug expenditures drop by 2,945 KRW (2.6%), but all the coefficients are loosely estimated and not significant at conventional level.[24] Note that these are short-term changes in response to recurring health screenings and they do not say anything about long-term effects.

Decomposing each outcome variable by the size of medical institutions reveals substitution effect in demand in favor of more advanced medical services. Visits to big hospitals, like general hospital, increases by 0.238 (9.6%) in response to screening, and so do hospital bills by 8,971 KRW (10.7%) and drug expenditures by 1,383 KRW (3.4%).[25] In contrast, visits to small local clinics decrease significantly by 1.157 (7.7%). Hospital and drug expenditures drop by 28,715 KRW (13.3%) and 3,896 KRW (6.3%), respectively. The demand for mid-sized local hospitals is in between. This implies there is a substitution effect from visiting local clinics to general hospitals. General hospital provides wide range of medical services that are also more specialized than local clinics. Screening induces short-term increase in demand for higher level services at the expense of more easily accessible local medical services.

Panel B presents the same analysis with the outcomes replaced by first visits, first hospital bills and first drug expenditures for a new illness. One goal of screening is to help people discover unknown illness, if any, at an early stage, and increase in visits to hospital for a new illness is evidence that this works. The empirical evidence supports the hypothesis. The total number of visits to hospitals, hospital bills and drug expenditures all rise significantly in response to screening. The estimates are also economically significant. Screening leads to 0.348 (9.4%) more visits to a hospital, and hospital bills and drug expenditures also increase by 24,238 KRW (27%) and 2,048 KRW (18.6%), respectively. Decomposition by hospital size confirms the substitution of local clinics with general hospitals. The comparison of local average treatment among different types of

[24]The currency is Korean Won (KRW), and the average exchange rate in 2015 was approximately 1 USD = 1130 KRW.

[25]Table A5 and A6 presenting the results using specifications with control variables show statistically significant coefficients for three outcome variables pertaining to general hospital.

hospitals shows smallest point estimates for local hospital across 3 outcomes. However, column 1 shows the base usage is lower for local hospitals. Converting the point estimates to percentage, there are 28.3%, 13.3% and 7.1% increase in visits to general hospital, local hospital and local clinic, respectively. This shows that screening creates a higher demand for more specialized medical services.

5.2.2 Event study

The previous analysis using IV establishes the causal effect of recurring health screening on outpatient care usage at yearly level. On top of this, the high frequency outpatient dataset makes it possible to examine more granular variation in time. Using the event study specification presented in Equation 3, I investigate the temporal variation in outpatient care utilization around the screening date. Combined with the IV regression results, this paints more complete picture of the impact of screening on demand for outpatient care.

Among the total 24 outcome variables, I focus on 4 outcomes: total and first visits for general hospitals and local clinics.[26] The event study figures with 12 pre- and post-periods are presented in Figure 4. There are 3 common patterns present in all 4 figures. First, there is a downward trend in the number of visits before screening. Next, all outcome variables sharply increase at the month of screening. Lastly, they gradually decline after screening and revert back to their original levels. The difference between the aggregate visits and the number of first visits seems minimal, except that first visit variables drop more quickly after screening to return to pre-screening level in about 2 3 months, while the aggregate variables take 4 5 months of decline. Comparison between general hospital and local clinic also shows little difference. One interesting pattern is that while visit to both institutions shows downward trend in pre-periods, visit to local clinics displays slight rebound just before the screening, unlike in general hospital.

Overall, all the figures display similar pattern of decline in visits before screening, large

[26] This is because the number of visits is highly correlated with both measures of medical expenditures, and they all move in the same direction. Furthermore, visits to local hospital always lie in between the other two types of hospitals. Hence, the comparison between general hospital and local clinic makes the distinction between different types of hospitals clear.

jump at the month of screening, and gradual decline in post-periods. This pattern is also present in medical expenditures whose event study figures are presented in the Appendix. The results suggest temporal allocation of health care usage. Expecting upcoming health screening, people show tendency to delay their hospital visits to after the screening, and this shows up for both the general hospital and local clinic. After the screening, with the screening results and evaluation of their health conditions in hand, they visit hospitals to get further consultation with the doctor and if necessary, treatment for their health conditions. This shows up as a large increase in the number of visits to hospitals and medical expenditures. It is followed by gradual decline in 4 to 5 months. All the figures make it clear that there is an anticipation for screening which reduces health care usage in pre-periods, and it induces temporary sharp increase in hospital visits and medical expenditures afterwards.

The increase in visit to local clinics at the month of screening shown in the event study figure 4b does not contradict the overall decline in the total number of visits to clinics as shown in IV regression table 4. The two are in essence measurement of different concepts. The IV analysis produces local average treatment effect by comparing the treatment and control group, and the difference in means stems from compliers in the treatment group that were induced to participate in screening due to fee waiver. This produces a causal estimate of screening. Event study regression, on the other hand, only uses those who at least participated once in screening during the study period. They are always-takers and compliers, and do not include never-takers. The coefficients only show temporal change in outcome variables across time using one month before the screening as the reference category. Combining the 2 pieces of evidence, screening leads to an overall reduction in clinic visit, but at the month of screening and following 1 to 2 months, it shows higher clinic visits than one month prior to screening.

5.3 Inpatient and emergency care

5.3.1 IV analysis

Using the same empirical strategy as in outpatient care analysis, I investigate the impact of screening on inpatient and emergency care usage. The rich data on inpatient and emergency care at the visit level enables IV analysis with focus on the causal impact of screening at the yearly level and the event study analysis examining the variation across time dimension. The outcome variables are similarly defined as in outpatient care analysis with the number of visits to hospitals, hospital bills and drug expenditures disaggregated by the type of medical institutions. The timing of inpatient care is determined by the first day of hospitalization and I do not consider the duration of hospitalization. As for emergency care, the hospital bills do not include the expense for ambulance service, since it is provided for free.[27]

Table **??** Panel A shows the naïve regression of inpatient outcome variables on screening participation. Unlike in outpatient usage where screening is positively correlated, there is negative association between screening inpatient care use. Those who get screened are less likely to use inpatient care. Table 5 Panel A shows the ITT and LATE estimates of screening. Most of the coefficients are close to zero and the standard errors are large. I fail to reject that screening has significant effect on inpatient care usage.

Next, I examine the impact of screening on emergency care. Panel B in Table **??** also shows that there is negative correlation between screening and the use of emergency care. Panel B of Table 5 presents the causal impact of screening on emergency care using free screening eligibility as IV. Unlike in inpatient care use, there is a significant reduction of emergency care use in general hospital due to screening. Note that local clinics or small hospitals rarely have emergency care and most of them are provided by general hospitals and big local hospitals. So, no impact on local clinics is to be expected. The majority of emergency care services are provided by big general hospitals and we see significant reduction in the visit to emergency room in general hospitals.

[27] While there are private ambulance services that incur cost, they only apply to special circumstances. In my data, more than 95% of people who used emergency service did not incur any fee for ambulance service.

5.3.2 Event study

Given that there is no large difference in inpatient and emergency care response to screening between different types of hospitals in Table 5, I present the event study plots for aggregate measure of total inpatient and emergency care usage. Figure 4e plots the event study coefficients for the total visits to hospitals for inpatient care with 12 pre- and post-periods in months. Consistent with the figures from outpatient care, there is a strong decreasing trend in inpatient care before screening, but the slope is much steeper than in outpatient care. Then, there is also a sharp increase in inpatient care at the month of screening and then it levels off. This is an indication of a strong temporal allocation of inpatient care. People delay getting hospitalized until after the screening, as confirmed with outpatient care data.

Corresponding figure for emergency care is presented in Figure 4f. Unlike in previous figures, the estimates are overall noisy and it is hard to discern a pattern. There seems to be a slight downward trend in pre-periods and increase in the month of screening, but it is not clear. Even after the screening, there is no sign of leveling off. If anything, the use of emergency care seems to show an increasing trend after screening. This implies screening does not have large impact on temporal variation in the use of emergency care.

5.4 Health behavior

This section presents the effect of screening on smoking, drinking and exercising behavior. For each behavior domain, I investigate the impact of screening on both the extensive and intensive margin. The intensive margin is decomposed into the frequency of engaging in a behavior and the consumption amount of cigarette or alcohol or the duration of exercising. I use the eligibility for free screening to estimate the local average treatment effect on behavioral outcomes using the specification presented in Equation 2. For each behavior domain, index variable reports the standardized treatment effect following Kling et al. (2007). Considering inference with multiple hypotheses, Westfall and Young stepdown p-values are also reported accounting for the same number of hypotheses as the number of outcome variables in each domain. All the baseline specifications do not include any control variable. The estimation results using various controls and individual fixed effects are presented in the Appendix.

Given difference in screening and survey date, the interpretation of the causal estimates is more nuanced. The responses to lifestyle questions measure the behavior of respondents in the recent past of the survey date. Screening, on the other hand, can take place either before or after the survey date. To understand what we are measuring, I assume that there are both anticipatory and ex-post effects of screening, that is, people can change their behavior in anticipation of screening and as a result of screening. If survey is carried out before screening, the behavioral responses capture anticipatory effects and if survey is carried out after screening, they capture ex-post effects. I also assume that the magnitude of anticipatory and ex-post effect is inversely related to the duration between screening and survey dates. The closer the survey happens to the screening date, regardless of before or after, the magnitude of the anticipatory or ex-post effects that it captures is stronger. Hence, the absolute value of the date interval between screening and survey date works as a channel through which screening affects health behaviors measured in surveys, with smaller date interval meaning stronger effect.

Based on these assumptions, I check the empirical distribution of the date interval variable, formally defined as survey date minus screening date in days such that negative number indicates survey happening before screening.[28] The kernel density of date interval is plotted for the treatment and the control group separately in Figure 5a. Figure 5b plots the same distribution for date interval between -360 and 360. The distribution is rather symmetric around zero and the treatment group has much higher mass concentrated around zero. This is due to compliers. Free screening provided to the treatment group induces higher participation as shown in the first stage regression, and this results in many surveys having date intervals close to zero. This is the channel through which free screening provision affects health behaviors measured in the survey. By having more individuals surveyed near the screening date, treatment group captures stronger

[28]This variable is defined for each survey using the screening date that is closest to the survey date. They do not necessarily need to happen in the same calendar year. It cannot be defined for those who never participated in screening during the study period.

anticipatory and ex-post effect of screening than control group, and this shows up as different behavioral responses in the survey.[29]

The naïve regression presented in the Appendix Table A4 show smoking and drinking are generally negatively correlated with screening behavior, while exercise is positively correlated. These are mere correlations and may be driven by a factor that those more invested in health are more likely to get screened and exercise, and less likely to drink and smoke. There is evidence that mild drinking has health benefits. This nonlinear effect of drinking on health is what drives positive correlation between screening and drinker dummy. Furthermore, it is surprising to see that after controlling for various demographic and socioeconomic factors, most of the drinking behavior is positively correlated with decision to get screened. As shown in the specification with controls, the inclusion of various control variables is not enough to drive down these selection effect. Hence, this paper uses random variation in screening rate provided by different birth year to estimate the causal effect of screening on health behavior.

The reduced form plot for the detrended drinker and everyday drinker variables are presented in Figure 3e and 3f. If the IV induces change in health behaviors, we should expect to see opposite change in health behavior between the treatment and control group over the years, with positive health behavior occurring in the group that is entitled to free screening. Figure **??** clearly shows there is an alternating pattern of ups and downs for the rate of everyday drinker that corresponds with the change in screening rate shown in the first stage Figure 3a. The group eligible for free screening shows decline, whereas those ineligible show increase in the rate of everyday drinker. Figure **??** exhibits reversed pattern for the current drinker. This is due to the fact that mild drinking has health benefits and people respond by starting to drink in the year of screening.

Panel A in Table 6 shows that screening has an overall negative effect on smoking behavior, although the estimates are generally loosely estimated. The smoker indicator variable captures change in extensive margin, and screening reduces the probability of being a smoker by 1.4 percentage point, though the estimate was not significant at con-

[29]Note that our estimates measure the sum of anticipatory and ex-post effects, and it is not possible to disentangle them.

ventional level. Next 3 outcome variables relate to the frequency of smoking. Smoking days decrease by 4.7 days a year (6.9%) and the probability of smoking once a week or more, or smoking everyday decreases by 1.3 percentage points, each. While the estimates are all negative, they were insignificant. Lastly, the amount of smoking on a smoking day also declines by 0.12 cigarettes (4.3%), but it was not significant. The coefficients on the dummy for smoking 3 or10 cigarettes or more were significant at 10%, and the probability decreases by about 1.5 percentage point. The index variable summarizes the change in smoking behavior, and overall, screening reduces smoking behavior, but the effect was marginally insignificant.

The corresponding estimation results for drinking are presented in Panel B in Table 6. Surprisingly, there is a 2.5 percentage point increase in the probability of being a current drinker. This increase in extensive margin is consistent with the explanation that mild drinking has beneficial impact on health, and non-drinking people respond by taking up occasional drinking. However, the coefficient immediately turns negative for the outcome variable: drinking once a month or more. This implies those who start drinking in response to screening drink less than once a month, or 12 times a year. Otherwise, the LATE coefficient for the outcome variable drinking once a month or more should have been positive. To pinpoint how often the new drinkers drink, I run the same two stage least square regressions shown in Equation 2 with the outcome variables $\mathbf{1}\Big[\text{Drinking frequency per year} \geqslant j\Big]$ for $1 \leqslant j \leqslant 11$. Those who responded to be a drinker but drink less than once a month were further asked to specify how many times they drink a year, with the answers ranging from 1 to 11. The LATE estimate would be positive for small j, but would turn negative at some point as j increases to 11. The Figure 6 plots the 11 LATE coefficients with 95% confidence interval. The coefficient shows downward trend as expected, gets very close to 0 at 4 times a year and turns negative at 6 times a year. Hence, screening induces new occasional drinking at most up to 4, 5 times a year.

The 3 variables on drinking frequency show that the probability of drinking (i) once a month or more, (ii) once a week or more, and (iii) everyday, decreases by 1.7, 2.1 and 1.8 percentage points in response to screening, respectively. Given that the control group averages of drinking at given frequency are 50%, 29% and 6% in the given order, it implies screening decreases the frequency of drinking more among frequent drinkers. This represents 3.4%, 7.3%, and 31.6% decline in the probability of drinking (i) once a month or more, (ii) once a week or more, and (iii) everyday. Screening also induces reduction in binge drinking frequency as shown by negative coefficients. Once again, there is a larger reduction in frequency among frequent binge drinkers. There are 4.6%, 13%, and 40.9% decline in the probability of binge drinking (i) once a month or more, (ii) once a week or more, and (iii) everyday.[30] The effect on drinking amount is mixed and not significant. This implies screening reduces drinking more through the frequency channel, rather than the amount of drinking. The index of drinking behavior suggests significant behavioral modifications in response to screening.[31]

Lastly, Panel C in Table 6 presents the change in 3 types of exercising behaviors differentiated by intensity. There is evidence that walking increases on the extensive margin and in frequency in response to screening. The days of walking increase by 10.6 days a year (5.1%), but the amount of walking seems to slightly decline. The coefficients for more intensive exercises are imprecisely estimated and I cannot reject the null effect. The index variable summarizes the positive but insignificant effect of screening on exercising behavior. One reason the study lacks power to detect significant effects could be measurement error. Unlike in cigarette and alcohol consumption where purchase of which produces receipts and purchase histories that can be used as memory aid device, exercising hardly entails any memory aid device. Remembering how much one did different types of exercise in minutes could be much more challenging. This produces measurement error in the outcome variable, which lowers precision. Another reason could be health screening preparation guidelines that recommend avoiding intensive exercise before medical tests. This would have a direct effect of reducing vigorous or moderate exercise and potentially increasing walking, the mildest form of exercise, before screening.

[30]Binge drinking is an act of drinking large amount of alcohol in one sitting and, given nonlinear effect of drinking on health, it is more directly related to harmful drinking behavior. The exact definition of binge drinking used in this study is drinking more than 7 (6) cups of soju or 5 (4) cans of beer in one sitting for men (women).

[31]To capture nonlinear change in drinking behavior, the drinker dummy was reversed in sign when calculating index variable.

The negative coefficients for intensive exercises and positive coefficients for walking are suggestive evidence that screening preparation guidelines may play a role.[32]

5.5 Complier and heterogeneity analysis

After establishing baseline results using the entire sample, I explore heterogeneous treatment effects for different subsamples. As in the baseline regression, I do it in two steps. First, I explore heterogeneity in the first stage regression. This is effectively a complier analysis. Angrist and Pischke (2008) explains that the relative likelihood of being a complier by subsamples can be learned from comparing the first stage coefficients. Hence, I re-estimate first stage equation given in Equation 1 by restricting the sample to a particular demographic group, and compare the first stage coefficient to the one estimated using the entire sample. Given that it was robust to the inclusion of control variables in 3, I estimate it without any covariate. Next, I investigate heterogeneous response in health care usage and behavior. This is done by estimating the LATE of screening using IV strategy without any covariate after restricting the sample to a particular subsample. For outpatient care usage outcomes, the coefficient estimates for visits to general hospitals, visits to local clinics and first visits will be presented. For inpatient and emergency care, the coefficient estimates for total hospital visits for inpatient and emergency care will be presented. For behavioral outcomes, I only present the index variables for each domain of behaviors.

The result for complier analysis is presented in Figure **??**. The bottom circle indicates screening rate in the absence of free screening and the length of arrow represents the increase in screening rate.[33] Compared to total increase estimated using the entire sample, there is larger increase in screening rate among female, non-working, old, married, less educated and poor subsamples. This reveals who are more likely to respond to

[32]Note that our estimates capture the sum of anticipatory and ex-post effects and we cannot disentangle them. So, it is hard to make any definitive claim.

[33]The estimate for constant term in Table A11 is the baseline screening rate without free screening, hence the location of the bottom circle in Figure A11. The coefficient corresponds to the length of the arrow, and the location of the top of the arrow equals to the sum of constant term and the coefficient term. Young and poor subsample represent those below median age and household income, respectively. Those with low education represent people below median years of schooling, that is, at most 12 years of schooling, and hence, high school graduate.

an economic incentive of free screening. Furthermore, these are compliers that drive our estimates for local average treatment effect.

Note that those with low education and poor people experience larger growth in screening rate than those with high education and rich people, despite lower base screening rate without incentive. Figure 8 plots the similar first stage coefficients for people with different education levels. While the baseline screening rate increases with education level, the increase in screening rate due to free provision is highest in the group with middle school education. Similarly, Figure 9 that plots the first stage coefficients for people in different income deciles makes it clear that the baseline screening rate rises with income, but the increase in screening rate does not. If anything, those with lower education and income experience larger increase in screening despite lower base rate to achieve the similar level of ultimate screening rate. Given that people with high education and income are more likely to be invested in their health and be always-takers, the provision of free screening as a public health policy is successful in inducing screening among disadvantaged population.

Lastly, I examine the heterogeneous treatment effect in medical service usage and health behavior by running the two stage least square regression after restricting the sample to a particular subsample. The figures that summarize the heterogeneous treatment effects are provided in the Appendix. Figure A1 plots the percentage change in the outcome variables, visits to general hospital, visits to local clinic and first visits with 95% confidence intervals.[34] The estimates for the total sample shows increase in visits to general hospitals, drop in visits to local and increase in first visits, as confirmed in Table 4. The coefficient estimates for different subsamples indicate slight differences in the magnitude or standard errors, but the general pattern is similar across groups, suggesting minimal heterogeneity. Figure A2 plots the coefficients for the number of hospital visits for inpatient and emergency care, and it also presents similar pattern across different subsamples.

Figure A3 summarizes the heterogeneous behavioral response to screening. The ex-

[34]Percentage change is calculated by LATE estimate divided by the mean of the control group.

ercise index is noisy, making it hard to discern any heterogeneity. Comparing smoking and drinking behaviors suggests the two behavioral changes occur in different pool of populations. Smoking is reduced significantly among non-working and old people, while drinking decreases significantly among working and young population. Given that our sample of biannual screening eligible people consists of formally employed young individuals and those with 40 years or above, there is a lot of overlap between non-working and old people. Lastly, heterogeneous responses along education and income dimension reveals that people with higher education and income tend to make larger adjustment in both the smoking and drinking behavior in response to screening.

5.6 Spillover effect

This section provides evidence that health screening behavior is subject to spillover effect using random variation in the composition of screening eligibilities within a couple. Current dataset is based on household level sampling, and this allows identifying relationships between household members. The dataset allows identifying all the couples consisting of a husband and a wife for all the currently married people. I take advantage of this information to investigate the spillover effects in screening and health behavior between spouses. The research question is the spillover effect in screening uptake, that is, whether spouse's screening increases one's likelihood of getting screened. Given fear of screening, one is likely to put off or avoid screening. However, doing it together can help overcome the fear or procrastination and increase take-up. If there is a positive spillover in screening take-up, the policy should take it into account and assign screening eligibility so as to maximize the take-up. I check this hypothesis by re-estimating the first stage regression with an added regressor, spouse's screening eligibility. Hence, I estimate the effect of one's own and the spouse's eligibility on one's screening take-up. An interaction term of the two eligibility terms will also be added to check if there is an additional increase when couples are eligible for screening in the same year. The specification is shown in the following equation.

$$Screening^A_{ict} = \gamma_0 + \gamma_1 Eligible^A_{ict} + \gamma_2 Eligible^B_{ict} + \gamma_3 Eligible^A_{ict} \times Eligible^B_{ict} + \mathbf{X^A_{ict}} + \mathbf{X^B_{ict}} + \psi_{ict} \tag{4}$$

The outcome variable, $Screening^A_{ict}$, is one's own screening status of individual i in couple c at year t. Independent variable $Eligible^A_{ict}$ refers to one's own eligibility and $Eligible^B_{ict}$ indicates the eligibility of the individual i's spouse belonging to couple c at year t. Coefficient γ_2 captures the spillover effect from spouse, and γ_3 reveals if there is any additional increase for couples eligible for screening in the same year.[35] ψ_{ict} is an error term and the standard errors are clustered at the couple level. Note that the analysis requires adjustment of the sample. I restrict the sample to the currently married couples, both of whom are subject to biannual health screening. This leaves me an effective sample size of 40,258 individual-year pairs.

The previous equation estimates the reduced form effect of spouse's screening eligibility on one's own screening take-up. To translate this into the effect of spouse's actual screening take-up on one's own take-up, I estimate regression Equation 5 using the spouse's screening take-up as a regressor instrumented by the screening eligibility. The coefficient of spouse's screening status, δ_2, provides the LATE estimate that summarizes the spillover effect in screening behavior.

$$Screening^A_{ict} = \delta_0 + \delta_1 Eligible^A_{ict} + \delta_2 Screening^B_{ict} + \mathbf{X^A_{ict}} + \mathbf{X^B_{ict}} + \varepsilon_{ict} \tag{5}$$

Table 7 shows the regression results for the first stage. Column 1 shows that the spouse's eligibility for free screening leads to 21 percentage point increase in spouse's screening rate. This provides justification for using spouse's eligibility as an IV for spouse's screening take-up. The first-stage coefficient is slightly larger than the original one obtained using the whole sample from Table 3, which is consistent with the finding that married people are more likely to be a complier. The coefficient estimate is robust to

[35] Note that a couple is subject to screening either in the same year (even-even or odd-odd) or in different years. (even-odd or odd-even) The first case is equivalent to an age difference being even-numbered.

the inclusion of one's own and spouse's covariates as shown in column 2. Columns 3 and 4 estimate the spillover effect in screening take-up. In addition to 22 percentage point increase in response to one's own eligibility, the spouse's eligibility additionally increases the screening rate by 1.8 percentage point. The estimate is tightly estimated and is robust to control variables. However, the interaction term was statistically indistinguishable from zero. This suggests there is an additively separable spillover effect. Spouse's screening eligibility increases one's own take-up by 1.8 percentage point regardless of whether they are eligible on the same year or on different years. This implies there indeed is a spillover effect between spouses, but it does not support the hypothetical policy recommendation that spouses should be eligible for free screening on the same year to maximize take-up. Lastly, we use the spouse eligibility as an IV for spouse screening to estimate the Wald estimate that summarizes the effect of spouse's screening on one's own screening rate. The estimates are shown in columns 5 and 6. It shows spouse's screening increases one's own screening by 8 percentage point. Given that baseline screening rate without any incentive was 10 percent, this is an economically sizable increase in screening rate.

6 Conclusion

This paper evaluates the public health screening program currently implemented in Korea. Using biannual free screening based on the year of birth as an IV, I estimate the causal effect of screening on health care utilization and health behaviors. The results show screening brings about changes in the pattern of outpatient, inpatient and emergency care usage. Screening leads to an increase in hospital visits for a new illness and the demand for specialized medical services rises at the cost of accessible local care. Its impact on inpatient care was small and not precisely estimated, but the use of emergency care dropped with screening. These IV estimates summarize the aggregate impact of screening in the one year period. Event study analyses based on granular health care history data suggest intertemporal reallocation of outpatient and inpatient care, that is, people delay the use until after the screening.

In response to regular screening, people also modify their health behaviors and display healthier habits. Smoking mildly decreases, while walking increases. There is a nonlinear effect on drinking. Screening induces some new drinker, albeit moderate. Nevertheless, it also helps more frequent drinkers and binge drinkers to significantly decrease their drinking frequency. Overall, the evidence indicates people respond by adjusting their behaviors in a way perceived to be healthy.

Free provision of screening induces a certain group of population more to participate, such as female, nonworking, old and married. It also helps people with low levels of education and household income to get screening, thereby raising their screening rate to be at par with those with high education and income. Screening take-up is also affected by peer effect. High rate of screening participation among married group reflects the fact that spousal spillover effect has a significant impact on participation.

This study documents many significant impacts on health care usage and behaviors than previous studies on screening. One potential explanation could be that it is hard to find random variation in screening participation, and thus, prior studies rely on other variation in the program design conditional on screening. Another could be the repeated implementation of the program. The fact that people perceive it to be a recurring event that will keep happening in the future could have a larger impact than when it is perceived to be a one-off event. The recurring aspect of the program limits the impact evaluation to focus on short term outcomes, but in return, could result in larger short term impact. Future studies providing improved understanding of the frequency of screening and the magnitude of impact could help design better policy.

References

Hyuncheol Bryant Kim, Suejin A Lee, and Wilfredo Lim. Knowing is not half the battle: Impacts of information from the national health screening program in korea. *Journal of health economics*, 65:1–14, 2019.

Hyuncheol Bryant Kim and Sun-mi Lee. 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, 2017.

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

Rebecca Thornton. The demand for and impact of learning hiv status: Evidence from a field experiment. *Informal paper, Harvard University, Cambridge, MA*, 2005.

Emily Oster, Ira Shoulson, and E Dorsey. Optimal expectations and limited medical testing: evidence from huntington disease. *American Economic Review*, 103(2):804–30, 2013.

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

Toshiaki Iizuka, Katsuhiko Nishiyama, Brian Chen, and Karen Eggleston. False alarm? estimating the marginal value of health signals. *Journal of Public Economics*, 195: 104368, 2021.

Marije Deutekom, Fleur Vansenne, Kirsten McCaffery, Marie-Louise Essink-Bot, Karien Stronks, and Patrick MM Bossuyt. The effects of screening on health behaviour: a summary of the results of randomized controlled trials. *Journal of Public Health*, 33 (1):71–79, 2011.

DA Wood, AL Kinmonth, GA Davies, J Yarwood, SG Thompson, SDM Pyke, Y Kok, R Cramb, C Le Guen, TM Marteau, et al. Randomised controlled trial evaluating cardiovascular screening and intervention in general practice: principal results of british family heart study. *Bmj*, 308(6924):313–320, 1994.

Imperial Cancer Research Fund OXCHECK Study Group. Effectiveness of health checks conducted by nurses in primary care: final results of the oxcheck study. *BMJ*, 310 (6987):1099–1104, 1995.

Inger K Larsen, Tom Grotmol, Kari Almendingen, and Geir Hoff. Impact of colorectal cancer screening on future lifestyle choices: a three-year randomized controlled trial. *Clinical Gastroenterology and Hepatology*, 5(4):477–483, 2007.

Irene M Strychar, François Champagne, Parviz Ghadirian, André Bonin, Milos Jenicek, and Thomas M Lasater. Impact of receiving blood cholesterol test results on dietary change. *American journal of preventive medicine*, 14(2):103–110, 1998.

Emily Oster. Diabetes and diet: Behavioral response and the value of health. Technical report, National Bureau of Economic Research, 2015.

Emily Oster. Hiv and sexual behavior change: Why not africa? *Journal of health economics*, 31(1):35–49, 2012.

Alexander N Slade. Health investment decisions in response to diabetes information in older americans. *Journal of Health Economics*, 31(3):502–520, 2012.

Michael Grossman. On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–55, 1972.

Gary S Becker and Kevin M Murphy. A theory of rational addiction. *Journal of political Economy*, 96(4):675–700, 1988.

Richard H Thaler and Hersh M Shefrin. An economic theory of self-control. *Journal of political Economy*, 89(2):392–406, 1981.

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

Donald S Kenkel and Jody Sindelar. Economics of health behaviors and addictions: contemporary issues and policy implications. In *The Oxford Handbook of Health Economics*. 2011.

Donald S Kenkel. Health behavior, health knowledge, and schooling. *Journal of Political Economy*, 99(2):287–305, 1991.

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

Christopher J Ruhm. Are recessions good for your health? *The Quarterly journal of economics*, 115(2):617–650, 2000.

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

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

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

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

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

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

Hyuncheol Bryant Kim, Booyuel Kim, Cristian Pop-Eleches, and Jaehyun Jung. Male circumcision, peer effects and risk compensation. *Unpublished Manuscript, Department of Policy Analysis and Management, Cornell University*, 2018.

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

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

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

Peter H Westfall and S Stanley Young. *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons, 1993.

Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics*. Princeton university press, 2008.

Table 1: Survey cohort

	Cohort 1	Cohort 2
Year	2009–2017	2014–2017
Number of households	6282	1852
Number of individuals	19153	5246
Individual attrition after 3 years	17%(2012)	14%(2017)
Average number of participating in surveys	6.5/9	3.6/4

Notes: This table presents the number of observations and attrition of the 2 survey cohorts in the dataset. The number of households and individuals is the original size of the two cohorts in the starting years 2009 and 2014. The maximum number of responses is 9 (=2017-2009+1) for cohort 1 and 4 (=2017-2014+1) for cohort 2.

7 Tables

Table 2: Balance table

	Treatment group	Control group	Difference
Individual characteristics			
Age	55.80	56.02	–0.23*
	(15.71)	(15.69)	(0.12)
Female	0.56	0.55	0.00
	(0.50)	(0.50)	(0.00)
Married	0.74	0.74	–0.00
	(0.44)	(0.44)	(0.00)
Years of schooling	11.03	11.01	0.03
	(4.78)	(4.79)	(0.04)
Working status	0.53	0.53	–0.00
	(0.50)	(0.50)	(0.00)
Handicapped	0.09	0.09	–0.00
	(0.28)	(0.29)	(0.00)
Employment-based insurance	0.56	0.56	–0.00
	(0.50)	(0.50)	(0.00)
Individual income	1461.42	1470.06	–8.65
	(2148.42)	(2153.68)	(15.87)
Household characteristics			
Income decile	5.78	5.75	0.03
	(2.94)	(2.94)	(0.02)
Household income	4331.70	4315.12	16.58
	(4044.67)	(3986.58)	(29.62)
House ownership	0.70	0.70	0.00
	(0.46)	(0.46)	(0.00)
Number of Household members	3.12	3.11	0.01
	(1.32)	(1.32)	(0.01)
Observations	37024	36511	
Share	(54.19)	(53.44)	
F-stat (12, 73370)			1.16
p-value			(0.30)

Notes: This table reports the balance of covariates between the treatment and control group defined by the eligibility for free general health screening. Treatment group consists of individuals who are born in even-numbered years in even-numbered years and those born in odd-numbered years in odd-numbered years. Another way to phrase it is the group of people whose age is even-numbered. Control group consists of those whose age is odd-numbered. Note that the unit of observation is individual-year pair. Treatment group is eligible for free general health screening and the control group is not. Every year, an individual takes turn belonging to treatment and control group. A */**/*** indicates significance at the 10/5/1% levels.

Table 3: First stage: Effect of free health screening eligibility on take-up

	(1)	(2)	(3)
		Outcome var: Health screening take-up	
Eligible	0.180***	0.182***	0.185***
	(0.003)	(0.003)	(0.003)
N	73,535	73,372	70,913
Adj R^2	0.053	0.077	0.160
Controls		Y	
Year FE		Y	Y
Individual FE			Y

Notes: This table presents the first stage regression result. Outcome variable is the take-up of general health screening. Independent variable is eligibility for free health screening. Control variables are listed in section 4. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table 4: Health screening and outpatient care

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Outpatient care								
Number of hospital visits								
Total	19.021	18.869	-0.153	-0.847	-4	0.573	0.400	73535
General hospital (bed 100+)	2.473	2.516	0.043	0.238	10	0.154	0.400	73535
Local hospital (bed 30-100)	1.448	1.461	0.013	0.072	5	0.144	1.000	73535
Local clinic (bed 0-30)	15.100	14.891	-0.209**	-1.157**	-8	0.524	0.200	73535
Hospital bill								
Total	340847	336962	-3885	-21553	-6	25571	1.000	73535
General hospital (bed 100+)	83558	85175	1617	8971	11	9242	0.800	73535
Local hospital (bed 30-100)	40985	40606	-378	-2098	-5	9061	1.000	73535
Local clinic (bed 0-30)	215702	210526	-5176	-28715	-13	22093	0.600	73535
Drug expenditures								
Total	113003	112472	-531	-2945	-3	3819	1.000	73535
General hospital (bed 100+)	41091	41341	249	1383	3	2879	1.000	73535
Local hospital (bed 30-100)	8168	8093	-75.018	-416	-5	966	1.000	73535
Local clinic (bed 0-30)	61677	60975	-702*	-3896*	-6	2345	0.400	73535
Panel B: Outpatient care for a new illness								
First hospital visit for a new illness								
Total	3.696	3.758	0.063***	0.348***	9	0.108	0.000	73535
General hospital (bed 100+)	0.353	0.371	0.018***	0.100***	28	0.032	0.000	73535
Local hospital (bed 30-100)	0.286	0.293	0.007	0.038	13	0.026	0.400	73535
Local clinic (bed 0-30)	2.847	2.884	0.037**	0.203**	7	0.093	0.200	73535
First hospital bill for a new illness								
Total	89527	93896	4369**	24238**	27	10707	0.200	73535
General hospital (bed 100+)	20347	21728	1380**	7658**	38	3689	0.200	73535
Local hospital (bed 30-100)	13010	13988	978	5426	42	4079	0.600	73535
Local clinic (bed 0-30)	55986	57996	2010	11153	20	9190	0.600	73535
First drug expenditures for a new illness								
Total	10982	11351	369***	2048***	19	662	0.000	73535
General hospital (bed 100+)	1924	2011	87.830	487	25	407	0.600	73535
Local hospital (bed 30-100)	1000	1035	35.198	195	20	188	0.800	73535
Local clinic (bed 0-30)	7947	8193	246***	1363***	17	466	0.000	73535

Notes: This table presents the effect of health screening on outpatient care usage and medical expenditure. Both the number of hospital visits and medical expenditures exclude the visits for screening and medical expenditures incurred for screening. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. A */**/*** indicates significance at the 10/5/1%.
[a] Westfall-Young adjusted p-values account for 24 hypotheses tested with 5 bootstrap replications.

Table 5: Health screening and inpatient/emergency care

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Inpatient care								
Number of hospital visits								
Total	0.236	0.237	0.001	0.004	2	0.027	1.000	73535
General hospital (bed 100+)	0.125	0.123	-0.002	-0.012	-9	0.021	1.000	73535
Local hospital (bed 30-100)	0.076	0.080	0.004	0.020	27	0.013	0.800	73535
Local clinic (bed 0-30)	0.036	0.035	-0.001	-0.005	-14	0.009	1.000	73535
Hospital bill								
Total	209413	208121	-1292	-7170	-3	40095	1.000	73535
General hospital (bed 100+)	128569	126258	-2312	-12824	-10	32903	1.000	73535
Local hospital (bed 30-100)	66662	68826	2164	12007	18	21596	1.000	73535
Local clinic (bed 0-30)	14113	13021	-1092	-6061	-43	6255	1.000	73535
Drug expenditures								
Total	85.934	90.767	4.833	26.811	31	90.724	1.000	73535
General hospital (bed 100+)	43.834	58.667	14.833	82.287	188	80.775	1.000	73535
Local hospital (bed 30-100)	17.758	14.806	-2.952	-16.374	-92	31.448	1.000	73535
Local clinic (bed 0-30)	24.343	17.294	-7.048	-39.101	-161	26.565	0.800	73535
Panel B: Emergency care								
Number of hospital visits								
Total	0.127	0.121	-0.006*	-0.032*	-25	0.018	0.400	73535
General hospital (bed 100+)	0.094	0.088	-0.006**	-0.035**	-37	0.015	0.200	73535
Local hospital (bed 30-100)	0.032	0.032	0.000	0.002	5	0.009	1.000	73535
Local clinic (bed 0-30)	0.001	0.001	0.000	0.001	125	0.002	1.000	73535
Hospital bill								
Total	7444	7372	-71.574	-397	-5	2352	1.000	73535
General hospital (bed 100+)	6589	6559	-29.872	-166	-3	2295	1.000	73535
Local hospital (bed 30-100)	814	791	-23.072	-128	-16	472	1.000	73535
Local clinic (bed 0-30)	40.059	21.429	-18.630	-103	-258	105	1.000	73535
Drug expenditures								
Total	50.376	54.718	4.343	24.091	48	49.191	1.000	73535
General hospital (bed 100+)	26.558	31.530	4.972	27.581	104	44.912	1.000	73535
Local hospital (bed 30-100)	23.776	22.940	-0.836	-4.640	-20	20.002	1.000	73535
Local clinic (bed 0-30)	0.041	0.248	0.207	1.151	2801	1.233	1.000	73535

Notes: This table presents the effect of health screening on inpatient and emergency care usage and medical expenditures. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 12 hypotheses tested in each inpatient and emergency care domain with 5 bootstrap replications.

Table 6: Health screening and behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Smoking								
Extensive margin								
Smoker	0.193	0.190	−0.003	−0.014	-7	0.009	0.400	71691
Frequency								
Smoking days per year	68.018	67.159	−0.860	−4.664	-7	3.053	0.400	71691
Smoking once a week or more	0.190	0.188	−0.002	−0.013	-7	0.009	0.400	71691
Smoking everyday	0.184	0.182	−0.002	−0.013	-7	0.008	0.400	71691
Amount								
Cigarettes per day	2.838	2.816	−0.023	−0.122	-4	0.155	0.600	71691
Smoking 3 cigarettes or more	0.184	0.182	−0.003 *	−0.014 *	-8	0.009	0.400	71691
Smoking 10 cigarettes or more	0.154	0.151	−0.003 *	−0.016 *	-10	0.008	0.400	71691
Standardized treatment effect[b]								
Smoking index				−0.033		0.021		
Panel B: Drinking								
Extensive margin								
Drinker	0.633	0.638	0.005**	0.025**	4	0.012	0.400	71814
Frequency								
Drinking once a month or more	0.497	0.494	−0.003	−0.017	-3	0.013	0.600	71814
Drinking once a week or more	0.287	0.283	−0.004 *	−0.021 *	-7	0.012	0.600	71814
Drinking everyday	0.057	0.054	−0.003**	−0.018**	-31	0.007	0.000	71814
Binge drinking once a month or more	0.217	0.215	−0.002	−0.010	-4	0.012	0.600	71796
Binge drinking once a week or more	0.130	0.126	−0.003	−0.017	-13	0.010	0.600	71796
Binge drinking everyday	0.022	0.020	−0.002 *	−0.009 *	-39	0.005	0.600	71796
Amount								
Drinking 5 cups or more	0.255	0.253	−0.002	−0.009	-4	0.012	0.600	71793
Drinking 10 cups or more	0.070	0.072	0.001	0.008	11	0.008	0.600	71793
Standardized treatment effect[b]								
Drinking index				−0.037**		0.015		
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.215	0.214	−0.001	−0.005	-2	0.014	1.000	71813
Doing moderate exercise	0.351	0.350	−0.001	−0.005	-2	0.017	1.000	71812
Doing walking exercise	0.772	0.777	0.005 *	0.025 *	3	0.015	0.200	71812
Frequency								
Days of vigorous exercise	36.886	36.859	−0.027	−0.146	-0	2.955	1.000	71813
Days of moderate exercise	71.154	71.085	−0.069	−0.375	-1	3.995	1.000	71812
Days of walking	206	208	1.951**	10.602**	5	5.031	0.200	71812
Amount								
30 min vigorous exercise	0.166	0.167	0.001	0.006	4	0.013	1.000	71813
30 min moderate exercise	0.256	0.259	0.003	0.019	7	0.015	0.600	71811
30 min walking	0.416	0.412	−0.004	−0.021	-5	0.018	0.600	71812
Standardized treatment effect[b]								
Exercise index				0.014		0.021		

Notes: This table presents the effect of health screening on health behaviors. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for the number of hypotheses tested in each behavior domain with 5 bootstrap replications.

[b] Standardized treatment effect gives equal weight to all the outcomes in each behavior domain. For drinking, the drinker variable's sign is reversed before calculation.

Table 7: Spillover effect in screening behavior

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome var: Spouse screening		Outcome var: Own screening			
Eligible			0.210***	0.210***	0.208***	0.209***
			(0.008)	(0.007)	(0.005)	(0.005)
Spouse eligible	0.210***	0.211***	0.018***	0.017***		
	(0.005)	(0.005)	(0.006)	(0.006)		
Eligible × Spouse eligible			–0.001	0.001		
			(0.012)	(0.011)		
Spouse screening					0.082***	0.082***
					(0.023)	(0.023)
N	40,258	40,170	40,258	40,170	40,258	40,170
Controls		Y		Y		Y
Year FE		Y		Y		Y
Specification	OLS	OLS	OLS	OLS	IV	IV

Notes: This table reports the spillover effect in screening take-up between spouses. The sample is restricted to married couples both of whom are eligible for biannual free health screening. Outcome variable is the spouse's health screening take-up in columns 1 and 2, and own take-up in columns 3-6. Independent variables are one's own and spouse's eligibility for health screening and spouse's screening take-up. Columns 1-4 use OLS regression. Column 5 and 6 use spouse's screening eligibility as an IV for spouse's screening take-up. Control variables include the set of variables listed in section 4 plus the same set of control variables pertaining to the spouse. Standard errors are clustered at couple level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

8 Figures

Figure 1: Screening result form

Cardiovascular Disease Risk Assessment

*Cardiovascular disease refers to ailments including stroke, myocardial infarction, etc.

Name **Sex** **Age** **Date of examination** 0000-00-00

Risk of cardiovascular disease

Your risk of cardiovascular disease (Compared to average for your (age) (sex))	Probability that you develop cardiovascular disease within 10 years	Cardiovascular age
0.00 times	Mr./Mrs./Ms. 0.0% Average for your (age)(sex) 0.0% 0 20 40 60 80 100	**00 years**

Learn about health related factors

Health related factors	Current condition →	Target condition	Health signals
Weight Waist line		Below 65kg Below 90cm	Danger
Exercise		Five or more times per week	Danger
Drinking		Not more than 2 glasses	Danger
Blood pressure		Below 120/80	Danger
Smoking		Sustain nonsmoking	Caution
Fasting blood sugar		Below 100	Caution
Total cholesterol LDL cholesterol		Below 200 Below 130	Safe

Mr./Mrs./Ms. , the results above are your current health conditions and goals based on your questionnaire answers and test results. You will need to actively improve the items corresponding to the health signals "Caution" or "Danger". If you are taking any medication for hypertension, diabetes mellitus, or dyslipidemia, your health signal result will be "Caution" even when your blood pressure, fasting blood sugar level, and cholesterol fall within the goal ranges. Please continue to manage your health conditions.

※ As the goals above follow the general guidelines, they may vary depending on your health conditions. Please consult your physician.

To improve health related factors

The probability that you develop cardiovascular disease within 10 years should decrease by	Cardiovascular age
00% (0.0% → 0.0%)	**00 → 00**

This certifies the information above as your health checkup results.

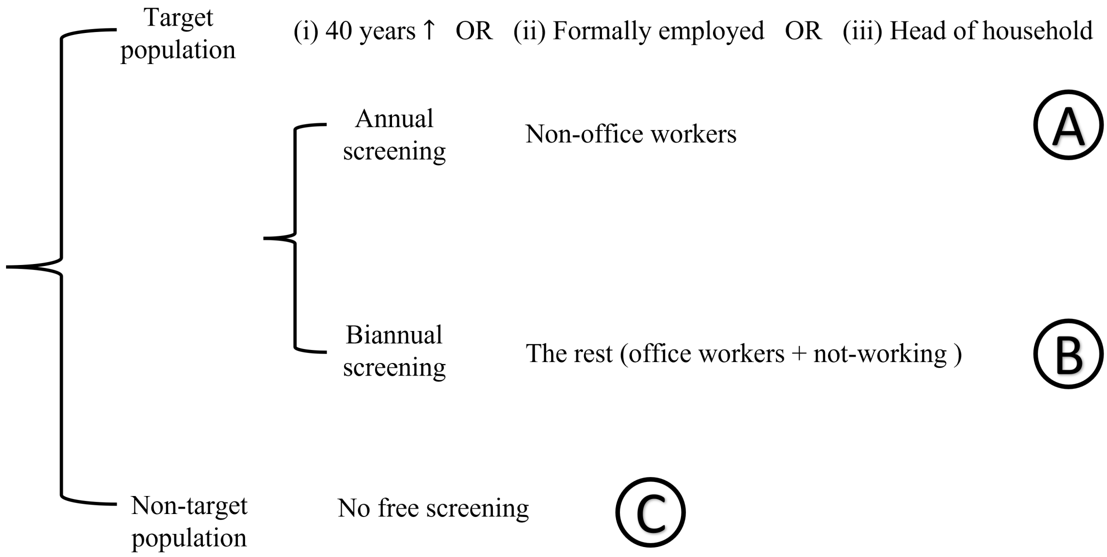
20 . . .

Date of assessment: Physician: License (Qualification) / No. Name (Signature)

(Medical Care Institution)

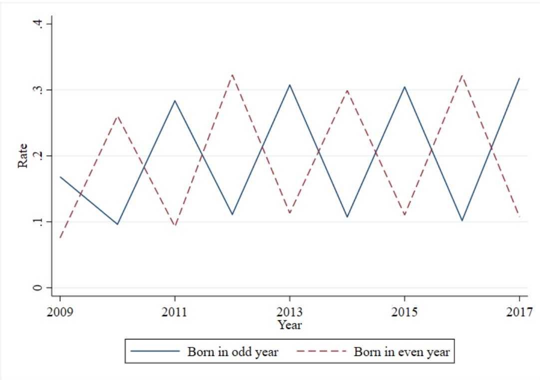
Notes: The figure shows an example screening result form that is given to the examinee with the results. It includes current health conditions and target conditions with respect to various health indicators.

Figure 2: Sample composition and analytical sample

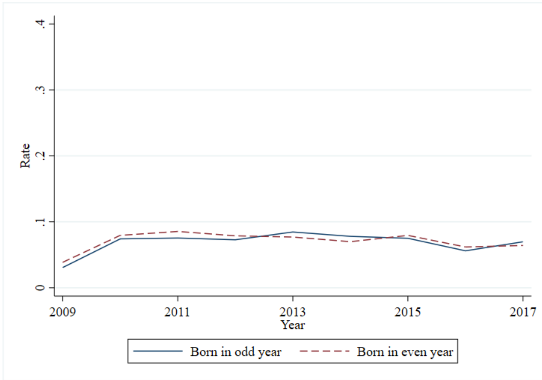


Notes: The figure shows the composition of the sample with regard to health screening eligibility and frequency. One of the three conditions should be satisfied to be a target population. Otherwise, one belongs to non-target population (C) without any free screening. The target population is decomposed into a group of non-office workers (A) entitled to annual free screening and the rest (B) entitled to biannual free screening. The analytical sample is group (B) eligible for biannual free screening. The robustness check using (A) + (B) as analytical sample is provided in the Appendix.

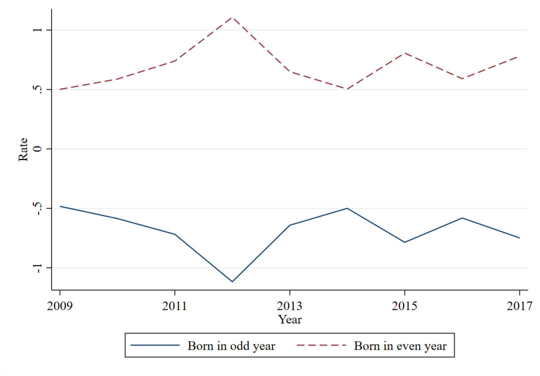
Figure 3: First stage and reduced form plots



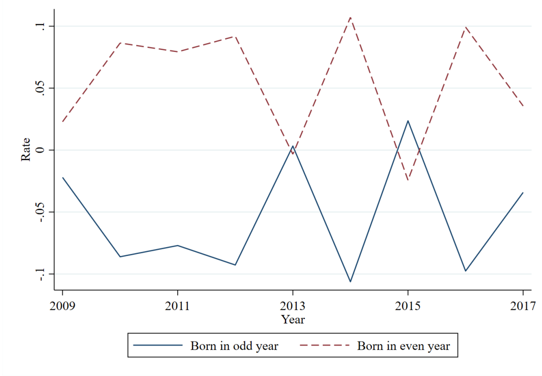
(a) Screening rate for analytical sample



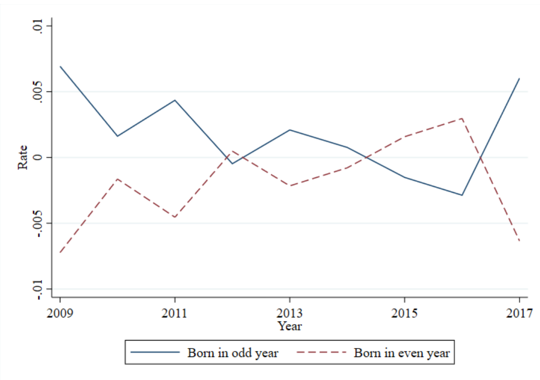
(b) Screening rate for non-target population



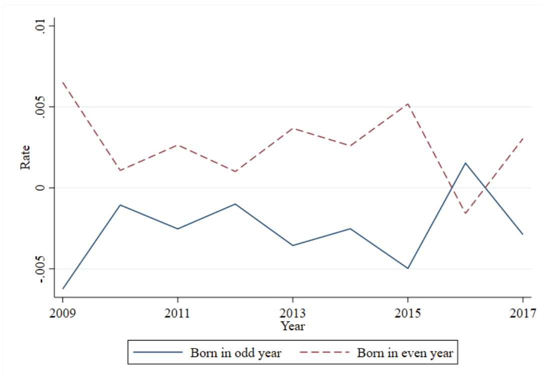
(c) Detrended aggregate hospital visits for outpatient care



(d) Detrended first hospital visits due to a new illness for outpatient care



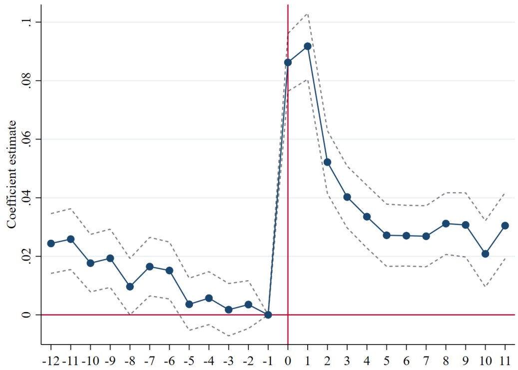
(e) Detrended current drinker



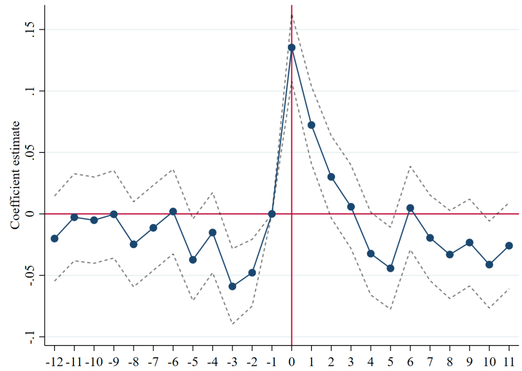
(f) Detrended everyday drinker

Notes: Figure (a) and (b) show the first stage plots for the analytical sample and the non-target population. Refer to Figure 2 for detail on the definition of analytical sample and non-target population. They plot the screening rate for those born in odd or even years, separately. Figure (c), (d), (e), and (f) show the reduced form plots for the analytical sample. They plot the detrended outcome variables for those born in odd or even years, separately. The outcome variables are detrended by demeaning the values at year-by-year basis. Each year values are subtracted by the mean of the corresponding year.

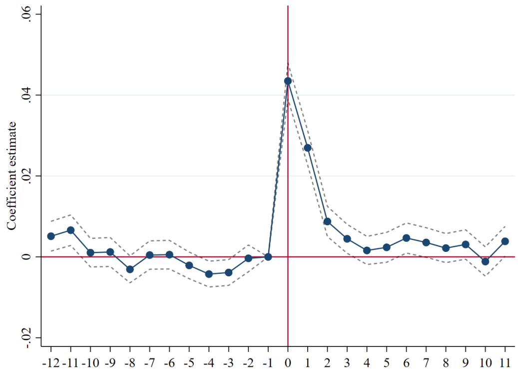
Figure 4: Event study figures for health care use



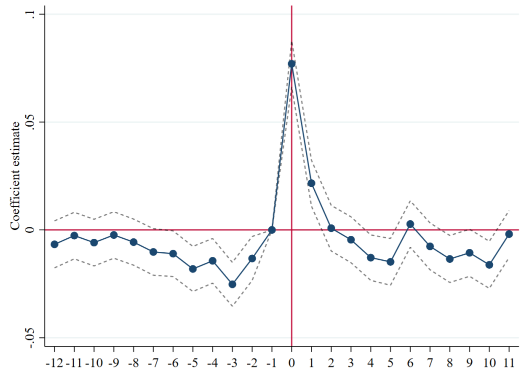
(a) Visit to general hospitals for outpatient care



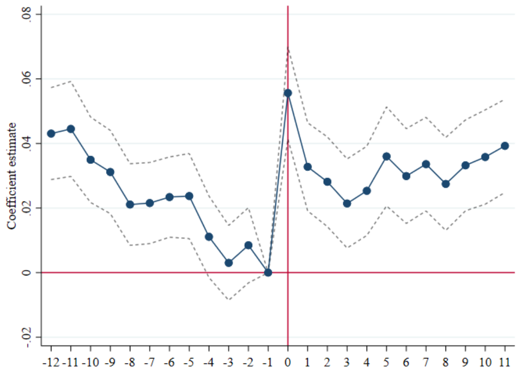
(b) Visit to local clinics for outpatient care



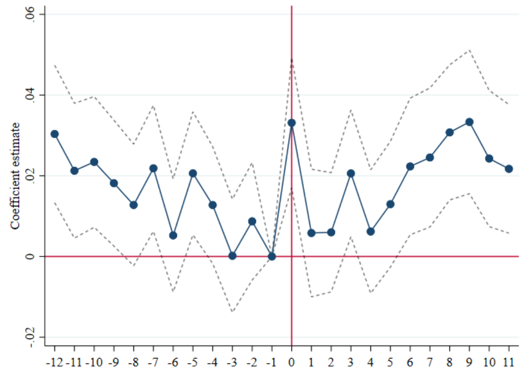
(c) First visit to general hospitals for outpatient care



(d) First visit to local clinics for outpatient care



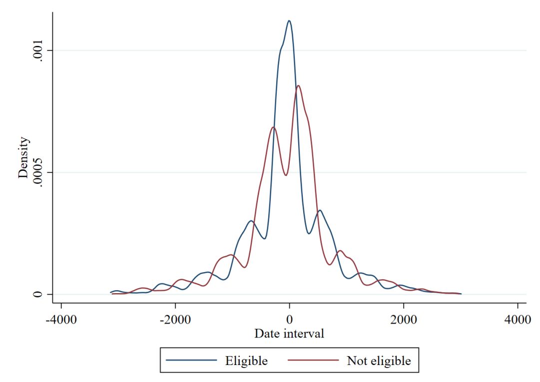
(e) Total visit for inpatient care



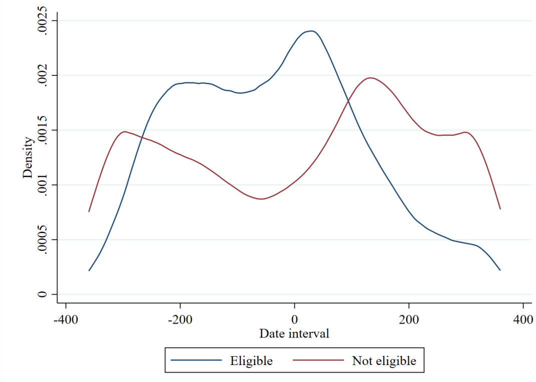
(f) Total visit for emergency care

Notes: Figure (a) and (b) plot the event study coefficients for the number of visits to general hospitals and local clinics for outpatient care with 12 pre- and post-periods in months around the screening date. Figure (c) and (d) plot the same event study coefficients for the first visits to general hospitals and local clinics for a new illness. Figure (e) and (f) plot the same event study coefficients for aggregate visits to any type of hospital for inpatient and emergency care. The month of inpatient care usage is determined by the date of hospitalization. Specification is given in Equation 3 and the standard errors are clustered at individual level. Health screening never-takers are not included and stacking design is used for those who get screening multiple times during the study period. See section 4 for detail.

Figure 5: Date interval between screening and survey date



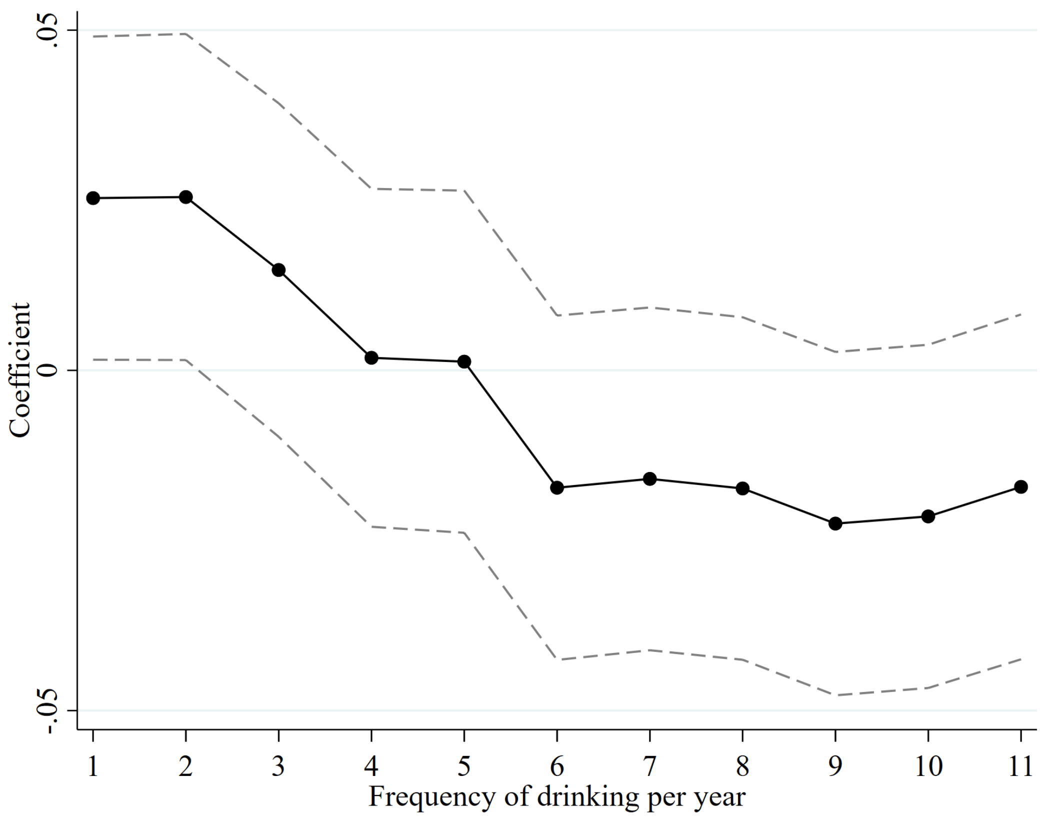
(a) Whole distribution



(b) |Date interval| < 360

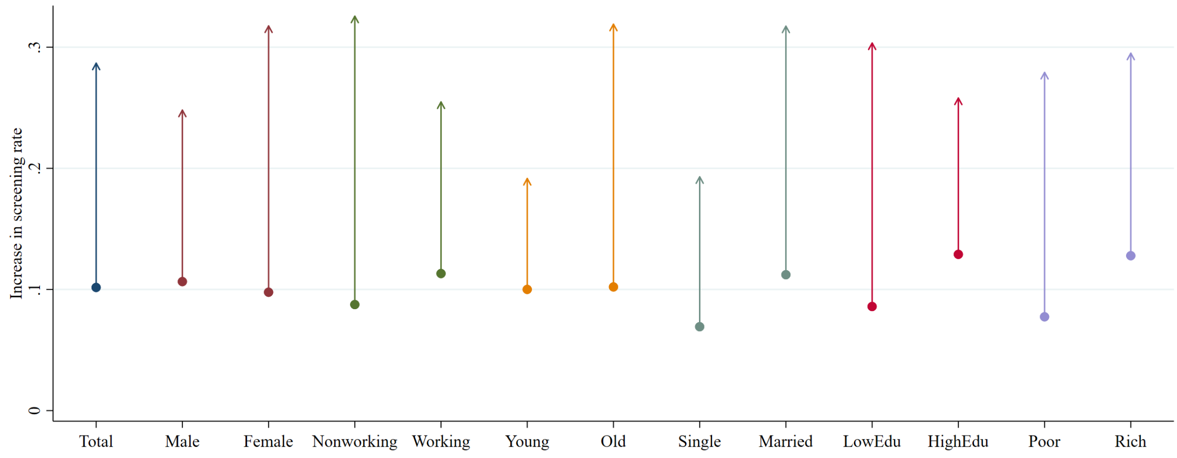
Notes: Figure (a) and (b) plots kernel density plot of the date interval. Date interval is defined as survey date minus screening date in days. Negative number indicates survey happening before screening capturing behavioral changes in anticipation of screening, and positive number indicates survey happening after screening capturing behavioral changes happening as a result of screening. Those eligible and ineligible for free screening are plotted separately to examine which group has more mass around 0.

Figure 6: LATE for drinking frequency less than once a month



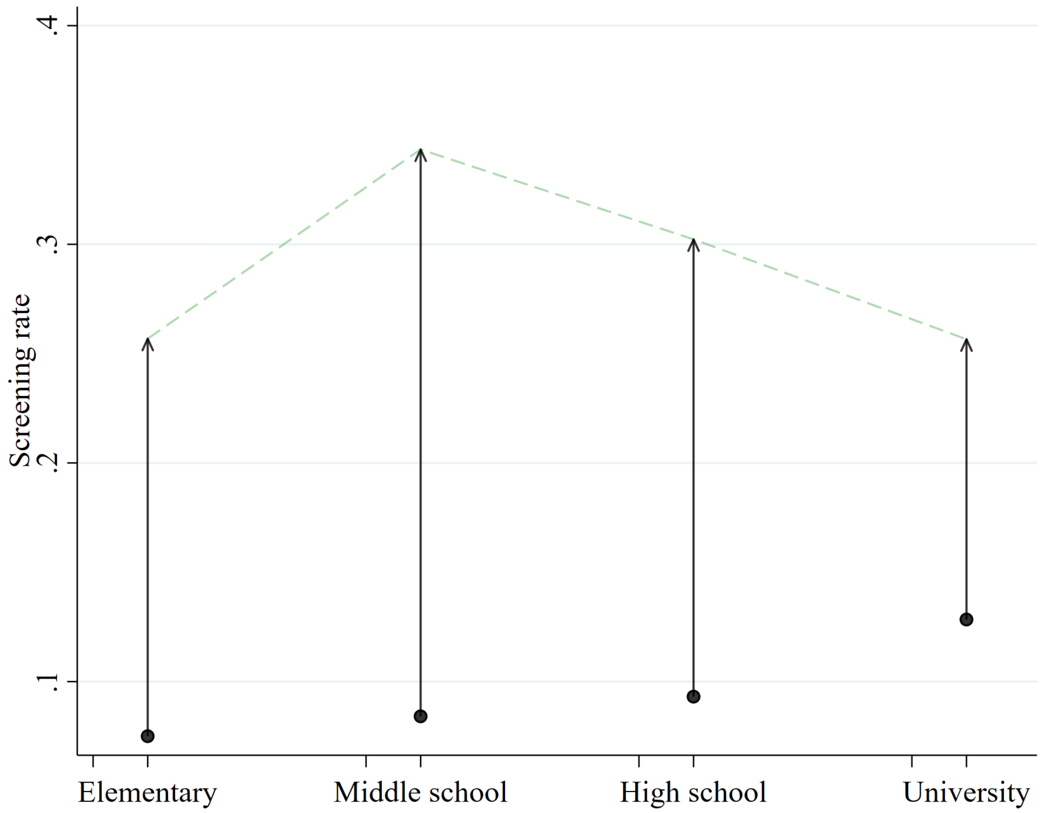
Notes: The LATE coefficients are from two-stage least square regressions (Equation 2) with the outcome variable $\mathbf{1}\Big[\text{Drinking frequency per year} \geqslant j\Big]$ for $1 \leqslant j \leqslant 11$. Current drinkers who answered they drink less than once a month were further asked to specify how many times they drink a year, with the answers ranging from 1 to 11. Screening is shown to increase drinking on the extensive margin. The 11 coefficients pinpoint at which point screening ceases to induce new drinking, but begins to reduce drinking frequency.

Figure 7: Complier analysis



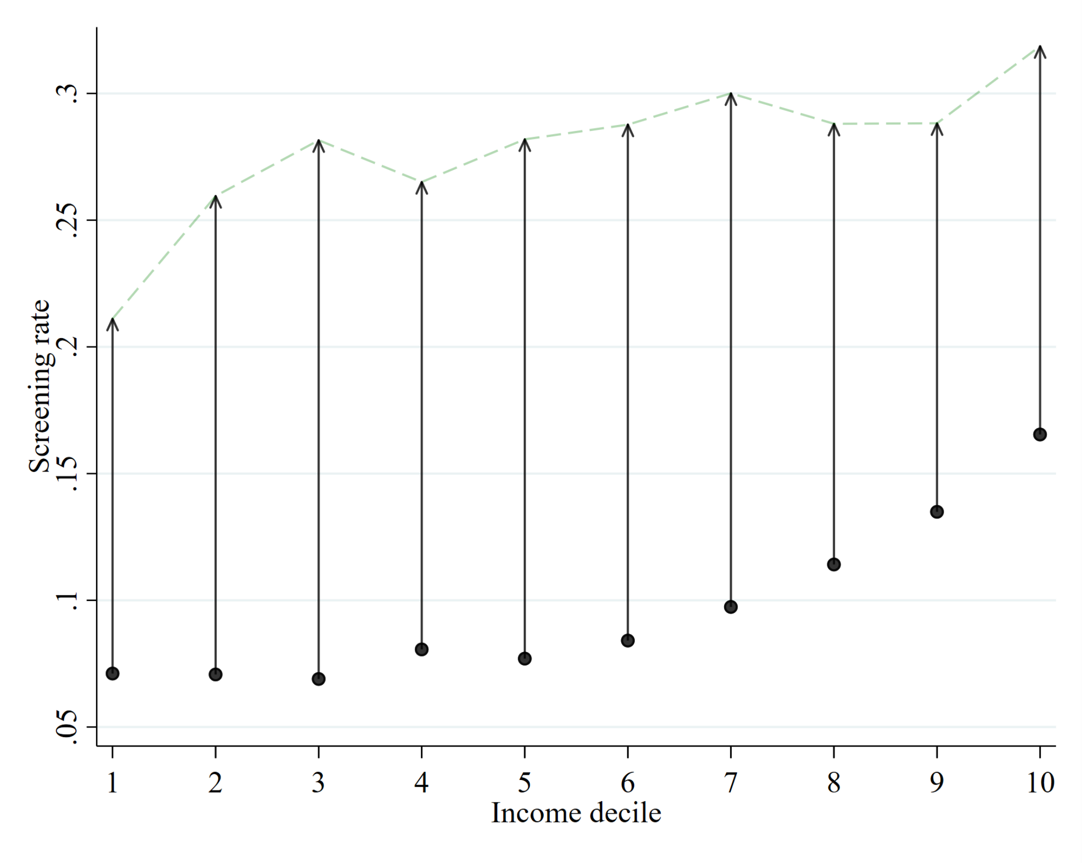
Notes: The figure shows the result for complier analysis. The bottom circles plot the baseline screening rate without any incentive for the total and subsamples. The length of the upward arrow indicates the increase in screening rate in response to free screening, and the top arrow shows the screening rate when free screening is provided.

Figure 8: Complier analysis by education level



Notes: The figure shows the result for complier analysis by education level. The upward arrows show the increase in screening rate from base rate without free screening to the one with free screening. Each education level includes those who dropped out before graduating.

Figure 9: Complier analysis by income level



Notes: The figure shows the result for complier analysis by income decile. The upward arrows show the increase in screening rate from base rate without free screening to the one with free screening. Income decile is defined using household level income.

Appendix A Tables

Table A1: First stage regression for various screenings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Any screening	General health screening	Any cancer screening	Stomach cancer screening	Liver cancer screening	Lung cancer screening	Colorectal cancer screening	Prostate cancer screening	Breast cancer screening	Cervical cancer screening
Eligible	0.189***	0.180***	0.184***	0.171***	0.024***	0.005***	0.028***	0.007***	0.170***	0.148***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)
cons	0.110***	0.100***	0.076***	0.071***	0.025***	0.007***	0.021***	0.008***	0.056***	0.046***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.001)	(0.001)	(0.002)	(0.002)
N	73,535	73,535	73,535	73,535	73,535	73,535	73,535	32,688	40,847	40,847
R^2	0.055	0.053	0.060	0.055	0.004	0.001	0.006	0.001	0.059	0.052
NHIS subsidy		100%		90%	90%	0%	90%	0%	90%	90%
Frequency		2 years		2 years	1 year		1 year		2 years	2 years
Eligible population		Target population		40 or older	40 or older AND high-risk group		50 or older		Women 40 or older	Women 30 or older

Notes: This table presents the first stage regression results for various screenings. Independent variable is the eligibility for free biannual general health screening. The first column considers all types of screening and the third column considers all types of cancer screening participation as outcome variables. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels. The National Health Insurance Service (NHIS) subsidizes 100% of the general health screening and 90% of the 5 types of cancer screenings. The frequency row shows the recommended frequency of screening at which one can be eligible for NHIS subsidy. All screenings with 2 years frequency are carried out based on even-odd design. Those born in even (odd) years are eligible on the even (odd) years. The target population for general health screening is explained in section 3. The frequency and eligibility conditions are based on the 2015 Korean health screening policy.

Table A2: Naive regression for outpatient care usage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Without any control				With controls			
Outcome	Point estimate	Standard error	R^2	N	Point estimate	Standard error	R^2	N
Panel A: Outpatient care								
Number of hospital visits								
Total	2.716***	0.312	0.001	73535	2.652***	0.282	0.179	73372
General hospital (bed 100+)	0.000	0.070	0.000	73535	-0.009	0.064	0.074	73372
Local hospital (bed 30-100)	0.173**	0.079	0.000	73535	0.200**	0.082	0.021	73372
Local clinic (bed 0-30)	2.542***	0.276	0.002	73535	2.462***	0.254	0.143	73372
Hospital bill								
Total	108052***	7983	0.003	73535	73622***	7963	0.031	73372
General hospital (bed 100+)	13619***	2937	0.000	73535	4965*	2973	0.023	73372
Local hospital (bed 30-100)	9704***	2362	0.000	73535	4669*	2418	0.004	73372
Local clinic (bed 0-30)	84581***	6765	0.003	73535	63834***	6777	0.017	73372
Drug expenditures								
Total	16837***	2018	0.001	73535	10109***	1834	0.187	73372
General hospital (bed 100+)	1805	1459	0.000	73535	-1079	1436	0.059	73372
Local hospital (bed 30-100)	1365***	446	0.000	73535	1001**	446	0.023	73372
Local clinic (bed 0-30)	12999***	1226	0.002	73535	9576***	1134	0.132	73372
Panel B: Outpatient care for a new illness								
Hospital visit for a new illness								
Total	1.287***	0.043	0.017	73535	1.087***	0.041	0.118	73372
General hospital (bed 100+)	0.125***	0.011	0.003	73535	0.108***	0.010	0.036	73372
Local hospital (bed 30-100)	0.091***	0.008	0.002	73535	0.076***	0.008	0.018	73372
Local clinic (bed 0-30)	1.031***	0.037	0.015	73535	0.860***	0.036	0.087	73372
Hospital bill for a new illness								
Total	41639***	2904	0.004	73535	31905***	2951	0.016	73372
General hospital (bed 100+)	10819***	1100	0.002	73535	8577***	1084	0.011	73372
Local hospital (bed 30-100)	6412***	1081	0.001	73535	4958***	1080	0.003	73372
Local clinic (bed 0-30)	24311***	2381	0.002	73535	18274***	2453	0.010	73372
Drug expenditures for a new illness								
Total	4409***	209	0.008	73535	3442***	209	0.047	73372
General hospital (bed 100+)	815***	112	0.001	73535	655***	116	0.010	73372
Local hospital (bed 30-100)	376***	51.451	0.001	73535	297***	52.888	0.007	73372
Local clinic (bed 0-30)	3186***	151	0.008	73535	2463***	149	0.044	73372

Notes: This table presents the naive regression result of outpatient care variables on screening participation. The first four columns report the regression results without any control variable. The next four columns report the regression results with control variables listed in section 4. Standard error is clustered at the individual level. A */**/*** indicates significance at the 10/5/1%.

Table A3: Naive regression for inpatient care outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Without any control				With controls			
Outcome	Point estimate	Standard error	R^2	N	Point estimate	Standard error	R^2	N
Panel A: Inpatient care								
Number of hospital visits								
Total	-0.039***	0.007	0.000	73535	-0.035***	0.007	0.036	73372
General hospital (bed 100+)	-0.021***	0.005	0.000	73535	-0.020***	0.005	0.023	73372
Local hospital (bed 30-100)	-0.020***	0.004	0.000	73535	-0.016***	0.003	0.021	73372
Local clinic (bed 0-30)	0.002	0.002	0.000	73535	0.002	0.002	0.006	73372
Hospital bill								
Total	-31891***	8800	0.000	73535	-40147***	8967	0.016	73372
General hospital (bed 100+)	-29182***	6560	0.000	73535	-34779***	6669	0.012	73372
Local hospital (bed 30-100)	-3031	5413	0.000	73535	-4913	5584	0.007	73372
Local clinic (bed 0-30)	374	1504	0.000	73535	-401	1497	0.002	73372
Drug expenditures								
Total	-22.186	16.991	0.000	73535	-19.573	17.682	0.001	73372
General hospital (bed 100+)	-24.681**	12.535	0.000	73535	-23.038*	13.782	0.001	73372
Local hospital (bed 30-100)	5.240	10.475	0.000	73535	5.416	10.160	0.001	73372
Local clinic (bed 0-30)	-2.745	4.681	0.000	73535	-1.951	4.433	0.001	73372
Panel B: Emergency care								
Number of hospital visits								
Total	-0.024***	0.005	0.000	73535	-0.022***	0.005	0.014	73372
General hospital (bed 100+)	-0.018***	0.004	0.000	73535	-0.017***	0.004	0.016	73372
Local hospital (bed 30-100)	-0.006*	0.003	0.000	73535	-0.005	0.004	0.003	73372
Local clinic (bed 0-30)	-0.000	0.000	0.000	73535	-0.000	0.000	0.001	73372
Hospital bill								
Total	-960**	488	0.000	73535	-1268**	500	0.005	73372
General hospital (bed 100+)	-880*	474	0.000	73535	-1215**	485	0.004	73372
Local hospital (bed 30-100)	-86.182	100	0.000	73535	-61.711	107	0.001	73372
Local clinic (bed 0-30)	6.435	31.045	0.000	73535	8.732	30.731	0.001	73372
Drug expenditures								
Total	-19.800**	7.983	0.000	73535	-23.278***	8.657	0.001	73372
General hospital (bed 100+)	-15.507***	5.581	0.000	73535	-17.491***	6.552	0.001	73372
Local hospital (bed 30-100)	-4.819	5.657	0.000	73535	-6.325	5.616	0.001	73372
Local clinic (bed 0-30)	0.526	0.572	0.000	73535	0.538	0.566	0.000	73372

Notes: This table presents the naive regression result of inpatient and emergency care variables on screening participation. The first four columns report the regression results without any control variable. The next four columns report the regression results with control variables listed in section 4. Standard error is clustered at the individual level. A */**/*** indicates significance at the 10/5/1%.

Table A4: Naive regression for behavioral outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Without any control				With controls			
Outcome	Point estimate	Standard error	R^2	N	Point estimate	Standard error	R^2	N
Panel A: Smoking								
Extensive margin								
Smoker	−0.059 ***	0.004	0.004	71691	−0.023***	0.004	0.253	71548
Frequency								
Smoking days per year	−21.186 ***	1.509	0.004	71691	−8.027***	1.303	0.248	71548
Smoking once a week or more	−0.059 ***	0.004	0.004	71691	−0.022***	0.004	0.250	71548
Smoking everyday	−0.057 ***	0.004	0.003	71691	−0.021***	0.004	0.243	71548
Amount								
Cigarettes per day	−0.915 ***	0.072	0.003	71691	−0.301***	0.063	0.218	71548
Smoking 3 cigarettes or more	−0.057 ***	0.004	0.003	71691	−0.022***	0.004	0.245	71548
Smoking 10 cigarettes or more	−0.049 ***	0.004	0.003	71691	−0.017***	0.003	0.220	71548
Panel B: Drinking								
Extensive margin								
Drinker	0.023 ***	0.005	0.000	71814	0.034***	0.005	0.209	71669
Frequency								
Drinking once a month or more	0.007	0.005	0.000	71814	0.027***	0.005	0.234	71669
Drinking once a week or more	−0.008 *	0.005	0.000	71814	0.010 **	0.004	0.195	71669
Drinking everyday	−0.006 ***	0.002	0.000	71814	0.000	0.002	0.062	71669
Binge drinking once a month or more	−0.017 ***	0.004	0.000	71796	0.006 *	0.004	0.220	71651
Binge drinking once a week or more	−0.009 **	0.004	0.000	71796	0.006 *	0.003	0.138	71651
Binge drinking everyday	−0.003 **	0.001	0.000	71796	0.000	0.001	0.029	71651
Amount								
Drinking 5 cups or more	−0.025 ***	0.005	0.001	71793	0.002	0.004	0.277	71648
Drinking 10 cups or more	−0.007 **	0.003	0.000	71793	0.003	0.002	0.091	71648
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.024 ***	0.004	0.001	71813	0.024***	0.004	0.108	71668
Doing moderate exercise	0.058 ***	0.005	0.002	71812	0.053***	0.005	0.080	71667
Doing walking exercise	0.038 ***	0.004	0.001	71812	0.034***	0.004	0.027	71667
Frequency								
Days of vigorous exercise	3.830 ***	0.875	0.000	71813	4.229***	0.845	0.062	71668
Days of moderate exercise	11.939 ***	1.164	0.002	71812	11.817***	1.143	0.044	71667
Days of walking	7.043 ***	1.396	0.000	71812	8.069***	1.384	0.023	71667
Amount								
30 min vigorous exercise	0.024 ***	0.004	0.001	71813	0.023***	0.004	0.076	71668
30 min moderate exercise	0.055 ***	0.005	0.002	71811	0.050***	0.004	0.050	71666
30 min walking	0.050 ***	0.005	0.002	71812	0.039***	0.005	0.028	71667

Notes: This table presents the naive regression result of health behavior on screening participation. The first four columns report the regression results without any control variable. The next four columns report the regression results with control variables listed in section 4. Standard error is clustered at the individual level. A */**/*** indicates significance at the 10/5/1%.

Table A5: Health screening and outpatient care with controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Outpatient care								
Number of hospital visits								
Total	19.021	18.869	0.027	0.151	1	0.517	1.000	73372
General hospital (bed 100+)	2.473	2.516	0.055**	0.304**	12	0.148	0.400	73372
Local hospital (bed 30-100)	1.448	1.461	0.021	0.117	8	0.145	1.000	73372
Local clinic (bed 0-30)	15.100	14.891	-0.049	-0.269	-2	0.469	1.000	73372
Hospital bill								
Total	340847	336962	-3335	-18372	-5	25276	1.000	73372
General hospital (bed 100+)	83558	85175	1732	9540	11	9004	1.000	73372
Local hospital (bed 30-100)	40985	40606	-685	-3775	-9	9135	1.000	73372
Local clinic (bed 0-30)	215702	210526	-4428	-24393	-11	21881	1.000	73372
Drug expenditures								
Total	113003	112472	526	2895	3	3451	1.000	73372
General hospital (bed 100+)	41091	41341	807*	4445*	11	2679	0.600	73372
Local hospital (bed 30-100)	8168	8093	-38.984	-215	-3	918	1.000	73372
Local clinic (bed 0-30)	61677	60975	-263	-1448	-2	2189	1.000	73372
Panel B: Outpatient care for a new illness								
First hospital visit for a new illness								
Total	3.696	3.758	0.074***	0.408***	11	0.103	0.000	73372
General hospital (bed 100+)	0.353	0.371	0.019***	0.105***	30	0.031	0.000	73372
Local hospital (bed 30-100)	0.286	0.293	0.009*	0.047*	16	0.026	0.600	73372
Local clinic (bed 0-30)	2.847	2.884	0.042***	0.234***	8	0.089	0.200	73372
First hospital bill for a new illness								
Total	89527	93896	4273**	23544**	26	10717	0.400	73372
General hospital (bed 100+)	20347	21728	1309**	7214**	35	3678	0.600	73372
Local hospital (bed 30-100)	13010	13988	969	5339	41	4130	0.600	73372
Local clinic (bed 0-30)	55986	57996	1989	10958	20	9204	1.000	73372
First drug expenditures for a new illness								
Total	10982	11351	398***	2192***	20	651	0.000	73372
General hospital (bed 100+)	1924	2011	100	552	29	407	0.600	73372
Local hospital (bed 30-100)	1000	1035	43.102	237	24	186	0.600	73372
Local clinic (bed 0-30)	7947	8193	253***	1393***	18	453	0.000	73372

Notes: This table presents the effect of health screening on outpatient care usage and medical expenditure. Both the number of hospital visits and medical expenditures exclude the visits for screening and medical expenditures incurred for screening. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include control variables listed in section 4. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 24 hypotheses tested with 5 bootstrap replications.

Table A6: Health screening and outpatient care with individual FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Outpatient care								
Number of hospital visits								
Total	19.021	18.869	0.104	0.563	3	0.479	0.600	70913
General hospital (bed 100+)	2.473	2.516	0.055**	0.295**	12	0.149	0.400	70913
Local hospital (bed 30-100)	1.448	1.461	0.056**	0.302**	21	0.141	0.400	70913
Local clinic (bed 0-30)	15.100	14.891	-0.006	-0.034	-0	0.434	1.000	70913
Hospital bill								
Total	340847	336962	-1278	-6906	-2	25193	1.000	70913
General hospital (bed 100+)	83558	85175	1457	7874	9	8992	1.000	70913
Local hospital (bed 30-100)	40985	40606	333	1801	4	8838	1.000	70913
Local clinic (bed 0-30)	215702	210526	-3092	-16711	-8	22014	1.000	70913
Drug expenditures								
Total	113003	112472	808	4367	4	2851	0.600	70913
General hospital (bed 100+)	41091	41341	795*	4295*	10	2335	0.400	70913
Local hospital (bed 30-100)	8168	8093	118	637	8	837	1.000	70913
Local clinic (bed 0-30)	61677	60975	-104	-559	-1	1681	1.000	70913
Panel B: Outpatient care for a new illness								
First hospital visit for a new illness								
Total	3.696	3.758	0.068***	0.365***	10	0.099	0.000	70913
General hospital (bed 100+)	0.353	0.371	0.022***	0.117***	33	0.031	0.000	70913
Local hospital (bed 30-100)	0.286	0.293	0.010**	0.055**	19	0.025	0.400	70913
Local clinic (bed 0-30)	2.847	2.884	0.036**	0.194**	7	0.086	0.400	70913
First hospital bill for a new illness								
Total	89527	93896	4405**	23806**	27	10544	0.400	70913
General hospital (bed 100+)	20347	21728	1312*	7092*	35	3735	0.400	70913
Local hospital (bed 30-100)	13010	13988	995	5375	41	3857	0.600	70913
Local clinic (bed 0-30)	55986	57996	2086	11274	20	9144	0.600	70913
First drug expenditures for a new illness								
Total	10982	11351	376***	2034***	19	631	0.200	70913
General hospital (bed 100+)	1924	2011	106	572	30	388	0.600	70913
Local hospital (bed 30-100)	1000	1035	42.803	231	23	186	0.600	70913
Local clinic (bed 0-30)	7947	8193	225***	1217***	15	444	0.200	70913

Notes: This table presents the effect of health screening on outpatient care usage and medical expenditure. Both the number of hospital visits and medical expenditures exclude the visits for screening and medical expenditures incurred for screening. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include individual and year fixed effects. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 24 hypotheses tested with 5 bootstrap replications.

Table A7: Health screening and inpatient/emergency care with controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Inpatient care								
Number of hospital visits								
Total	0.236	0.237	0.003	0.016	7	0.027	1.000	73372
General hospital (bed 100+)	0.125	0.123	-0.001	-0.007	-5	0.020	1.000	73372
Local hospital (bed 30-100)	0.076	0.080	0.005*	0.026*	34	0.013	0.400	73372
Local clinic (bed 0-30)	0.036	0.035	-0.001	-0.003	-8	0.009	1.000	73372
Hospital bill								
Total	209413	208121	1486	8184	4	40563	1.000	73372
General hospital (bed 100+)	128569	126258	198	1092	1	33319	1.000	73372
Local hospital (bed 30-100)	66662	68826	2418	13323	20	21827	1.000	73372
Local clinic (bed 0-30)	14113	13021	-1067	-5876	-42	6148	1.000	73372
Drug expenditures								
Total	85.934	90.767	3.868	21.310	25	92.246	1.000	73372
General hospital (bed 100+)	43.834	58.667	15.631	86.114	196	81.963	1.000	73372
Local hospital (bed 30-100)	17.758	14.806	-3.704	-20.406	-115	33.142	1.000	73372
Local clinic (bed 0-30)	24.343	17.294	-8.059*	-44.398*	-182	26.186	0.600	73372
Panel B: Emergency care								
Number of hospital visits								
Total	0.127	0.121	-0.006*	-0.031*	-24	0.018	0.400	73372
General hospital (bed 100+)	0.094	0.088	-0.006**	-0.034**	-36	0.015	0.200	73372
Local hospital (bed 30-100)	0.032	0.032	0.000	0.001	4	0.011	1.000	73372
Local clinic (bed 0-30)	0.001	0.001	0.000	0.001	134	0.002	1.000	73372
Hospital bill								
Total	7444	7372	0.238	1.311	0	2406	1.000	73372
General hospital (bed 100+)	6589	6559	38.548	212	3	2349	1.000	73372
Local hospital (bed 30-100)	814	791	-22.228	-122	-15	478	1.000	73372
Local clinic (bed 0-30)	40.059	21.429	-16.081	-88.598	-221	108	1.000	73372
Drug expenditures								
Total	50.376	54.718	6.752	37.201	74	54.211	1.000	73372
General hospital (bed 100+)	26.558	31.530	6.852	37.752	142	49.829	1.000	73372
Local hospital (bed 30-100)	23.776	22.940	-0.342	-1.886	-8	21.244	1.000	73372
Local clinic (bed 0-30)	0.041	0.248	0.242	1.335	3250	1.374	1.000	73372

Notes: This table presents the effect of health screening on inpatient and emergency care usage and medical expenditures. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include control variables listed in section 4. A */**/*** indicates significance at the 10/5/1%.
[a] Westfall-Young adjusted p-values account for 12 hypotheses tested in each inpatient and emergency care domain with 5 bootstrap replications.

Table A8: Health screening and inpatient/emergency care with individual FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Inpatient care								
Number of hospital visits								
Total	0.236	0.237	0.007	0.040	17	0.027	0.800	70913
General hospital (bed 100+)	0.125	0.123	0.001	0.005	4	0.020	1.000	70913
Local hospital (bed 30-100)	0.076	0.080	0.007***	0.036***	47	0.013	0.000	70913
Local clinic (bed 0-30)	0.036	0.035	-0.000	-0.001	-3	0.009	1.000	70913
Hospital bill								
Total	209413	208121	3744	20235	10	39703	1.000	70913
General hospital (bed 100+)	128569	126258	1029	5561	4	32339	1.000	70913
Local hospital (bed 30-100)	66662	68826	3585	19372	29	21915	1.000	70913
Local clinic (bed 0-30)	14113	13021	-796	-4302	-30	6159	1.000	70913
Drug expenditures								
Total	85.934	90.767	10.171	54.963	64	92.973	1.000	70913
General hospital (bed 100+)	43.834	58.667	17.191	92.903	212	83.552	1.000	70913
Local hospital (bed 30-100)	17.758	14.806	-2.155	-11.646	-66	32.209	1.000	70913
Local clinic (bed 0-30)	24.343	17.294	-4.866	-26.294	-108	24.829	1.000	70913
Panel B: Emergency care								
Number of hospital visits								
Total	0.127	0.121	-0.005	-0.026	-20	0.018	0.600	70913
General hospital (bed 100+)	0.094	0.088	-0.006**	-0.034**	-37	0.015	0.200	70913
Local hospital (bed 30-100)	0.032	0.032	0.001	0.007	21	0.010	1.000	70913
Local clinic (bed 0-30)	0.001	0.001	0.000	0.002	156	0.002	1.000	70913
Hospital bill								
Total	7444	7372	-31.493	-170	-2	2442	1.000	70913
General hospital (bed 100+)	6589	6559	-33.152	-179	-3	2393	1.000	70913
Local hospital (bed 30-100)	814	791	10.764	58.167	7	442	1.000	70913
Local clinic (bed 0-30)	40.059	21.429	-9.104	-49.200	-123	105	1.000	70913
Drug expenditures								
Total	50.376	54.718	10.194	55.088	109	63.375	1.000	70913
General hospital (bed 100+)	26.558	31.530	9.457	51.104	192	59.669	1.000	70913
Local hospital (bed 30-100)	23.776	22.940	0.478	2.584	11	21.244	1.000	70913
Local clinic (bed 0-30)	0.041	0.248	0.259	1.401	3409	1.451	1.000	70913

Notes: This table presents the effect of health screening on inpatient and emergency care usage and medical expenditures. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include individual and year fixed effects. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 12 hypotheses tested in each inpatient and emergency care domain with 5 bootstrap replications.

Table A9: Health screening and behavior with controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Smoking								
Extensive margin								
Smoker	0.193	0.190	−0.002	−0.010	-5	0.007	0.600	71548
Frequency								
Smoking days per year	68.018	67.159	−0.570	−3.075	-5	2.638	1.000	71548
Smoking once a week or more	0.190	0.188	−0.002	−0.009	-4	0.007	1.000	71548
Smoking everyday	0.184	0.182	−0.002	−0.009	-5	0.007	1.000	71548
Amount								
Cigarettes per day	2.838	2.816	−0.013	−0.070	-2	0.136	1.000	71548
Smoking 3 cigarettes or more	0.184	0.182	−0.002	−0.010	-5	0.007	0.600	71548
Smoking 10 cigarettes or more	0.154	0.151	−0.002	−0.012	-8	0.008	0.400	71548
Standardized treatment effect[b]								
Smoking index				−0.023		0.018		
Panel B: Drinking								
Extensive margin								
Drinker	0.633	0.638	0.003	0.017	3	0.011	0.600	71669
Frequency								
Drinking once a month or more	0.497	0.494	−0.004*	−0.020 *	-4	0.012	0.600	71669
Drinking once a week or more	0.287	0.283	−0.003	−0.019	-6	0.011	0.600	71669
Drinking everyday	0.057	0.054	−0.002*	−0.013 *	-23	0.007	0.400	71669
Binge drinking once a month or more	0.217	0.215	−0.002	−0.008	-4	0.011	0.800	71651
Binge drinking once a week or more	0.130	0.126	−0.003	−0.014	-11	0.010	0.600	71651
Binge drinking everyday	0.022	0.020	−0.001	−0.006	-29	0.005	0.600	71651
Amount								
Drinking 5 cups or more	0.255	0.253	−0.002	−0.010	-4	0.011	0.800	71648
Drinking 10 cups or more	0.070	0.072	0.002	0.009	13	0.008	0.800	71648
Standardized treatment effect[b]								
Drinking index				−0.029**		0.013		
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.215	0.214	−0.001	−0.007	-3	0.014	1.000	71668
Doing moderate exercise	0.351	0.350	−0.001	−0.007	-2	0.016	1.000	71667
Doing walking exercise	0.772	0.777	0.003	0.018	2	0.015	0.800	71667
Frequency								
Days of vigorous exercise	36.886	36.859	−0.050	−0.267	-1	2.893	1.000	71668
Days of moderate exercise	71.154	71.085	−0.115	−0.622	-1	3.924	1.000	71667
Days of walking	206	208	1.720*	9.291 *	5	4.985	0.200	71667
Amount								
30 min vigorous exercise	0.166	0.167	0.001	0.007	4	0.013	1.000	71668
30 min moderate exercise	0.256	0.259	0.003	0.017	6	0.015	0.800	71666
30 min walking	0.416	0.412	−0.004	−0.024	-6	0.017	0.600	71667
Standardized treatment effect[b]								
Exercise index				0.008		0.020		

Notes: This table presents the effect of health screening on health behaviors. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include control variables listed in section 4. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for the number of hypotheses tested in each behavior domain with 5 bootstrap replications.

[b] Standardized treatment effect gives equal weight to all the outcomes in each behavior domain. For drinking, the drinker variable's sign is reversed before calculation.

Table A10: Health screening and behavior with individual FE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Smoking								
Extensive margin								
Smoker	0.193	0.190	−0.000	−0.001	-0	0.006	1.000	69054
Frequency								
Smoking days per year	68.018	67.159	0.064	0.338	0	2.077	1.000	69054
Smoking once a week or more	0.190	0.188	0.000	0.001	0	0.006	1.000	69054
Smoking everyday	0.184	0.182	0.000	0.001	1	0.006	1.000	69054
Amount								
Cigarettes per day	2.838	2.816	0.021	0.109	4	0.110	0.800	69054
Smoking 3 cigarettes or more	0.184	0.182	−0.000	−0.001	-1	0.006	1.000	69054
Smoking 10 cigarettes or more	0.154	0.151	−0.000	−0.000	-0	0.006	1.000	69054
Standardized treatment effect[b]								
Smoking index				0.003		0.014		
Panel B: Drinking								
Extensive margin								
Drinker	0.633	0.638	0.003	0.014	2	0.010	0.600	69179
Frequency								
Drinking once a month or more	0.497	0.494	−0.003	−0.016	-3	0.011	0.600	69179
Drinking once a week or more	0.287	0.283	−0.002	−0.010	-4	0.011	0.800	69179
Drinking everyday	0.057	0.054	−0.002*	−0.012*	-21	0.006	0.200	69179
Binge drinking once a month or more	0.217	0.215	−0.001	−0.008	-3	0.011	1.000	69162
Binge drinking once a week or more	0.130	0.126	−0.002	−0.010	-7	0.010	0.800	69162
Binge drinking everyday	0.022	0.020	−0.001	−0.006	-25	0.005	0.600	69162
Amount								
Drinking 5 cups or more	0.255	0.253	−0.001	−0.006	-2	0.010	1.000	69158
Drinking 10 cups or more	0.070	0.072	0.001	0.007	10	0.008	0.800	69158
Standardized treatment effect[b]								
Drinking index				−0.023*		0.012		
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.215	0.214	0.000	0.002	1	0.014	1.000	69179
Doing moderate exercise	0.351	0.350	0.001	0.003	1	0.016	1.000	69178
Doing walking exercise	0.772	0.777	0.003	0.014	2	0.015	0.800	69178
Frequency								
Days of vigorous exercise	36.886	36.859	0.089	0.473	1	2.888	1.000	69179
Days of moderate exercise	71.154	71.085	0.215	1.137	2	3.889	1.000	69178
Days of walking	206	208	1.437	7.603	4	4.947	0.200	69178
Amount								
30 min vigorous exercise	0.166	0.167	0.002	0.011	7	0.013	0.800	69179
30 min moderate exercise	0.256	0.259	0.004	0.019	8	0.015	0.200	69177
30 min walking	0.416	0.412	−0.005	−0.028	-7	0.017	0.200	69178
Standardized treatment effect[b]								
Exercise index				0.015		0.020		

Notes: This table presents the effect of health screening on health behaviors. The first and second columns report the mean of outcome variables in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports percentage change calculated as LATE divided by control group mean. The sixth column reports standard errors of local average treatment effect clustered at individual level. All regressions include individual and year fixed effects. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for the number of hypotheses tested in each behavior domain with 5 bootstrap replications.

[b] Standardized treatment effect gives equal weight to all the outcomes in each behavior domain. For drinking, the drinker variable's sign is reversed before calculation.

Table A11: Complier analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Outcome var: Health screening take-up												
	Total	Male	Female	Non working	Working	Young	Old	Single	Married	Low Edu	High edu	Poor	Rich
Eligible	0.180***	0.138***	0.214***	0.224***	0.142***	0.139***	0.220***	0.123***	0.201***	0.208***	0.128***	0.186***	0.175***
	(0.003)	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
cons	0.100***	0.104***	0.096***	0.085***	0.113***	0.110***	0.090***	0.069***	0.111***	0.085***	0.128***	0.074***	0.122***
	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
N	73,535	32,688	40,847	34,670	38,857	35,887	37,648	19,348	54,187	48,146	25,389	33,897	39,494
R^2	0.053	0.033	0.070	0.079	0.033	0.033	0.076	0.033	0.060	0.070	0.026	0.062	0.046

Notes: The table presents the first stage regression results for the total sample and subsamples. Outcome variable is the indicator variable for getting screening. Independent variable is eligibility for free screening. Subsamples young and poor indicate a group of individuals with age and household income below median. Subsample Low Edu refers to individuals with at most 12 years of schooling, and hence, at most high school graduate. All specifications do not include any control. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1%.

Table A12: Robustness: First stage using cancer screening

	(1)	(2)	(3)
		Outcome var: Cancer screening take-up	
Eligible	0.184***	0.185***	0.188***
	(0.003)	(0.003)	(0.003)
N	73,535	73,372	70,913
Adj R^2	0.060	0.087	0.164
Controls		Y	
Year FE		Y	Y
Individual FE			Y

Notes: This table presents the first stage regression result using cancer screening. Outcome variable is the take-up of cancer screening. Independent variable is eligibility for biannual free health screening. Control variables are listed in section 4. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A13: Robustness: Outpatient care usage and medical expenditures using cancer screening

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Outpatient care								
Number of hospital visits								
Total	19.021	18.869	-0.153	-0.832	-4	0.562	0.400	73535
General hospital (bed 100+)	2.473	2.516	0.043	0.234	9	0.152	0.400	73535
Local hospital (bed 30-100)	1.448	1.461	0.013	0.070	5	0.141	1.000	73535
Local clinic (bed 0-30)	15.100	14.891	-0.209**	-1.136**	-8	0.514	0.200	73535
Hospital bill								
Total	340847	336962	-3885	-21150	-6	25091	1.000	73535
General hospital (bed 100+)	83558	85175	1617	8803	11	9069	0.800	73535
Local hospital (bed 30-100)	40985	40606	-378	-2059	-5	8891	1.000	73535
Local clinic (bed 0-30)	215702	210526	-5176	-28178	-13	21677	0.600	73535
Drug expenditures								
Total	113003	112472	-531	-2890	-3	3748	1.000	73535
General hospital (bed 100+)	41091	41341	249	1357	3	2825	1.000	73535
Local hospital (bed 30-100)	8168	8093	-75.018	-408	-5	948	1.000	73535
Local clinic (bed 0-30)	61677	60975	-702*	-3823*	-6	2301	0.400	73535
Panel B: Outpatient care for a new illness								
First hospital visit for a new illness								
Total	3.696	3.758	0.063***	0.342***	9	0.106	0.000	73535
General hospital (bed 100+)	0.353	0.371	0.018***	0.098***	28	0.031	0.000	73535
Local hospital (bed 30-100)	0.286	0.293	0.007	0.038	13	0.025	0.400	73535
Local clinic (bed 0-30)	2.847	2.884	0.037**	0.199**	7	0.091	0.200	73535
First hospital bill for a new illness								
Total	89527	93896	4369**	23784**	27	10507	0.200	73535
General hospital (bed 100+)	20347	21728	1380**	7514**	37	3619	0.200	73535
Local hospital (bed 30-100)	13010	13988	978	5324	41	4003	0.600	73535
Local clinic (bed 0-30)	55986	57996	2010	10944	20	9019	0.600	73535
First drug expenditures for a new illness								
Total	10982	11351	369***	2010***	18	649	0.000	73535
General hospital (bed 100+)	1924	2011	87.830	478	25	399	0.600	73535
Local hospital (bed 30-100)	1000	1035	35.198	192	19	184	0.800	73535
Local clinic (bed 0-30)	7947	8193	246***	1338***	17	457	0.000	73535

Notes: This table presents the effect of cancer screening on outpatient care usage and medical expenditure. Both the number of hospital visits and medical expenditure exclude the visits for screening and medical expenditure incurred for screening. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 24 hypotheses tested with 5 bootstrap replications.

Table A14: Robustness: Inpatient/emergency care usage and medical expenditures using cancer screening

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Inpatient care								
Number of hospital visits								
Total	0.236	0.237	0.001	0.004	2	0.027	1.000	73535
General hospital (bed 100+)	0.125	0.123	-0.002	-0.012	-9	0.021	1.000	73535
Local hospital (bed 30-100)	0.076	0.080	0.004	0.020	27	0.013	0.800	73535
Local clinic (bed 0-30)	0.036	0.035	-0.001	-0.005	-14	0.009	1.000	73535
Hospital bill								
Total	209413	208121	-1292	-7170	-3	40095	1.000	73535
General hospital (bed 100+)	128569	126258	-2312	-12824	-10	32903	1.000	73535
Local hospital (bed 30-100)	66662	68826	2164	12007	18	21596	1.000	73535
Local clinic (bed 0-30)	14113	13021	-1092	-6061	-43	6255	1.000	73535
Drug expenditures								
Total	85.934	90.767	4.833	26.811	31	90.724	1.000	73535
General hospital (bed 100+)	43.834	58.667	14.833	82.287	188	80.775	1.000	73535
Local hospital (bed 30-100)	17.758	14.806	-2.952	-16.374	-92	31.448	1.000	73535
Local clinic (bed 0-30)	24.343	17.294	-7.048	-39.101	-161	26.565	0.800	73535
Panel B: Emergency care								
Number of hospital visits								
Total	0.127	0.121	-0.006*	-0.032*	-25	0.018	0.400	73535
General hospital (bed 100+)	0.094	0.088	-0.006**	-0.035**	-37	0.015	0.200	73535
Local hospital (bed 30-100)	0.032	0.032	0.000	0.002	5	0.009	1.000	73535
Local clinic (bed 0-30)	0.001	0.001	0.000	0.001	125	0.002	1.000	73535
Hospital bill								
Total	7444	7372	-71.574	-397	-5	2352	1.000	73535
General hospital (bed 100+)	6589	6559	-29.872	-166	-3	2295	1.000	73535
Local hospital (bed 30-100)	814	791	-23.072	-128	-16	472	1.000	73535
Local clinic (bed 0-30)	40.059	21.429	-18.630	-103	-258	105	1.000	73535
Drug expenditures								
Total	50.376	54.718	4.343	24.091	48	49.191	1.000	73535
General hospital (bed 100+)	26.558	31.530	4.972	27.581	104	44.912	1.000	73535
Local hospital (bed 30-100)	23.776	22.940	-0.836	-4.640	-20	20.002	1.000	73535
Local clinic (bed 0-30)	0.041	0.248	0.207	1.151	2801	1.233	1.000	73535

Notes: This table presents the effect of cancer screening on inpatient and emergency care usage and medical expenditure. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.
[a] Westfall-Young adjusted p-values account for 12 hypotheses tested in each impatient and emergency care domain with 5 bootstrap replications.

Table A15: Robustness: Cancer screening and behavior

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Smoking								
Extensive margin								
Smoker	0.193	0.190	−0.003	−0.014	-7	0.008	0.400	71691
Frequency								
Smoking days per year	68.018	67.159	−0.860	−4.584	-7	2.999	0.400	71691
Smoking once a week or more	0.190	0.188	−0.002	−0.013	-7	0.008	0.400	71691
Smoking everyday	0.184	0.182	−0.002	−0.013	-7	0.008	0.400	71691
Amount								
Cigarettes per day	2.838	2.816	−0.023	−0.120	-4	0.152	0.600	71691
Smoking 3 cigarettes or more	0.184	0.182	−0.003 *	−0.014 *	-7	0.008	0.400	71691
Smoking 10 cigarettes or more	0.154	0.151	−0.003 *	−0.015 *	-10	0.008	0.400	71691
Standardized treatment effect[b]								
Smoking index				−0.033		0.020		
Panel B: Drinking								
Extensive margin								
Drinker	0.633	0.638	0.005**	0.025**	4	0.012	0.400	71814
Frequency								
Drinking once a month or more	0.497	0.494	−0.003	−0.017	-3	0.013	0.800	71814
Drinking once a week or more	0.287	0.283	−0.004 *	−0.021 *	-7	0.012	0.800	71814
Drinking everyday	0.057	0.054	−0.003**	−0.017**	-30	0.007	0.000	71814
Binge drinking once a month or more	0.217	0.215	−0.002	−0.009	-4	0.012	0.800	71796
Binge drinking once a week or more	0.130	0.126	−0.003	−0.017	-13	0.010	0.800	71796
Binge drinking everyday	0.022	0.020	−0.002 *	−0.009 *	-39	0.005	0.600	71796
Amount								
Drinking 5 cups or more	0.255	0.253	−0.002	−0.009	-4	0.012	0.800	71793
Drinking 10 cups or more	0.070	0.072	0.001	0.008	11	0.008	0.800	71793
Standardized treatment effect[b]								
Drinking index				−0.036**		0.014		
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.215	0.214	−0.001	−0.005	-2	0.014	1.000	71813
Doing moderate exercise	0.351	0.350	−0.001	−0.005	-1	0.016	1.000	71812
Doing walking exercise	0.772	0.777	0.005 *	0.025 *	3	0.015	0.200	71812
Frequency								
Days of vigorous exercise	36.886	36.859	−0.027	−0.144	-0	2.904	1.000	71813
Days of moderate exercise	71.154	71.085	−0.069	−0.368	-1	3.925	1.000	71812
Days of walking	206	208	1.951**	10.416**	5	4.942	0.200	71812
Amount								
30 min vigorous exercise	0.166	0.167	0.001	0.006	4	0.013	1.000	71813
30 min moderate exercise	0.256	0.259	0.003	0.019	7	0.015	0.600	71811
30 min walking	0.416	0.412	−0.004	−0.020	-5	0.017	0.600	71812
Standardized treatment effect[b]								
Exercise index				0.013		0.020		

Notes: This table presents the effect of cancer screening on health behaviors. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of cancer screening participation using biannual free screening eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.

[a] Standardized treatment effect gives equal weight to all the outcomes in each behavior domain. For drinking, the drinker variable's sign is reversed before calculation.

[b] Westfall-Young adjusted p-values account for the number of hypotheses tested in each behavior domain with 5 bootstrap replications.

Table A16: Robustness: First stage including non-office workers

	(1)	(2)	(3)
	Outcome var: Health screening take-up		
Eligible	0.170***	0.171***	0.175***
	(0.003)	(0.003)	(0.003)
N	104,130	103,921	101,935
Adj R^2	0.046	0.071	0.152
Controls		Y	
Year FE		Y	Y
Individual FE			Y

Notes: This table presents the first stage regression result with the sample including non-office workers who are subject to annual health screening. Outcome variable is the take-up of general health screening. Independent variable is eligibility for biannual free health screening. Control variables are listed in section 4. Standard errors are clustered at individual level and reported in parentheses. A */**/*** indicates significance at the 10/5/1% levels.

Table A17: Robustness: Outpatient care usage and medical expenditure including non-office workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Outpatient care								
Number of hospital visits								
Total	17.795	17.701	-0.094	-0.551	-3	0.449	1.000	104130
General hospital (bed 100+)	2.173	2.204	0.031	0.181	8	0.123	1.000	104130
Local hospital (bed 30-100)	1.375	1.383	0.008	0.048	4	0.120	1.000	104130
Local clinic (bed 0-30)	14.247	14.114	-0.133*	-0.780*	-5	0.409	0.800	104130
Hospital bill								
Total	335470	330357	-5113	-30005	-9	22968	1.000	104130
General hospital (bed 100+)	76187	78025	1838	10786	14	7758	1.000	104130
Local hospital (bed 30-100)	41147	41138	-8.749	-51.348	-0	8500	1.000	104130
Local clinic (bed 0-30)	217438	210473	-6965**	-40878**	-19	19883	0.600	104130
Drug expenditures								
Total	107173	107237	64.200	377	0	3011	1.000	104130
General hospital (bed 100+)	36955	36978	23.276	137	0	2276	1.000	104130
Local hospital (bed 30-100)	7913	8008	95.182	559	7	781	1.000	104130
Local clinic (bed 0-30)	60154	60088	-66.466	-390	-1	1833	1.000	104130
Panel B: Outpatient care for a new illness								
First hospital visit for a new illness								
Total	3.635	3.679	0.044***	0.256***	7	0.091	0.200	104130
General hospital (bed 100+)	0.331	0.345	0.014***	0.083***	25	0.027	0.000	104130
Local hospital (bed 30-100)	0.285	0.294	0.008**	0.049**	17	0.022	0.400	104130
Local clinic (bed 0-30)	2.795	2.816	0.021	0.126	4	0.078	1.000	104130
First hospital bill for a new illness								
Total	89348	92065	2717*	15948*	18	9537	0.800	104130
General hospital (bed 100+)	19551	20942	1391***	8166***	42	3143	0.400	104130
Local hospital (bed 30-100)	13229	14372	1143*	6706*	51	3849	0.800	104130
Local clinic (bed 0-30)	56367	56548	180	1059	2	8095	1.000	104130
First drug expenditures for a new illness								
Total	10999	11363	365***	2141***	19	577	0.000	104130
General hospital (bed 100+)	1854	1938	83.959	493	27	352	1.000	104130
Local hospital (bed 30-100)	1026	1086	60.325**	354**	35	166	0.600	104130
Local clinic (bed 0-30)	8003	8227	224***	1313***	16	411	0.000	104130

Notes: This table presents the effect of health screening on health care usage and medical expenditures. The sample includes non-office workers who are subject to annual health screening. Both the number of hospital visits and medical expenditure exclude the visits for screening and medical expenditure incurred for screening. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of health screening participation using biannual health screening eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 24 hypotheses tested with 5 bootstrap replications.

Table A18: Robustness: Inpatient/emergency care usage and medical expenditures including non-office workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Inpatient care								
Number of hospital visits								
Total	0.214	0.216	0.002	0.011	5	0.022	1.000	104130
General hospital (bed 100+)	0.112	0.111	-0.001	-0.005	-4	0.017	1.000	104130
Local hospital (bed 30-100)	0.068	0.072	0.003*	0.019*	27	0.011	0.600	104130
Local clinic (bed 0-30)	0.034	0.034	-0.000	-0.003	-8	0.008	1.000	104130
Hospital bill								
Total	194203	190925	-3278	-19237	-10	33415	1.000	104130
General hospital (bed 100+)	118103	115114	-2989	-17539	-15	27389	1.000	104130
Local hospital (bed 30-100)	62088	62966	878	5151	8	18224	1.000	104130
Local clinic (bed 0-30)	13963	12833	-1130	-6632	-47	5696	0.600	104130
Drug expenditures								
Total	81.820	76.438	-5.382	-31.584	-39	74.213	1.000	104130
General hospital (bed 100+)	44.829	48.456	3.627	21.285	47	66.546	1.000	104130
Local hospital (bed 30-100)	14.613	12.590	-2.024	-11.876	-81	24.409	1.000	104130
Local clinic (bed 0-30)	22.377	15.392	-6.985*	-40.993*	-183	21.831	0.600	104130
Panel B: Emergency care								
Number of hospital visits								
Total	0.118	0.116	-0.003	-0.015	-13	0.015	1.000	104130
General hospital (bed 100+)	0.086	0.082	-0.004*	-0.022*	-25	0.012	0.800	104130
Local hospital (bed 30-100)	0.032	0.032	0.001	0.004	13	0.008	1.000	104130
Local clinic (bed 0-30)	0.001	0.001	0.000*	0.002*	271	0.001	0.800	104130
Hospital bill								
Total	7059	6974	-84.988	-499	-7	1993	1.000	104130
General hospital (bed 100+)	6217	6127	-89.897	-528	-8	1940	1.000	104130
Local hospital (bed 30-100)	813	820	6.595	38.706	5	406	1.000	104130
Local clinic (bed 0-30)	28.226	26.539	-1.686	-9.897	-35	86.674	1.000	104130
Drug expenditures								
Total	47.099	52.606	5.507	32.317	69	38.733	1.000	104130
General hospital (bed 100+)	23.748	29.232	5.484	32.185	136	34.883	1.000	104130
Local hospital (bed 30-100)	23.322	23.060	-0.262	-1.537	-7	16.834	1.000	104130
Local clinic (bed 0-30)	0.029	0.314	0.285	1.670	5769	1.016	0.800	104130

Notes: This table presents the effect of health screening on inpatient and emergency care usage and medical expenditures. The sample includes non-office workers who are subject to annual health screening. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of free screening eligibility. The fourth column (LATE) reports the local average treatment effect of screening participation using eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.

[a] Westfall-Young adjusted p-values account for 12 hypotheses tested in each impatient and emergency care domain with 5 bootstrap replications.

Table A19: Robustness: Health screening and behavior including non-office workers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	Control group	Treatment group	ITT	LATE	Percentage change	Standard error	Adjusted p-value[a]	Obs
Panel A: Smoking								
Extensive margin								
Smoker	0.215	0.215	0.000	0.000	0	0.007	1.000	101678
Frequency								
Smoking days per year	76.092	76.338	0.247	1.418	2	2.562	1.000	101678
Smoking once a week or more	0.213	0.213	0.000	0.002	1	0.007	1.000	101678
Smoking everyday	0.206	0.207	0.001	0.005	3	0.007	0.800	101678
Amount								
Cigarettes per day	3.217	3.223	0.005	0.031	1	0.133	1.000	101678
Smoking 3 cigarettes or more	0.207	0.207	0.000	0.000	0	0.007	1.000	101678
Smoking 10 cigarettes or more	0.175	0.175	−0.000	−0.002	-1	0.007	1.000	101678
Standardized treatment effect[b]								
Smoking index				0.004		0.017		
Panel B: Drinking								
Extensive margin								
Drinker	0.664	0.668	0.005***	0.027***	4	0.010	0.000	101829
Frequency								
Drinking once a month or more	0.529	0.528	−0.002	−0.009	-2	0.011	1.000	101829
Drinking once a week or more	0.313	0.312	−0.000	−0.003	-1	0.011	1.000	101829
Drinking everyday	0.065	0.062	−0.003 **	−0.016 **	-25	0.006	0.000	101829
Binge drinking once a month or more	0.242	0.241	−0.001	−0.007	-3	0.011	1.000	101807
Binge drinking once a week or more	0.144	0.143	−0.001	−0.008	-5	0.010	1.000	101807
Binge drinking everyday	0.025	0.023	−0.001 *	−0.008 *	-31	0.005	0.400	101807
Amount								
Drinking 5 cups or more	0.283	0.283	−0.000	−0.000	-0	0.010	1.000	101804
Drinking 10 cups or more	0.080	0.081	0.001	0.004	5	0.008	1.000	101804
Standardized treatment effect[b]								
Drinking index				−0.025 **		0.012		
Panel C: Exercise								
Extensive margin								
Doing vigorous exercise	0.227	0.226	−0.002	−0.009	-4	0.013	1.000	101825
Doing moderate exercise	0.374	0.372	−0.002	−0.011	-3	0.015	1.000	101824
Doing walking exercise	0.776	0.779	0.003	0.018	2	0.013	0.800	101824
Frequency								
Days of vigorous exercise	39.736	39.464	−0.272	−1.567	-4	2.788	1.000	101825
Days of moderate exercise	78.186	77.623	−0.563	−3.240	-4	3.748	1.000	101824
Days of walking	210	211	1.475 *	8.490 *	4	4.443	0.400	101824
Amount								
30 min vigorous exercise	0.174	0.174	−0.000	−0.002	-1	0.012	1.000	101824
30 min moderate exercise	0.277	0.277	0.000	0.002	1	0.014	1.000	101823
30 min walking	0.417	0.414	−0.003	−0.015	-4	0.016	1.000	101824
Standardized treatment effect[b]								
Exercise index				−0.002		0.019		

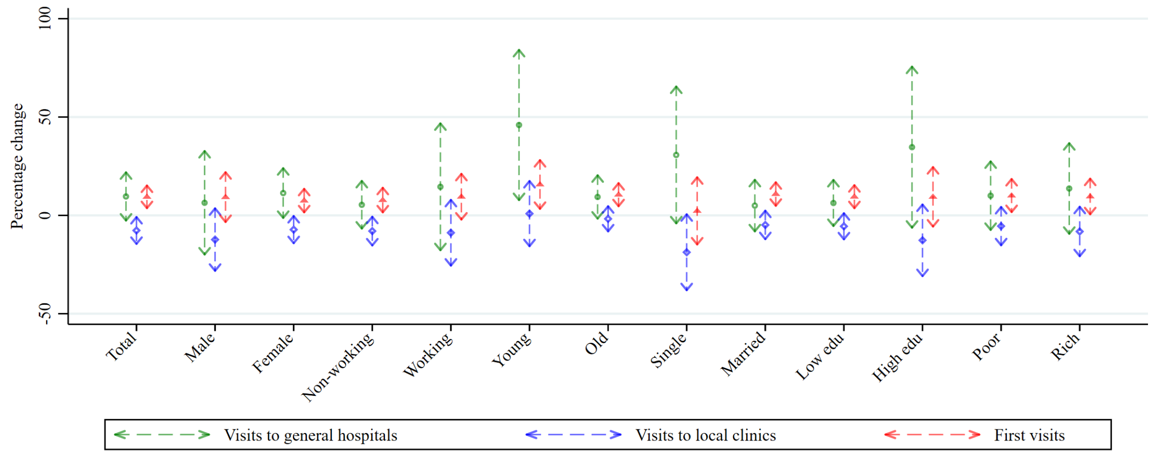
Notes: This table presents the effect of health screening on health behaviors. The sample includes non-office workers who are subject to annual health screening. The first and second columns report the mean of an outcome variable in the control and treatment group. The third column (ITT) reports the intention-to-treat effect of biannual free screening eligibility. The fourth column (LATE) reports the local average treatment effect of health screening participation using biannual health screening eligibility as an instrument. The fifth column reports standard errors of local average treatment effect clustered at individual level. All regressions do not include any control variable. Percentage change is calculated as LATE divided by control group mean. A */**/*** indicates significance at the 10/5/1%.

[a] Standardized treatment effect gives equal weight to all the outcomes in each behavior domain. For drinking, the drinker variable's sign is reversed before calculation.

[b] Westfall-Young adjusted p-values account for the number of hypotheses tested in each behavior domain with 5 bootstrap replications.

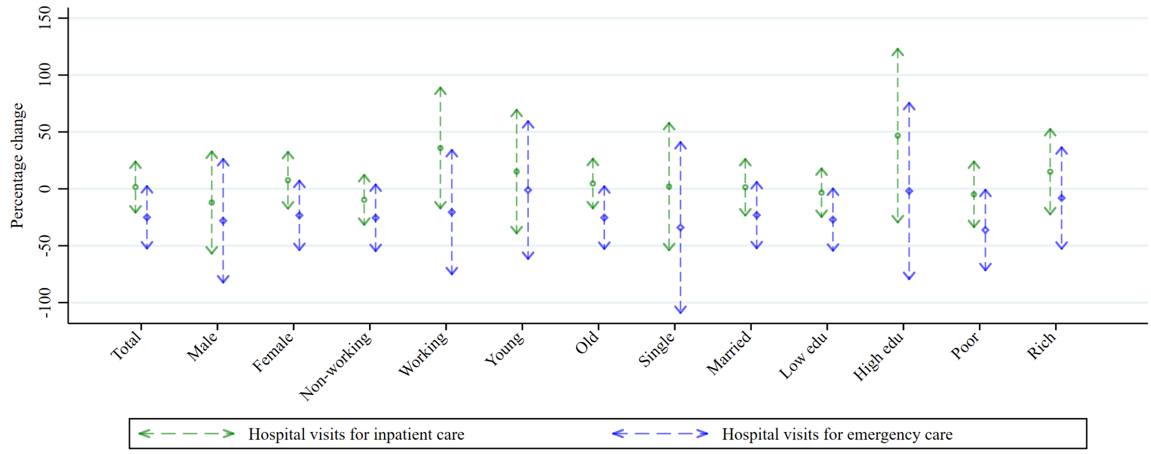
Appendix B Figures

Figure A1: Heterogeneous treatment effect in outpatient care



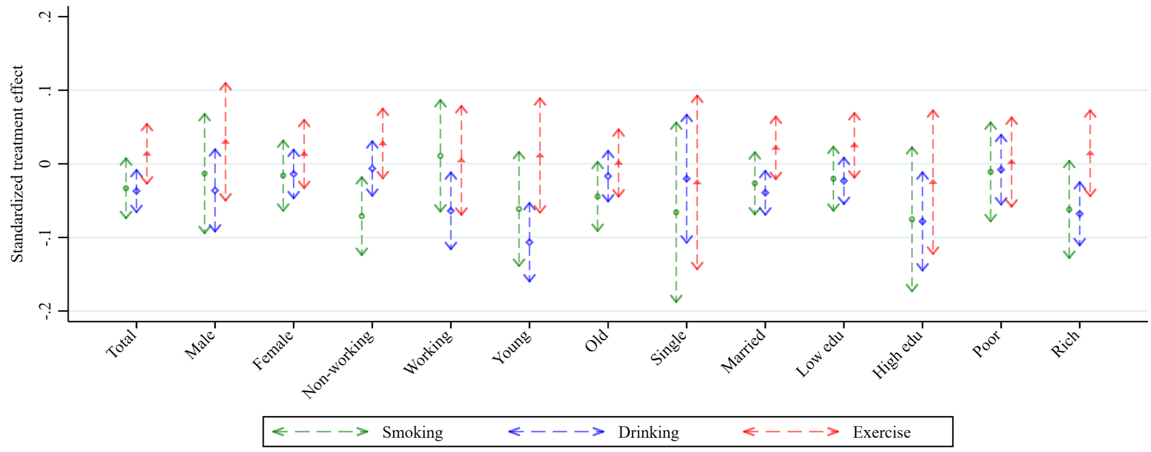
Notes: The figure plots the percentage change in visits to general hospitals, local clinics and first visits for outpatient care for the total and subsamples along with 95% confidence intervals. The percentage change is defined as the change in outcome variable divided by the control group mean.

Figure A2: Heterogeneous treatment effect in inpatient/emergency care



Notes: The figure plots the percentage change in hospital visits for inpatient and emergency care for the total and subsamples along with 95% confidence intervals. The percentage change is defined as the change in outcome variable divided by the control group mean.

Figure A3: Heterogeneous treatment effect in health behavior



Notes: The figure plots the standardized treatment effect for smoking, drinking and exercising behaviors for the total and subsamples along with 95% confidence intervals. The standardized treatment effect is calculated following Kling et al. (2007).