Three Essays on Health Economics

Hyuncheol Kim (Bryant)

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2013

ABSTRACT

Three Essays on Health Economics

Hyuncheol Kim (Bryant)

This dissertation consists of three essays on health economics. The first chapter evaluates the impacts of and behavioral responses to cost-sharing in population-based public cancer screening using Korea's National Cancer Screening Program (NCSP), which provides free stomach and breast cancer screenings to those below the insurance contribution cutoff. Free cancer screening substantially increases the cancer screening take up rate, yielding more cancer detections. Nevertheless, the program was unsuccessful along other key dimensions. First, the initial increase in cancer detections was quickly crowded out by the decrease in cancer detections through other channels, such as private screening and diagnostic testing. Second, those who were induced to take up cancer screening by the cash incentive (compliers) were relatively healthy. These compliers' baseline cancer prevalence is as high as those who take up screening regardless of the availability of free cancer screening (always takers). Those who do not undergo screening regardless of the availability of free cancer screening (never takers) had the highest cancer mortalities, and thus stood to benefit the most from the screening they did not receive. Taken together, free public cancer screening has a limited impact on cancer- and all-cause mortalities. This analysis demonstrates that even when take up is significantly responsive, population based cancer screening can be ineffective due to the behavioral responses to cancer screening such as crowd out and self-selection. More broadly, my study suggests that the impact of health programs, even when they display large participation responses, crucially depend upon the potential behavioral responses of the agents involved.

The second chapter provides empirical evidence on the impacts of government reimburse-

ment of long-term care. We apply a regression discontinuity design using administrative data from South Korea to estimate the impact of subsidies for formal home and institutional care on informal care use and medical expenditures. These subsidies lead to increases in formal long-term care utilization, even accounting for crowd out of private spending. Our main finding is that the benefits of home and facility care are heterogeneous across physical function level and therefore setting policy accordingly has the potential to dramatically reduce medical expenses. We also find that formal long-term care is not a strong substitute for informal long-term care at the extensive margin. Specifically, among individuals who are partially dependent for some activities of daily living (ADLs), we find that increased use of formal home care has no impact on the use of informal care at the extensive margin or on medical expenses. Among individuals who are partially dependent for several ADLs, we find that increased use of institutional care leads to reductions in informal care and medical expenses. From a policy perspective, these results suggest that publicly financed long-term care may have limited impact among the more able, and that home care may be both more cost effective and beneficial than institutional care for the least able.

The third chapter provides empirical evidence on both outcomes and potential mechanisms resulting from information obtained from health screening. We apply a regression discontinuity design using administrative data from South Korea to estimate the impact of different classifications of overall health that vary discontinuously with blood sugar level. We find that "disease suspected" classification leads to increase clinic visit for the secondary examinations and future screening take-ups, and decrease of outpatient days and medical expenditure, however few impacts on health outcomes such as future blood sugar level and mortality. We also find that the responsiveness to the classifications among the highest income quintiles is lower than among the other quintiles, consistent with more educated individuals incorporating information directly from the blood sugar measure itself.

Table of Contents

List of Figures

List of Tables

Acknowledgments

A decision to study economics to become a social doctor was not my own, but was an inspiration from my patients: a farmer who made me cry by asking me"I do not have a cancer, do I" despite clear presentation of terminal stage breast cancer, a mother of young children who was dying as a stomach cancer at the age of 36, and elderly living alone in Asan city whom I served for by home visiting medical program. These experiences have led me to have a heart for the poor and distressed, and every chapter in this dissertation is motivated from my time as a doctor.

It is a humbling experience to acknowledge many people who have helped enormously along the journey of studying economics. I especially thank Kiki Cristian Pop-Eleches who is the true embodiment of a mentor for a life as well as economics. I do not think I can ever repay the debt I owe you. I am also thankful to Douglas Almond who has provided much feedback, support, and time. I have been in luck to have you as a supervisor whom I can respect. I am grateful to Leigh Linden who has encouraged and convinced me especially at the early stage of this dissertation. I want to thank Tal Gross who has provided much guidance and support during my dissertation. This dissertation has benefited greatly from discussions with and comments from Wojciech Kopczuk, Miguel Urquiola, Til von Wachter and seminar participants at Columbia University. For support of my research, I thank Taewha Lee, Woojin Chung, and Sunmi Lee who has helped to find these wonderful data which was crucial in completing this dissertation.

This dissertation would not have been possible without my true friend and colleague, Wilfredo Lim. I am hugely indebted to your support, feedback, and patience. I have always been impressed with your sincerity and selflessness. I would like to express my gratitude

to Booyuel Kim for your sincere commitments. To have a friend whom I can share my life is one of the most important fruits from the PhD program in Columbia. I also thank my library crew from the first year of the program, Youngwoo Koh and Minkee Song, for helping me out endless tough homework.

I am extremely grateful to my parents, who has always respected and supported whatever path I have pursued. This achievement would not have been possible without your supports. To Yae Eun, my life partner, thank you so much for your love, support, patience, and encouragement. I could not have done this without you.

Most of all, I thank and praise God, a lamp for my feet and a light on my path. My life is just a standing on the shoulder of the biggest of all giants.

to my parents

who have given everything for us

and to YaeEun

my life partner

Chapter 1

Public Cancer Screening: Impact and Behavioral Responses

1.1 Introduction

Cancer screening, a testing for cancer in the absence of symptoms, is often thought to play a central role in the fight against cancer.[1] For example, Cutler (2008) argued that cancer screening is the most important factor in explaining the recent cancer mortality reduction in the US. The stakes surrounding cancer screening are large. In 2008, there were 12.4 million new cancer diagnoses and 7.6 million deaths (13% of total deaths).[2] The US spends $10 to $15 billion annually on cancer screening; Korea spends around $400 million for public cancer screening (NCI (2007), NCC (2009)).

Despite its popularity, the effect of cancer screening is surprisingly poorly understood. Evidence on cancer screening from RCTs has been increasing in the last few years,[3] but evidence on population-based cancer screening is still extremely rare even though the effect of population-based cancer screening might differ from that in RCTs (Kadiyala and Strumpf (2011)). For example, the take-up rate in RCTs is close to 100%, which is far higher than in the population setting. If population-based cancer screening (unwittingly) encourages specific groups of people to take up screening, the effect on these selected people might differ from an experimental setting. The take-up rates in a population-based breast cancer screening program were 55.2%, 67.0%, and 73.8% in Korea, the US, and the UK, respectively

[1] It is distinguished from the diagnostic testing that people undergo to detect cancer in the presence of relevant symptoms based on a doctor's recommendation.

[2] Cancers are the second leading cause of death in developed countries and one of the three leading causes of death for adults in developing countries. In terms of mortality, lung cancer is the most common cause of death (1.31 million deaths), followed by stomach (780,000), liver(699,000), colorectal (610,000), and breast (460,000) cancer (Boyle and Levin (2008))

[3] Mammography for breast cancer and fecal occult blood test (FOBT) for colorectal cancer were the only screenings with evidence from RCTs before 2008. Recently, RCTs on the PSA test for prostate cancer (Andriole et al. (2009), Schröder et al. (2009)), low dose computed tomography (CT) (Gross (2011)) and chest X-ray (Oken et al. (2011)) for lung cancer, and sigmoidoscopy (Atkin et al. (2010)) and colonoscopy (Zauber et al. (2012)) for colorectal cancer have been published.

(NCC (2009), NCI (2007), NHS (2008)).[4]

Moreover, at the time of past RCTs, cancer screening was not as popular as it is today. This means that a provision of cancer screening today could be crowded out more easily by outside options. Conceptually, the increase in cancer detections by cancer screening should be expected to erode completely over time. This erosion is predicted from the stylized framework where cancer is eventually detected sometime before death (e.g., through diagnostic testing) and screening per se does not cause cancer. Therefore, the actual effectiveness of cancer screening on health outcomes would depend upon the difference between the timing of cancer detection by screening and that of detection without screening. If cancer detection is quickly crowded out by diagnostic testing, for example, screening would be ineffective. Previous analyses have often ignored both the erosion prediction and the important interplay between initial screening and subsequent testing and detection. Sustained "effects" of screening on cancer detection may be an artifact of endogenous coding, where deaths with but not from cancer are seen to respond (Black, Haggstrom, and Welch, 2002).

In this paper, I evaluate the impacts of and behavioral responses to cost-sharing in population-based public stomach and breast cancer screenings. I use a regression discontinuity (hereafter RD) design that takes advantage of the National Cancer Screening Program (NCSP) in Korea, which provides free cancer screening to those under the insurance contribution cutoff and charges a 50% copayment to those above.[5] I investigate a dynamic aspect of cancer detections through various channels by utilizing data covering the all cancer

[4]Interestingly, even evidence from RCTs is mixed and controversial. For example, four out of the eight RCT studies on breast cancer screening reported breast cancer mortality reductions, while the other four reported no impact. Furthermore, only three out of eight studies were adequately randomized, and no adequately randomized study showed a reduction in breast cancer mortality (Schopper and de Wolf (2009)). Evidence on prostate cancer is also mixed. One in the US found no prostate cancer mortality reduction (Andriole et al. (2009)), while a European study found a 20% prostate cancer mortality reduction (Schröder et al. (2009)).

[5]Insurance contribution is a fixed percentage of basic salary for those with employee insurance.

detections regardless of detection channels. Furthermore, I explore the characteristics and cancer mortality of those induced to take up cancer screening by the program (compliers). Specifically, I compare them to other sub-populations such as those who take up screening regardless of the availability of free cancer screening (always takers) and those who do not take up screening regardless of the availability of free cancer screening (never takers) to shed light on why population-based estimates may depart from RCTs.

I reach two conclusions. First, I find that cancer screening take up increases by around 10 percentage points - more than doubling - when the price of cancer screening decreases from 50% copayment to zero. This also results in significantly more detections in the short-run. However, this detection bump quickly erodes over time through decreases in cancer detections via other channels such as private screening and diagnostic testing. This finding means that public cancer screenings were provided to those who would have taken testing through other channels of cancer detection. Second, I find that there is no difference in baseline cancer prevalence between compliers and always-takers. Moreover, I find that never takers are significantly less healthy compared to compliers and always takers in terms of cancer mortality. This finding implies that the population-based public cancer screening provision did not reach people more in need of cancer screening during the study period.

Taken together, despite its large effect on screening take up and initial increase in cancer detections, subsidizing cancer screening had a limited impact on cancer- and all-cause mortalities up to 6 years after cancer screening because of behavioral responses such as crowd out and selection. My results also imply that the finding from RCTs might be quite different from that in population-based programs due to behavioral responses. To be more successful, cancer screening programs should promote a sufficiently high take up rate in order to reach people in need of cancer screening. Given that cancer screening is already popular, additional provision of cancer screening should be considered with care because such screening can be easily crowded out.

The rest of this paper is organized as follows. Section 2 briefly discusses the previous literature and contribution of my study. Section 3 explains the institutional details. Section 4 describes the data and presents the descriptive statistics. Section 5 shows the estimation strategy. Section 6 presents the results. Finally section 7 provides the conclusions.

1.2 Contribution to Literature

This study contributes to the understanding of the causal effect of population-based cancer screening. Evidence on population-based cancer screening is extremely limited not only due to scarcity of exogenous variation in cancer screening take up but also due to lack of data. Access to population level data with information on cancer screening take up, cancer detections, and related outcomes for long term periods is extremely limited. An exception to this is Kadiyala and Strumpf (2011) who find that U.S. guidelines that recommend screening for breast and colorectal screening starting at age 40 and 50, respectively, generate discontinuous increases in screening rates that result in significant increases in early cancer detection at these ages.

This study has several distinct advantages that improve upon the previous literature. First, it provides reliable estimates on the causal effect of cancer screening by using plausibly exogenous variation in access to cancer screening. Current evidence on the causal effect is limited because the take up of cancer screening is associated with omitted variables (e.g., health-seeking behavior and genetic background) that are also related to health outcomes. I employ an RD design around insurance contribution cutoffs that determine free cancer screening eligibility. This design allows comparison across people with very similar characteristics but sharply different cost-sharing, and thus cancer screening take up rates.

Second, this study finds evidence from a nationwide population-based cancer screening program. This setting provides a unique opportunity to examine selection to screening

effects on various outcomes by exploring characteristics of compliers, always takers, and never takers. This selection can be a reason why effects in experimental settings differ from that in population-based cancer screening.

Third, taking advantage of a large administrative panel dataset covering all cancer detections, this study presents evidence on a dynamic aspect of cancer detections in response to a free public cancer screening offer. I measure not only the increase in cancer detections by public cancer screening, but also the crowd out of cancer detections by other channels over time. The unique setting of a public cancer screening program allows me to evaluate the dynamic feature of cancer detections through various channels. Especially, the crowd out could be timing and setting specific, which could be, at least partially, the reason for the mixed findings in the previous RCT literature.

This study also contributes to the understanding of the impact of cost-sharing of preventive health services. Cost-sharing is a double-edged sword; charging a non-zero price for health services could improve effectiveness by curbing unnecessary demand, but it may also reduce necessary demand, which could lead to worse health outcomes and higher medical expenditures in the future (selection effect). For example, Goldman, Joyce, and Zheng (2007) show that cost-sharing might reduce treatment compliance, which could lead to worse health outcomes and higher future medical expenditures.

In the context of preventive health care, the existence of a selection effect has not been made clear since individuals often are not aware of how much preventive health care they need. For this reason, price sensitivity in preventive health care is potentially different from other therapeutic health care. However, there is remarkably little evidence on the effects of cost-sharing in preventive health care. One of the few exceptions is evidence by Cohen and Dupas (2010), which shows that cost-sharing of insecticide-treated nets (ITN)s for malaria prevention decreases demand without inducing selection of people who are more vulnerable.

Baicker and Goldman (2011) explain that the overall cost-sharing effects consist of the

own-price effect, the cross-price effect, and the effect on health. As many studies have already shown, the demand for health care decreases due to cost-sharing (Newhouse et al. (1981), Manning et al. (1987), Newhouse and Group (1993), Hsu et al. (2006), Chandra, Gruber, and McKnight (2010)). In addition, a change in the price of a particular health care service may affect demand not only for that health service but also for a complementary or substitutable service. For example, Chandra, Gruber, and McKnight (2010) found "offset effects", specifically an increase in hospitalization in response to higher cost-sharing for outpatient or pharmaceutical use in Medicare. Lastly, the previous literature has found that greater utilization of health care is not related with better health outcomes for the average population (Manning et al. (1987), Wennberg and Cooper (1996)), but rather an increased cost-sharing is associated with adverse health outcomes for the vulnerable population (Swartz (2010)). Despite these findings, the evidence on the effect of cost-sharing on health is still scarce overall.

In the context of public cancer screening, I explore how much cost-sharing decreases demand for public cancer screening (own-price effect). In addition, I examine whether cost-sharing in cancer screening invites people who are more likely to have cancer to the screening, or just reduces screening take up without increasing the detection rate (selection effect). I also check whether changes in public cancer screening take up are crowded out by other sources of cancer detections such as private screening and diagnostic testing (cross price effect). Finally, I evaluate whether changes in public cancer screening take up have an impact on health outcomes (effect on health).

1.3 The National Cancer Screening Program (NCSP) in Korea

Korea provides universal health insurance coverage through the National Health Insurance (NHI) and the Medical Care Assistance (MCA). The NHI is available to 95% of the total population, and the MCA covers the rest of the population, that is, the poorest 5%. The National Health Insurance Corporation (NHIC), a single insurer, manages both the NHI and the MCA programs. There are two categories of insurance in the NHI program: employee and self-employed insurance. Employees and their dependents are eligible for the employee insurance, and those who are excluded from employee insurance are eligible for self-employed insurance.[6]

The financial resources of the NHI system mainly come from insurance contributions paid by the insured and their employers. The insurance contribution amount is calculated differently by type of insurance. The contribution rate of the employee insurance, which this study investigates, is based solely upon a fixed percentage of the basic wage.[7]

It is important to note that there are three types of resources for cancer detection in Korea: public cancer screening organized by the NHIC, private opportunistic screening, and diagnostic testing. The first two are screenings for detecting cancer in the absence of symptoms, and the last is a clinically recommended procedure when relevant symptoms are present. Public cancer screening and diagnostic testing are covered by health insurance, while private opportunistic cancer screening is not.

The NHI operates the National Health Screening Program (NHSP) and the National

[6]There are 31.4 and 17.2 million people in Korea with employee and self-employed insurance, respectively. Employee insurance applies to regular employees, but daily wage workers with less than one month of continuous employment are excluded from this category. Spouses, lineal ascendants and descendants, and siblings of employees who do not have remunerations or income are dependents of employee insurance.

[7]The contribution rate was 3.40% in 2001, 3.62% in 2002, and 3.94% in 2003, respectively.

Cancer Screening Program (NCSP). The NHIC implements a national campaign and sends letters to households to promote public health and cancer screenings. The NHSP provides a general health screening, including measurement of Body Mass Index (BMI), blood pressure, blood sugar level, and cholesterol. The NCSP provides cancer screenings. Both programs offer screenings every two years. People born in even/odd-numbered years are strongly encouraged to undertake screenings in an even/odd-numbered year, but those who missed the offer are allowed to undertake screenings in the next year.

Table 1.1 summarizes the NCSP. The NCSP for NHI beneficiaries started in 2002 with stomach, breast, and cervical cancer screenings (Kim et al. (2011)).[8] An upper gastrointestinal (UGI) series, which is a radiologic examination, and an Esophagogastroduodenoscopy (EGD), which is an endoscopic procedure, are used for stomach cancer screening. Screening takers are allowed to choose either of them based on their preferences. EGD, a confirmatory test, is provided to those who received cancer suspicion results from UGI. Mammography is used for breast cancer screening. The price of stomach and breast screenings were approximately $38 and $20, respectively, during the study period (Appendix Table A.1). The prices of public cancer screening and diagnostic testing, both of which are covered by the NHI, are the same, and private screening is more expensive.

The NCSP offers subsidized cancer screenings. The amount of the subsidy is determined by age and insurance contribution amount. Health and cervical cancer screenings are free of charge for people satisfying the age criteria regardless of insurance contribution amount. However, free stomach, breast, and liver cancer screenings are available to those satisfying both the age and insurance contribution criteria shown in Table 1.1.[9] The age cutoff is 40

[8]The NCSP for MCA recipients began in 1999. I limit my sample to only NHI beneficiaries since MCA beneficiaries are also eligible for other social programs such as the National Basic Livelihood Security (NBLS) program.

[9]Liver cancer screenings were introduced in 2003. Liver cancer screening is not a mass screening because it is offered to people with chronic liver disease who account for less than 1% of the population. Moreover, the

years old for stomach and breast cancer screenings. A 50% copayment is applied to those who satisfy the age criteria but not the insurance contribution criteria. The maximum cash incentive is \$19 (=50%×\$38) in males and \$29 (=50%×(\$38+\$20)) in females, respectively. Insurance contribution cutoffs are updated every year based on the government budget situation. During the study period, free cancer screening is available to those with around the lowest 30% of income.[10] The cutoff insurance contribution for employee type insurance was 26,180 and 24,630 Korean Won (KRW) in 2002 and 2003, respectively.

The identification strategy of this study is to compare people just below and above insurance contribution cutoffs among those who satisfy the age criteria. Specifically, the causal effect of stomach cancer screening in males can be estimated by comparing people just below and above the cutoffs. Not a single cancer screening effect can be isolated in females. For females, the combined effect of stomach and breast cancer screenings can be estimated around cutoffs.

1.4 Data

1.4.1 Data Description

The primary analysis relies on the NHI data for those with employee insurance for the years 2001-2008.[11] My empirical analysis requires data on the running variable, level of insurance contribution; indicator for take up of cancer screening; and relevant intermediate and final

cutoff for liver cancer screening in 2003 was 16,750 Korean Won (KRW), far enough from those of stomach and breast cancer screenings.

[10]Basic salary level (without including allowance, bonuses, and incentives) around the cutoff was \$713, and annual medical expenditure is \$702 and \$774 in males and females, respectively. Therefore, the \$19 - \$29 cash incentive for cancer screening is not large.

[11]The data is longitudinal in nature but not a perfectly balanced panel because of deaths and drop outs from the NHI (i.e., becoming an MCA recipient).

outcome variables explaining the effect of and behavioral responses to cancer screening.

The NHI data consists of three parts: eligibility, medical records, and screening.[12] The eligibility component contains basic demographic information such as gender, age, type of insurance, and monthly insurance contribution. Mortality (without cause-of-death information) is included. Also included is individual and household labor market participation. Medical records include medical expenditure based on the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10), which allows me to measure cancer detection and treatment. Lastly, the screening data includes information from health and cancer screenings.

1.4.2 Study Sample

The study sample consists of those with employee-type insurance at the time of screening offer. Males who were previously diagnosed with stomach cancer are excluded from the sample, as are females who were previously diagnosed with stomach or breast cancer. Insurance contribution is the running variable. Specific-year cohort is defined as people with employee health insurance in a specific year. "Even" and "odd" mean those born in even- and odd-numbered years, respectively. For example, "2002 even cohort" refers to people who are aged 40 and over, were born in an even-numbered year, and have employee health insurance in 2002. The main sample is a stacked-up sample of the 2002 even and 2003 odd cohorts aged 40 and over. Cohorts are stacked up by standardized insurance contribution, which measures how far each individual's insurance contribution is from the cutoff (=(Insurance contribution - Cutoff)/Standard deviation).

The outcome variables are cancer detections, cancer mortality, and all-cause mortality.

[12]The insurance cutoff for free cancer screening is determined based on November of the previous year. Therefore, I match November eligibility for the years 2000-2007 to the medical records and screening data of 2001-2008.

Cancer detections are based on ICD-10 information, which captures all cancers regardless of detection channels. One concern with using ICD-10 information is over-diagnosis.[13] To prevent misinterpretation, I restrict cancer detections to only if medical expenditure on cancer in the first year of detection is greater than 300,000KRW (≈$300).[14] Alternative definitions with different restrictions are used for the robustness check: no restriction and with non-zero medical expenditure on cancer in two or more sequential years.[15] Results are similar across definitions. Cancer detection by public cancer screening is defined as cancer detection with take up of public cancer screening either in the same or previous year[16]; otherwise it is categorized as cancer detection by other channels.[17]

The final health outcomes are cancer-related and all-cause mortalities. Cancer-related mortality is defined as death with non-zero medical expenditure on cancer in the last year of death.[18] Cancer-related mortality is potentially a more comprehensive concept than cancer-specific mortality in that cancer-specific mortality captures death where cancer is a main cause of death, while cancer-related mortality encompasses death where cancer is a comorbid

[13]An anecdotal story is that doctors are likely to input the "stomach cancer" ICD-10 code even when they are still just suspicious of the cancer (i.e., a malignant-looking stomach cancer). This is because it is preferable to record a more serious disease, as doing so means more procedures can be covered by the insurance.

[14]This definition is recommended by the National Cancer Center (NCC) of Korea. According to the NCC, more than 90-95% of cancer cases meeting this definition is matched with the national cancer registry, an official record of cancer cases in Korea. Unfortunately, data from the national cancer registry is not available in this study. The $300 restriction excludes 10.8% and 15.3% of cancer cases in males and females, respectively

[15]The two sequential year restriction excludes 27.3% and 33.4% of cancer cases in males and females, respectively. The two sequential year restriction may exclude early stomach cancer cases that are not necessary for the subsequent chemo- or radio-therapy. Results are available by request.

[16]Cancer detected by public screening late in the previous year can be captured in next year.

[17]Cancer detections with take up of both channels are categorized as detection by public cancer screening. Public cancer screening take ups might lead to extra take up of diagnostic testing for confirmation, but not the other way around.

[18]It also must satisfy the cancer detection restriction that medical expenditure in the first year of cancer detection is greater than $300.

condition such as suicide due to depression accompanied by cancer. All-cause mortality equals to one if an individual died for any reason, and zero otherwise. It is, of course, the most comprehensive outcome in my analysis.

1.4.3 Summary Statistics

Table 1.2 presents the descriptive statistics of the study sample. Panels A to F of Table 1.2 describe the general information, baseline medical expenditure, cancer screening take up, cumulative cancer incidence, and cumulative mortality. Around 10% of the population took up cancer screening within the first two years (panel C) and the cumulative screening take up within 6 years is over 50%. The cumulative stomach cancer incidence (up to 6 years) is 1.3% for males and .5% for females. In terms of mortality, the cumulative all cause mortality (up to 6 years) is 7.3% for males and 5.7% for females. The cumulative cancer mortality (also up to 6 years) is 4.7% for males and 4.3% for females.

Panels A, B, and C of Table 1.3 provide the results of the stomach and breast cancer screenings. Each panel consists of two sub-panels. The first sub-panel presents the statistics of the entire sample, which includes all individuals regardless of whether they get screened for cancer or not. The next sub-panel presents statistics by the cancer screening result of screening takers. Column (1) presents the total number of people in each category, and column (2) presents the number of cancer detections within two years in each category. Cancer incidence, the proportion of new cancer cases out of the total number of people, is presented in column (3). For example, the two-year incidences of stomach cancer in males and females were 0.44% (=17,447/3,948,584) and 0.19% (=8,482/4,41,1321), respectively.

Table 1.3 reveals four important facts about the efficiency of cancer screening. The first is that the rate of false negatives is low. As an example, the probability that a screening reports a stomach cancer-free result even though a patient has cancer is 0.23% for men and 0.06%

for women.[19] The second is that the rate of false positives, the probability that a screening reports a cancer suspicion for a cancer-free patient, is high. For example, the false positive rate is 93.53% (=100-6.47%) for stomach cancer screening in males. Such a high false positive rate is not surprising given that cancer screenings tend to minimize false negatives and largely tend to ignore false positives. The third is that cancer screening detects other types of diseases in addition to cancer. For example, stomach cancer screening detects benign diseases such as gastritis, stomach ulcer, and duodenal ulcer. The fourth is that the total number of new cancer detections through public cancer screening is low. For example, public cancer screening detects 749 out of 17,447 stomach cancers in males. Stomach cancer detection by public cancer screening accounts for only 4.3% of total detections. The corresponding numbers in female stomach and breast cancers are 3.8%(=320/8,482) and 4.8%(=347/7,299), respectively.

1.5 Estimation Strategy

1.5.1 Empirical Analysis Setup

I take advantage of the insurance contribution cutoff for free cancer screening eligibility in order to estimate the effect of cancer screening. This corresponds to the intent-to-treat effect of offering free cancer screening versus charging a 50% copayment without controlling for any subsequent take up of cancer screening. I consider the following main regression equation:

$$Y_{it} = \beta \cdot \mathbf{1}(I_i > \tau) + f(I_i) + \psi + \epsilon_i \tag{1.1}$$

where Y_{it} is outcomes for an individual i, such as cancer mortality or all cause mortality, t

[19] In reality, the number of false negative can be smaller than the suggested statistics because new cancer cases developed after the screenings are included in these statistics.

years after the cancer screening offer. $\mathbf{1}(\cdot)$ is an indicator function for whether an individual's insurance contribution (I) is greater than or equal to the cutoff, τ, which determines eligibility for free cancer screening. $f(\cdot)$ is a flexible polynomial function of I. ψ is a cohort fixed effect and ϵ is an error term. Considering that the distribution of standardized insurance contribution is not continuous, errors are clustered by the level of the normalized insurance contribution as suggested by Lee and Card (2008). The analysis is done separately by gender as different types of cancer screenings are offered based on gender.

The idea behind the RD design is that the discontinuity measured by β measures the causal effect of cancer screening, if all other factors except cancer screening take up are smooth around the cutoff. If this assumption holds, people right above and below the cut-off can serve as proper control and treatment groups, respectively, and therefore any difference in outcomes, which is captured by β, can be attributed to eligibility for cancer screening.

1.5.2 Bandwidth Selection and Modeling f(Ii)

Bandwidth selection is one of the critical decisions in the RD model. Since there is no universally accepted convention for how to choose the optimal bandwidth, I try several ways that have been proposed in the literature. In my analysis, I use a bandwidth of 0.3 as well as the Imbens-Kalyanarman (IK) optimal bandwidth suggested by Imbens and Kalyanaraman (2012). My preferred bandwidth is 0.3, since it is wide enough not to be too imprecise and narrow enough to compare observations around the cutoff. Furthermore, I use a rectangular kernel and the local linear regression method suggested by Hahn, Todd, and Van der Klaauw (2001) for modeling f(Ii).

1.5.3 Smoothness of Predetermined Characteristics around the Cutoff

An important assumption of RD design is that individuals just below and just above the cutoff can be compared with each other. There are several reasons why this assumption might not hold. One concern might be that those slightly above the cutoff may reduce their income level in order to become eligible for free cancer screening. However, such manipulation of income reporting is extremely unlikely. First, the cutoff for the program is decided annually based on the government budget and the cutoff is not announced in advance.[20] Secondly, it is not likely for people to manipulate their income level in order to get such small cash incentive.

Appendix Figure A.1 provides a visual illustration of the density of observations by the standard insurance contribution using the smallest bin size around the cutoff. In addition, I test for differences in observable characteristics around the cutoff. Appendix Table A.2 presents estimates of the discontinuity around the cutoff for predetermined variables such as age, general screening take up, employment status, and medical expenditure. Most variables appear to be continuous around the cutoff.

1.5.4 Compliers, Always Takers, and Never Takers

My analysis estimates the local average treatment effect (LATE) for compliers around the cutoff. Since compliers are not randomly selected from the population, the impacts of cancer screening in this study do not necessarily represent that for the average population. Moreover, in my specific case more than 80% of the sample remain never takers during the study period.

[20]For example, eligibility for cancer screening in 2002 was decided by the insurance contribution of November 2001, and screening was offered starting in January 2002.

I propose two different ways to compare the characteristics of compliers, always takers, and never takers. First, I compare compliers and always takers by restricting the sample to screening takers. Since everyone has undergone public cancer screening in the restricted sample, any difference around the cutoff is due to the compositional change of screening takers around the cutoff. In this sample, those right below the cutoff consist of always takers and compliers while those right above the cutoff are always takers. Thus, the analysis with the restricted sample allows me to compare the characteristics of compliers and always takers. Similarly, I restrict the sample to screening non-takers. This allows me to compare compliers and never takers.

Another way to compare compliers characteristics is suggested by Almond and Doyle (2011). Under the assumption that other things are equal around the cutoff, always takers and never takers are identified at just above and below the cutoff, respectively. Even though compliers are not identifiable, observable characteristics of compliers can be calculated from the sample (Abadie (2003)). To do so, I first define a binary variable F, an indicator for free cancer screening eligibility:

$$F = \begin{cases} 0 & \text{above cutoff} \\ 1 & \text{below cutoff} \end{cases}$$

Next, I also define a binary variable S, an indicator for cancer screening take up:

$$S = \begin{cases} 0 & \text{not take up of cancer screening} \\ 1 & \text{take up of cancer screening} \end{cases}$$

Lastly, I define S_F, as the value S would have if F were either 0 or 1. For example, $E(X|S_1 = 1)$ presents the mean value of screening takers in the eligible group.

To estimate complier characteristics, three conditions are required: the existence of a first stage, monotonicity and independence. First, the existence of first stage implies that the

probability of cancer screening take up is higher in the eligible group than in the non-eligible group. This is empirically testable. Second, the monotonicity assumption implies that $S_1 > S_0$ for everyone with probability 1. In other words, anyone who takes cancer screening in the absence of the cash incentive would also undertakes cancer screening in the presence of the cash incentive. This is not directly testable since I do not observe S_1 and S_0, but it is reasonable to assume monotonicity in my setting. Third, independence implies that S_1 and S_0 are independent of F and the potential outcomes. This is not directly testable either, but it is plausible not only because eligibility is determined by the government ex ante but also because people are not likely to manipulate income in order to get a small cash incentive. To see this, the smoothness of the observable characteristics around the cutoff is shown in section 5.4.

Let's first consider $E(X|S_1 = 1)$. It can be written as:

$$E(X|S_1 = 1) = E(X|S_1 = 1, S_0 = 1)\cdot P(S_0 = 1|S_1 = 1)+E(X|S_1 = 1, S_0 = 0)\cdot P(S_0 = 0|S_1 = 1) \tag{1.2}$$

Equation (1.2) implies that $E(X|S_1 = 1)$ is divided by always takers and compliers components. $E(X|S_1 = 1, S_0 = 0)$ represent the characteristics of compliers I am interested in. $E(X|S_1 = 1, S_0 = 1) = E(X|S_0 = 1)$ holds by the monotonicity assumption. P(S_0=1) and P(S_1=0) can be directly measured from the sample. P(S_0=1), the proportion of always-takers, can be measured by π_A the proportion of screening takers in the non-eligible group. Similarly, the proportion of never-takers, P(S_1=0) also can be measured by π_N, the proportion of screening non-takers in the eligible group. The proportion of compliers (π_C) is 1- π_A - π_N. Therefore, $P(S_0 = 1|S_1 = 1)$ and $P(S_0 = 0|S_1 = 1)$ are $\frac{\pi_A}{\pi_C+\pi_A}$ and $\frac{\pi_C}{\pi_C+\pi_A}$, respectively.

Finally, by rearranging the components of equation (1.2), the mean characteristics of

compliers are presented by the terms that can be calculated with the sample:

$$E(X|S_1 = 1, S_0 = 0) = \frac{\pi_C + \pi_A}{\pi_C} \cdot \Big[E(X|S = 1, F = 1) - \frac{\pi_A}{\pi_C + \pi_A} \cdot E(X|S = 1, F = 0) \Big] \quad (1.3)$$

1.6 Results

This section presents the results from equation (1.1). I first explore the effects of cost-sharing on screening take up and crowd out behaviors. Specifically, I present evidence of discrete changes in eligibility and a subsequent increase in cancer screening take ups. I then describe dynamic changes in cancer detections by public as well as other channels. I also explore the characteristics of compliers, always takers and never takers to explore self-selection. Lastly, I estimate the causal impacts of the increase in public cancer screening take up on mortality, and other behavioral responses.

1.6.1 Effect of Cost-sharing on Screening Take Up

I first illustrate that the eligibility for free cancer screening increases from 0 to 1, as shown in Figure 1.1. I plot the standardized insurance contribution that determines eligibility on the x-axis, and the outcomes on the y-axis. The solid lines present the fitted values from equation (1.1) with local linear regression using a 0.3 bandwidth and a rectangular kernel. The open circles in the figure display the means of the fitted values that are collapsed into bins containing individuals who are within 0.05 of a standardized insurance contribution. The vertical difference between two points right below and over the cutoff (vertical line) is an analog of β in equation (1.1).[21] Its regression analog is shown in column (1) of Table 1.4.

Columns (2) to (5) of Table 1.4 show how much eligibility for the cash incentive translates

[21]Figure 1.1- 1.15 have similar structures where the standardized insurance contribution is plotted on the x-axis, and the outcome variable on the y-axis and the open circles are the mean of the outcome in each bin.

into an increase in cancer screening take up. Panels A, B, and C of Table 1.4 show an increase in male stomach, female stomach, and female breast cancer screening take ups. Columns (2) and (3) present the cancer screening take up in the first year with bandwidth 0.3 and IK optimal bandwidth, respectively. Columns (4) and (5) present the cumulative cancer screening take up until the second year from the screening offer.[22] Figure 1.2 corresponds to column (4) of Table 1.4. As expected, eligible people took up public cancer screening mostly in the first year. Up to the second year, male stomach, female stomach, and female breast cancer screening take ups increased by 8.3%, 10.9%, and 10.7% points, respectively. This corresponds to an 83.6, 86.9, and 84.4 percent increase.[23]

The estimated arc-elasticities of demand is around -0.47.[24] The estimated arc elasticity is close to the elasticity of preventive health products in developing countries, such as -0.6 for chlorine, a disinfectant that prevents water-borne diseases in Zambia (Ashraf, Berry, and Shapiro (2010)), and -0.37 for ITNs for malaria prevention in Kenya (Cohen and Dupas (2010)). On the other hand, it is much bigger than the elasticity in therapeutic care in developed countries, such as -0.07 to -0.21 for ambulatory utilization in Korea (Kim, Ko, and Yang (2005)), around -0.2 for health care for the non-elderly in the US (Newhouse and Group (1993)), -0.10 for clinic visits for the elderly in the US(Chandra, Gruber, and McKnight (2010)), and -0.15 to -0.17 for the elderly in Japan (Shigeoka (2011)).

Next, I examine the impact of past public cancer screening take up on future public

[22]Remember that people born in even/odd-numbered years are strongly encouraged to take cancer screening in an even/odd-numbered year, but those who missed the offer are allowed to take up the screening in the next year. Therefore, the offer is actually valid for two years.

[23]Percentage increase is calculated by the formula $\frac{A}{B}$, where A is a the β from equation (1.1), and B is the mean of predicted value at just below and above the cutoff from the local linear regression with bandwidth 0.3

[24]The arc-elasticities are calculated as $((Q_2 - Q_1)/(Q_1 + Q_2)/2)/((P_2 - P_1)/(P_1 + P_2)/2)$. Comparing the arc-elasticity in a zero price setting to those in other settings could be problematic because the denominator, $(P_2 - P_1)/(P_1 + P_2)/2$, is always 2 if P_1=0. Moreover, people treat a zero price as not only a decrease of cost but also as an extra benefit (Shampanier, Mazar, and Ariely (2007)). This must be interpreted with this caveat.

cancer screening take up. If past and future cancer screenings are substitutes, future public cancer screening take up would decrease. On the other hand, if they are complements, future cancer screening take ups would increase. I first check whether there is a change in eligibility for future free public cancer screening. Unless the free cancer screening offer influences future wage levels (and thus insurance contribution), the eligibility for future cancer screening should be smooth around the cutoff. Figure 1.3 and columns (6) and (7) of Table 1.4 confirm the limited change in future eligibility.[25] I find no impact on future public cancer screening take ups as shown in columns (8) and (9) of Table 1.4 and Figure 1.4.

In sum, free cancer screening increased the demand for public cancer screening dramatically and take up of previous public cancer screening does not influence future take up.

1.6.2 Effect on Cancer Detections: Dynamic Aspect of Cancer Detections

In this section, I study whether public cancer screening actually promotes cancer detections, and explore whether the increased cancer detections diminish over time. As mentioned before, if cancers are eventually detected before death, the initial increase in cancer detections by the cancer screening program should be crowded out over time by other channels such as diagnostic testing. Therefore, while the crowd out by private cancer screening or diagnostic testing is expected, what is important is the time it takes for the crowd out to occur. The effect of cancer screening would depend upon the difference between the timing of cancer detection by screening and that of detection without screening.[26]

[25]The dependent variable is a summation of eligibility between years 3 and 6. Since the screening is offered every two years, it ranges between 0 and 2.

[26]It is worth mentioning that cancer detections (and medical expenditures) are observed only if individuals are under the NHI. It is important to address the concern of systematic sample selection by dropping out of the NHI, which could account for my finding. Therefore, whether public cancer screening had any impact on eligibility for the NHI is another relevant outcome. To check this possibility, I look at the NHI status

Table 1.5, Table 1.6, and Table 1.7 present the dynamic change of cumulative stomach cancer detections in males, stomach cancer in females, and breast cancer in females, respectively. Panels A, B, and C in each table present cumulative cancer detections by public cancer screening and by other channels, and overall cumulative detections. Overall detection is the summation of detection by public cancer screening and other channels. Columns (1) to (6) show cumulative cancer detections over a six year period. Columns (7) and (8) present cumulative cancer detections between 3 and 6 years after the cancer screening offer. Figure 1.5 and Figure 1.6 are analogs of columns (2) and (7) in each table.

Table 1.5 and Table 1.6 reveal that impacts on stomach cancer detections are similar for males and females. First, stomach cancer detections by public cancer screening significantly increases by 0.045% points for males (a 30.3% increase) and 0.022% points for females (a 36.0% increase) up to the second year of the screening offer (panel A). Second, cancer detections by other channels decrease as well (i.e., crowd out). As a result, overall cancer detections in males and females increase by 0.020% points (an 8.8% change) and 0.018% points (a 4.4% change) in the first year of the screening offer, but both decrease to zero within a year (panel C). The time it takes for the crowd out to occur is no more than a year. I also find similar result from breast cancer screening (Table 1.7). To summary, increased cancer detections by free public cancer screening were quickly crowded out, less than one year, by other channels including private screening and diagnostic testing.

1.6.3 Stage at Which Cancer is Detected

In order to be effective the screening should find the cancer at the earlier stages, but detecting cancer early does not necessarily mean detecting it at the earlier stages. If the target cancer grows slowly, for example, thyroid cancer, the difference in timing in cancer detections would

directly. I find no statistically significant difference in the NHI status.

not translate to detection at the earlier stages.[27] The question, then, is whether this short advance in timing of cancer detection translates into cancer detection at the earlier stage. In this section, I explore whether the difference in the timing of cancer detection I find in section 6.2 actually translated into cancer detection at the earlier stages.

The dependent variable that I use is the amount of medical expenditure in the first year of cancer detection. I believe that it is a good proxy for the stage of the cancer because it reflects the intensity of cancer treatment. Higher medical expenditures may imply a more advanced cancer stage. However, given the low incidence of cancer, the sample for which medical expenditure in the first year of cancer detection could be measured is much smaller than the initial sample,[28] and this might limit the precision of the estimates.

Figure 1.7 provides a graphical illustration of the level of medical expenditure in the first year of stomach and breast cancer detection during the first two years (panel A) as well as between 3 to 6 years after the screening offer (panel B). Table 1.8 is the regression analog. I find no evidence on stomach and breast cancer detection at the earlier stages. This suggests that cancer detection one year early did not actually translate to cancer detection at the earlier stages.

1.6.4 Selection to Cancer Screening

Selection Effect by Cost-Sharing

I examine whether cost-sharing in cancer screening changes the types of testers. To do so, I compare compliers and always takers by restricting the sample to screening takers as

[27]Moreover, if the target cancer is too malignant, for example, pancreatic cancer, screening could detect cancer at the earlier stage but doing so might not translate to mortality reduction because earlier intervention is less likely to be successful.

[28]Even though stomach and breast cancers are one of the most common cancers in Korea, the annual incidence for people aged 40 and over is no greater than 0.5%

suggested in section 5.3. Table 1.9, Figure 1.8 and Figure 1.9 illustrate the cancer screening results and cancer detections among screening takers. Panels A, B, and C of Figure 1.8 present the probabilities of having normal results, cancer suspicion results, and results for other diseases, respectively. They show that compared to always takers, compliers are more likely to have normal results (panel A), less likely to have cancer suspicion results (panel B, female stomach cancer), and less likely to have results for other diseases. Always takers are actually are more likely to have symptoms than compliers.

However, these relationships do not apply to stomach and breast cancers: I find no difference in cancer detections between compliers and always takers. This implies that the baseline health status of compliers in terms of cancer prevalence is as good as that of always takers. From a different perspective, it also implies that cost-sharing reduces the demand for cancer screening without increasing the efficiency of cancer detection.[29]

Characteristics of Compliers, Always Takers, and Never Takers

As mentioned above, the effects I measure stem from compliers. Since compliers are not randomly selected from the sample, understanding the characteristics of compliers, always takers, and never takers is important. Table 1.10 presents summary statistics of the entire sample for bandwidth [-0.3,0.3], compliers, always takers in bandwidth [0,0.3], and never takers in bandwidth [-0.3,0]. As explained in section 5.3, the characteristics of compliers can be estimated with the proportion of always takers (π_A) and never takers (π_N), as well as the average characteristics of always takers ($E(X|S=1,F=0)$), and eligible screening takers ($(X|S=1,F=1)$). The estimated proportions of compliers, always takers, and never takers are presented in panel A1, B1, and C1 of Table 1.10. The proportion of compliers is between

[29]This finding is similar to the result that cost-sharing of ITNs for malaria prevention in Sub-Saharan Africa decreases the demand without inducing a selection of people who actually need the ITNs more (Cohen and Dupas (2010)).

9% to 12%, and more than 80% of the sample remained never takers.

I find that never takers, those who did not undergo public cancer screenings even with a cash incentive, are significantly different from compliers and always takers. In contrast to the belief that people with higher risk are more likely to utilize medical services, never takers in cancer screening have the highest risk in terms of cancer mortality. Even though the 6-year cumulative stomach cancer detection rate is lowest among never takers (with always takers at 1.7%, compliers at 1.5%, and never takers at 1.3%)[30], the 6-year cumulative stomach cancer mortality is highest among males in this group (with always takers at 0.24%, compliers at 0.34%, and never takers at 0.54%).[31] In female stomach and breast cancers, the 6-year cumulative cancer mortality in never takers is much greater than that in compliers and always takers.

Appendix Figure A.2 illustrates that cancer screening take up is negatively correlated with health status, which in turn might also be related to cancer incidence and mortality. This finding implies that public cancer screening did not reach people who needed cancer screening the most during the study period. This potentially explains why, as described in a later section, I find no evidence of reductions in mortality.

Figure 1.10 illustrates another aspect of the characteristics of compliers, always takers, and never takers. Panel A compares compliers with always takers by using the sample of screening takers, and panel B compares compliers with never takers by using the sample of screening non-takers. Table 1.11 is its regression analog. It is important to note that panels A and B compare compliers with always takers and never takers indirectly because compliers are not identifiable. For example, I compare never-takers right below the cutoff

[30]Low detection among never-takers does not mean that cancer prevalence is lowest among never-takers. Further, never-takers are not diagnosed with cancer through public cancer screenings but through other channels or future public screening.

[31]Since I exclude previous cancer patients from the study sample, the cancer mortality that I measure is death from cancer that developed after the screening offer.

with the combined sample of 5.6%(= 5.1/(85.3 + 5.1)) compliers and 94.4% never-takers right below the cutoff in male stomach cancer screening. Even though the power of the test significantly decreases, I find a significant difference in female breast cancer screening mortality between non-takers and compliers. This result confirms that non-takers have the highest breast cancer mortality.

1.6.5 Effect on Health Behaviors: Health Screening and Medical Expenditure

I also explore additional behavioral responses: future general health screening and medical expenditure. Take up of free general health screening, which is offered every two years, is shown in Figure 1.11 and columns (1) and (2) of Table 1.12. The results indicate no significant impact on general health screening take up. I also evaluate the impact on medical expenditure, as shown in Figure 1.12 and columns (3) to (8) of Table 1.12. To increase precision, I use the differences in medical expenditure from baseline medical expenditure as dependent variables. I find no significant change in medical expenditure.

1.6.6 Effect on Health Outcomes: Bio-markers and Mortality

First, I explore intermediate health outcomes such as the probability of being obese, blood sugar level (Diabetes Mellitus (DM) indicator), and total cholesterol level (hyperlipidemia indicator).[32] Table 1.13 reveals no evidence of a change in intermediate health outcomes.[33]

Finally, I explore the impact on a number of mortality measures. Figure 1.15 presents

[32]I additionally have results for blood pressure, γGTP, and hemoglobin level. I find no change in these outcomes. These results are available by request.

[33]Intermediate health outcomes can be measured only for the general health screening takers. It is possible that this selection is a source of bias. Appendix Figure A.3 reveals that screening takers are more likely to have better health status (upward bias). Thus, the coefficient estimates on intermediate health outcomes would be upper bound.

the average effect on cancer-, non-cancer-, and all-cause mortalities, respectively. Table 1.14 presents the corresponding estimates of β from equation (1.1). I do not find evidence of an effect on stomach and breast cancer mortalities. This is not a surprise at all given that cancer screening has a limited impact on the stage at which cancer is detected. I do not find any statistically significant changes in all-cause mortality either. Figure 1.16 illustrates a changing trend in mortalities over time. Even though none of the estimates is statistically significant, it presents a decreasing pattern in males and an increasing pattern in females, suggesting that 6 years could be too short to measure mortality outcomes.

1.6.7 False Positive

Behavioral responses to cancer screening might differ by screening result. Especially, behavioral responses for those with false positive results are of a particular interest. To see this, I implement another RD regression with the additional interaction terms of eligibility and screening result. Screening result is categorized as result of "normal","cancer detection","false positive", and "other type of diseases". Cancer detection is the case when an individual is diagnosed with cancer after receiving a cancer suspicious result; otherwise it is categorized as a false positive. I construct an additional RD model in the following way:

$$Y_{it} = \beta \cdot \mathbf{1}(I_i > \tau) + \delta \cdot \mathbf{1}(I_i > \tau) \cdot \sum R_i + \eta \cdot \sum R_i + f(I_i) + \gamma X_i + \epsilon_i$$

where R is a dummy for the four types of cancer screening results. Since the cancer screening results are endogenous, results of this analysis need to be interpreted with care.

The results are presented in Table 1.15. Columns (1) and (2) show that males with false positive result from stomach cancer screening are more likely to take health and stomach cancer screenings in the future. Columns (6) to (8) show that females with false positive result from breast cancer screening are more likely to take stomach and breast cancer screen-

ings in the future. I find no significant change in medical expenditure both for males and females with false positive result. It implies that false positives only induce additional clinic visits and procedures, which are covered by the public cancer screening program.

1.6.8 Cost Analysis

This section presents estimates of the cost-effectiveness of the program, in terms of cancer detection, from a societal perspective.[34] Panels A, B, and C of Table 1.16 present calculations of the cost per cancer detected for male stomach, female stomach, and breast cancer screenings. Column (1) shows the number of screenings, travel days, and lost work days per cancer detected. Column (2) indicates the corresponding unit costs.[35] Column (3) presents the total cost of each category, and column (4), the proportion of that categorys cost to the overall cost. The overall cost of each cancer screening is the sum of direct screening costs, transportation costs, and opportunity costs.

Those who were induced to screen by the program are compliers. Accordingly, the numbers and costs presented in columns (1) and (2) are determined from equation (1.3). These are further weighted by the proportion of compliers in 2002 and 2003.[36] Medical expenditures induced by false positives and cancer detections are not included since there are no significant impacts as shown in columns (5) and (11) of Table 1.15.

I first consider the direct screening cost. Since individual level biopsy data are not available, I use the average biopsy rate in 2002 and 2003, and assume that it is constant around

[34]There is no "benefit" in terms of mortality since I do not find change in cancer- and all-cause mortality as shown in Table 1.14.

[35]The price of cancer screening is described in Appendix Table A.1.

[36]Proportions of compliers in male stomach cancer are 6.7% and 11.6% in 2002 and 2003, respectively. Weighted average is calculated by the formula, $\frac{X_{2002}\times 6.7 + X_{2003}\times 11.6}{6.7+11.6}$. Corresponding figures in female stomach cancer are 9.2% and 13.3%, and in breast cancer are 8.5% and 13.6%.

the cutoff.[37] Screening is often accompanied by procedures paid for out-of-pocket expenses by the patients such as conscious sedation in the case of EGD and ultrasonography to perform a biopsy of breast tissue. I assume that 30% of EGD procedures are undergone with conscious sedation. I also assume that the cost of conscious sedation and breast ultrasonography are $30 and $100, respectively.

Transportation costs are also considered. Transportation costs are based on the average transportation cost for a clinic visit computed from the Korean National Health and Nutrition Examination Survey in 2005.[38] The number of first visits is equal to the number of screenings undertaken, and I assume that those with "cancer suspicion" result from cancer screening made a follow-up visit.

Lastly, I consider the opportunity cost resulting from lost labor productivity. Since the insurance contribution is a fixed percentage of wage, foregone labor productivity can be directly calculated.[39] Since the insurance contribution is based only on base salary (not including bonuses and benefits), this leads to conservative estimates of the cost. The number of lost work days is the number of screenings undergone by employees in each sample.

The total estimated cost costs for identifying one additional case of male stomach, female stomach, and female breast cancers are $15,073, $59,590, and $63,811, respectively. Direct screening costs, transportation costs, and lost labor productivity account for 58%-66%, 20%-22%, and 13%-19% of total cost, respectively.

---

[37] I use the average biopsy rate of the entire sample. Average biopsy rate of stomach cancer screening were 26.7% and 39.8% in 2002 and 2003, respectively. Average biopsy rate of breast cancer screening were 25.5% and 38.5% in 2002 and 2003, respectively.

[38] The average transportation costs for a round trip clinic visit are $19.31 and $15.61 for males and females, respectively, in 2005. These cost are consumer price index (CPI) adjusted. Based upon a 2005 base of 100, the CPI is 90.747 and 93.946 in 2002 and 2003, respectively.

[39] I assume that screening only takes one day. The monthly wage is divided by 23, the average number of working days in a month, to compute daily labor productivity.

1.7 Conclusion

My paper presents empirical evidence on the impacts of and behavioral responses to cost-sharing in population-based public cancer screening. I use an RD design that takes advantage of the unique experience of the NCSP in Korea, which provides free stomach and breast cancer screenings to those below the insurance contribution cutoff, while charging a 50% copayment to those above.

My results suggest that although free cancer screening substantially increases the cancer screening take up rate and the number of cancer detections, there is no evidence that the program had an impact on cancer- and all-cause mortality rate. My analysis provides two main explanations for these results. First, the initial increase in cancer detections due to the public screening program was quickly crowded out by the decrease in cancer detections through other channels, such as private screening and diagnostic testing. Second, those induced into screening by the cash incentive (compliers) were relatively healthy. These compliers' baseline cancer prevalence and cancer mortality is as high as of those who take up screening regardless of the availability of free cancer screening (always takers), implying no selection effect of cost-sharing. In addition, those who do not undergo screening regardless of the availability of free cancer screening (never takers) had the poorest health and stood to benefit the most from the screening they did not receive.

My result suggests that in order to be successful, a population-based cancer screening program should promote a sufficiently high take up rate in order to reach the people most in need of cancer screening. My study also provides implications on the additional provision of cancer screening, given that cancer screening is already popular. First, provision of cancer screening can be crowded out easily. Crowd out is more likely the more popular cancer screening is, which means better access to outside options. Second, people who are more likely to have cancer would be less likely to participate in cancer screening. Therefore,

incentives for cancer screening must be well-designed in order to reach these people.

My results also imply that findings from RCTs might be quite different from that in population-based programs due to the behavioral responses to the programs. More broadly, even though the findings of this study may reflect responses that are specific to cancer screening in Korea, this analysis demonstrates that the impacts of health programs, even when they display large participation responses, crucially depend on the potential behavioral responses of the agents involved.

Table 1.1: National Cancer Screening Program (NCSP)

Year	Male	Female	Age	Insurance contribution cutoff (KRW)
2002	Stomach	Stomach	40 and over	26,180
		Breast	40 and over	26,180
		Cervix	30 and over	n/c
2003	Stomach	Stomach	40 and over	24,630
		Breast	40 and over	24,630
		Cervix	30 and over	n/c

Note: This table presents cancer screenings covered by National Cancer Screening Program (NCSP) and the cutoffs for free cancer screening. Cervical cancer screening was free of charge for all. For stomach, breast and colorectal cancer were free for those with insurance contribution below the cutoff, while 50% copayment was charged those above. Liver cancer screening is offered since 2003 with a cutoff 16,750 KRW. Liver cancer screening targets on people with chronic liver disease such as liver cirrhosis, and HBV and HCV related liver diseases, explaining less than 1% of the population. Unit is KRW. \$1 $\approx$ 1,000KRW.

Table 1.2: Basic Statistics

	Male			Female		
	N	Mean	Std.Dev	N	Mean	Std.Dev
Panel A. General Information						
Age	4,041,275	53.9	11.2	4,460,789	56.2	12.3
Cancer Screening Eligibility	4,041,275	0.347	0.476	4,460,789	0.374	0.484
Standard insurance contribution	4,041,275	0.468	1.006	4,460,789	0.441	1.006
Employment	4,041,275	0.625	0.484	4,460,789	0.157	0.364
Panel B. Medical expenditure (Unit:1000KRW=$1)						
Total	3,641,741	709.9	1709.2	4,217,969	796.9	1524.9
Non cancer	3,641,741	643.7	1479.7	4,217,969	760.0	1395.1
Cancer	3,641,741	66.3	812.4	4,217,969	37.0	595.3
Panel C. Screening take up (Year 1-2)						
Stomach cancer screening	4,041,275	0.097	0.295	4,460,789	0.110	0.312
EGD	4,041,275	0.042	0.200	4,460,789	0.043	0.202
UGI	4,041,275	0.058	0.233	4,460,789	0.068	0.251
Breast cancer screening				4,460,789	0.114	0.317
General health screening	4,041,275	0.467	0.499	4,460,789	0.287	0.452
Panel D. Cumulative screening take up (Year 1-6)						
Stomach cancer screening	4,041,275	0.512	0.868	4,460,789	0.586	0.840
Breast cancer screening				4,460,789	0.626	0.865
General health screening	4,041,275	1.819	1.744	4,460,789	1.198	1.307
Panel E. Cumulative Cancer Incidence (up to Year 6)						
Stomach	4,021,374	0.013	0.112	4,440,967	0.005	0.074
Breast				4,440,967	0.005	0.073
Panel F. Cumulative mortality (up to Year 6)						
All-cause	4,041,275	0.073	0.261	4,460,789	0.057	0.231
Non-cancer	4,041,275	0.047	0.211	4,460,789	0.043	0.202
Cancer-related	4,041,275	0.027	0.161	4,460,789	0.014	0.117
Stomach cancer-related	4,041,275	0.005	0.071	4,460,789	0.003	0.059
Breast cancer-related				4,460,789	0.002	0.046

Note: This table shows summary statistics of study samples. N is the sample size and Std. Dev refers a standard deviation. The data covers universe Korean people with employee health insurance. Screening take up is defined as 1 for individuals took cancer screening within two-years from the offer. See text for definitions of variables. All measures are at the baseline

Table 1.3: Result of Public Cancer Screening

	(1)	(2)	(3)
Panel A. Male stomach cancer screening			
	Total	Cancer	Cancer incidence
Panel A1. Whole sample			
Total	3,948,584	17,447	0.44%
Screening non-takers	3,687,115	15,599	0.42%
Screening takers	261,469	1,848	0.71%
Panel A2. By screening result (among screening takers)			
Normal	123,462	317	0.26%†
Cancer suspicion	11,585	749	6.47%‡
Other stomach disease	126,422	782	0.62%
Panel B. Female stomach cancer screening			
Panel B1. Whole sample			
Total	4,411,321	8,482	0.19%
Screening non-takers	4,001,177	7,674	0.19%
Screening takers	410,144	808	0.20%
Panel B2. By screening result (among screening takers)			
Normal	228,523	150	0.07%†
Cancer suspicion	11,590	320	2.76%‡
Other stomach disease	170,031	338	0.20%
Panel C. Female breast cancer screening			
Panel C1. Whole sample			
Total	4,411,321	7,299	0.17%
Screening non-takers	3,972,497	6,325	0.16%
Screening takers	437,922	974	0.22%
Panel C2. By screening result (among screening takers)			
Normal	331,144	250	0.08%†
Cancer suspicion	39,518	347	0.88%‡
Other stomach disease	67,260	377	0.56%

Note: This table shows the results of the stomach and breast cancer screenings. Each panel consists of two sub-panels. The first sub-panel presents the statistics of the entire sample, which contains all individuals regardless if they get screened for cancer or not. The next sub-panel presents statistics by the cancer screening result of screening takers. Column (1) presents the total number of people in each category, and column (2) presents the number of cancer detections within two years in each category. Cancer incidence, the proportion of new cancer cases out of the total number of people, is presented in column (3). False negative is statistics with '†' and false positive is 1 - statistics with '‡'.

Table 1.4: Effect of Cost-Sharing on Cumulative Cancer Screening Take up

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependant variable	Eligibility Year 1-2	Screening take up Year 1		Screening take up Year 1-2		Eligibility Year 3-6		Screening take up Year 3-6	
Panel A. Male stomach cancer screening									
bandwidth	0.3	0.3	IK(0.14)	0.3	IK(0.12)	0.3	IK(0.09)	0.3	IK(0.12)
	1.0000**	0.0791**	0.0695**	0.0829**	0.0725**	0.0543	-0.0252*	0.0007	-0.0119*
	(0.000)	(0.0078)	(0.0106)	(0.0114)	(0.0174)	(0.064)	(0.009)	(0.011)	(0.005)
N	1,260,729	1,260,729	606,937	1,260,729	465,339	1,260,729	408,959	1,260,729	465,335
Panel B. Female stomach cancer screening									
bandwidth	0.3	0.3	IK(0.07)	0.3	IK(0.09)	0.3	IK(0.09)	0.3	IK(0.11)
	1.0000**	0.1115**	0.1711**	0.1087**	0.0919**	0.0702	-1.0550**	-0.0015	-0.0185**
	(0.000)	(0.0085)	(0.0098)	(0.0101)	(0.0182)	(0.070)	(0.010)	(0.012)	(0.003)
N	1,396,081	1,396,081	271,655	1,396,081	445,477	1,396,081	271,655	1,396,081	506,252
Panel C. Female breast cancer screening									
bandwidth	0.3	0.3	IK(0.18)	0.3	IK(0.06)	0.3	IK(0.09)	0.3	IK(0.24)
	1.0000**	0.1101**	0.1085**	0.1065**	0.1936**	0.0702	-1.0550**	-0.0021	-0.0086
	(0.000)	(0.0106)	(0.0148)	(0.0124)	(0.0110)	(0.070)	(0.010)	(0.011)	(0.008)
N	1,396,081	1,396,081	823,434	1,396,081	271,655	1,396,081	271,655	1,396,081	1,026,628

Note: Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.5: Effect on Cumulative Stomach Cancer Detection, Male

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year	Year 1	Year 1-2	Year 1-3	Year 1-4	Year 1-5	Year 1-6	Year 3-6	
Panel A. Cumulative stomach cancer detection by public screening								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.35)
	0.00013	0.00045*	0.00044*	0.00052*	0.00069**	0.00051*	0.00006	0.00003
	(0.00015)	(0.00018)	(0.00017)	(0.00019)	(0.00020)	(0.00022)	(0.00010)	(0.00011)
N	1,212,427	1,212,427	1,212,427	1,212,427	1,212,427	1,212,427	1,244,400	1,458,485
Panel B. Cumulative stomach cancer detection by other channels								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.39)
	0.00003	-0.00046**	-0.00068*	-0.00080*	-0.00072+	-0.00063	-0.00010	-0.00010
	(0.00004)	(0.00012)	(0.00028)	(0.00037)	(0.00040)	(0.00041)	(0.00034)	(0.00033)
N	1,212,427	1,212,427	1,212,427	1,212,427	1,212,427	1,212,427	1,244,400	1,733,954
Panel C. Overall cumulative stomach cancer detection								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.37)
	0.00020	0.00004	-0.00017	-0.00019	0.00002	-0.00002	-0.00004	-0.00011
	(0.00017)	(0.00024)	(0.00039)	(0.00050)	(0.00058)	(0.00057)	(0.00037)	(0.00038)
N	1,212,427	1,234,492	1,244,400	1,249,522	1,252,297	1,254,031	1,244,400	1,559,108

Note: Dependent variables in Panels A, B, and C are cumulative stomach cancer detection by public screening, by other channels, and overall cumulative detections, respectively. Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.6: Effect on Cumulative Stomach Cancer Detection, Female

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year	Year 1	Year 1-2	Year 1-3	Year 1-4	Year 1-5	Year 1-6	Year 3-6	
Panel A. Cumulative stomach cancer detection by public screening								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.25)
	0.00013**	0.00022**	0.00015	0.00022+	0.00027*	0.00028*	0.00007	0.00005
	(0.00003)	(0.00005)	(0.00009)	(0.00011)	(0.00012)	(0.00013)	(0.00010)	(0.00011)
N	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,385,151	1,018,478
Panel B. Cumulative stomach cancer detection by other channels								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.18)
	0.00002	-0.00016+	-0.00024	-0.00023	-0.00005	0.00004	0.00018	0.00051**
	(0.00005)	(0.00009)	(0.00015)	(0.00017)	(0.00016)	(0.00015)	(0.00014)	(0.00010)
N	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,385,151	816,816
Panel C. Overall cumulative stomach cancer detection								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.27)
	0.00018**	0.00008	-0.00007	0.00000	0.00022	0.00032+	0.00025	0.00026+
	(0.00006)	(0.00010)	(0.00019)	(0.00018)	(0.00015)	(0.00016)	(0.00015)	(0.00015)
N	1,368,472	1,380,172	1,385,151	1,387,618	1,389,023	1,389,878	1,385,151	1,332,383

Note: Dependent variables in Panels A, B, and C are cumulative stomach cancer detection by public screening, by other channels, and overall cumulative detections, respectively. Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.7: Effect on Cumulative Breast Cancer Detection, Female

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Year	Year 1	Year 1-2	Year 1-3	Year 1-4	Year 1-5	Year 1-6	Year 3-6	
Panel A. Cumulative breast cancer detection by public screening								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.18)
	0.00007+	0.00017**	0.00013+	0.00010	-0.00004	-0.00010	-0.00026+	-0.00035**
	(0.00004)	(0.00004)	(0.00007)	(0.00007)	(0.00013)	(0.00016)	(0.00013)	(0.00011)
N	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,385,151	816,816
Panel B. Cumulative stomach cancer detection by other channels								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.24)
	0.00013+	-0.00014	-0.00018	-0.00037*	-0.00041*	-0.00047*	-0.00034	-0.00009
	(0.00007)	(0.00008)	(0.00011)	(0.00017)	(0.00020)	(0.00022)	(0.00020)	(0.00021)
N	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,368,472	1,385,151	1,018,478
Panel C. Overall cumulative breast cancer detection								
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	IK(0.28)
	0.00027**	0.00008	0.00001	-0.00021	-0.00039**	-0.00051**	-0.00060**	-0.00058**
	(0.00009)	(0.00009)	(0.00012)	(0.00015)	(0.00014)	(0.00017)	(0.00018)	(0.00018)
N	1,368,472	1,380,172	1,385,151	1,387,618	1,389,023	1,389,878	1,385,151	1,332,384

Note: Dependent variables in Panels A, B, and C are cumulative breast cancer detection by public screening, by other channels, and overall cumulative detections, respectively. Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.8: Effect on Early Detection

	(1)	(2)	(3)	(4)	(5)	(6)
Year	Cancer detected within 2 year			Cancer detected after 3-6 year		
Gender	Male	Female	Female	Male	Female	Female
Cancer	Stomach	Stomach	Breast	Stomach	Stomach	Breast
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3
	40.8	169.7	-220.3	-84.0	374.4	-29.9
	(182.5)	(436.1)	(310.6)	(166.7)	(342.2)	(364.5)
N	5,237	2,373	2,163	10,595	4,785	4,562

Note: The dependent variable is the amount of medical expenditure at the first year of cancer detection. Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.9: Selection Effect by Cost-Sharing

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Male stomach cancer								
Dependent variable	Stomach cancer screening result						Cancer detection	
	Normal		Cancer suspicion		Other disease			
Bandwidth	0.3	IK(0.29)	0.3	IK(0.56)	0.3	IK(0.33)	0.3	IK(0.24)
	0.06878**	0.06883**	0.00123	0.00296	-0.07001**	-0.06214**	0.00024	0.00048
	(0.01132)	(0.01134)	(0.00315)	(0.00247)	(0.01103)	(0.01034)	(0.00030)	(0.00031)
N	97,186	97,185	97,186	162,648	97,186	110,433	130,413	111,160
Panel B. Female stomach cancer								
Dependent variable	Stomach cancer screening result						Cancer detection	
	Normal		Cancer suspicion		Other disease			
Bandwidth	0.3	IK(0.31)	0.3	IK(0.64)	0.3	IK(0.33)	0.3	IK(0.16)
	0.05642**	0.05642**	-0.00598**	-0.00692**	-0.05044**	-0.04710**	0.00012	0.00016
	(0.00704)	(0.00704)	(0.00089)	(0.00094)	(0.00731)	(0.00762)	(0.00021)	(0.00025)
N	162,907	162,907	162,907	289,627	162,907	186,289	185,371	95,666
Panel C. Female Breast cancer								
Dependent variable	Breast cancer screening result						Cancer detection	
	Normal		Cancer suspicion		Other disease			
Bandwidth	0.3	IK(0.13)	0.3	IK(0.28)	0.3	IK(0.16)	0.3	IK(0.15)
	0.03899**	0.03467*	-0.00025	-0.00023	-0.03874**	-0.03519**	0.00006	0.00051
	(0.00825)	(0.01300)	(0.00113)	(0.00113)	(0.00759)	(0.00896)	(0.00026)	(0.00029)
N	162,979	56,439	162,979	153,127	162,979	86,649	186,922	95,985

Note: The sample is restricted to screening takers. Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.10: Characteristics of Compliers, Always Takers and Never Takers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Male, stomach cancer screening	Total	Compliers	Always Takers	Never Takers	(2)=(3)	t-stat (2)=(4)	(3)=(4)
Panel A1. Proportion	1.000	0.096	0.051	0.853			
Panel A2. Public screening take-ups							
stomach cancer (Year 1-2)	0.103	1.000	1.000	0.000			
stomach cancer (Year 1-6)	0.542	2.042	2.090	0.363	8.8	2038.7	304.0
general health (Year 1-2)	0.568	1.106	1.134	0.457	10.0	698.4	224.8
general health (Year 1-6)	1.713	2.561	2.811	1.455	35.6	303.1	104.3
Panel A3. Mortality							
Cumulative stomach cancer detection (6 year)	0.0130	0.0167	0.0146	0.0125	3.0	28.8	2.9
Cumulative stomach cancer mortality (6 year)	0.0051	0.0034	0.0024	0.0054	3.5	21.2	10.1
Cumulative all-cause mortality (6 year)	0.0727	0.0493	0.0332	0.0760	15.4	77.4	38.9
Panel B. Female, stomach cancer screening	Total	Compliers	Always Takers	Never Takers	(2)=(3)	T-stat (2)=(4)	(3)=(4)
Panel B1. Proportion	1.000	0.118	0.066	0.816			
Panel B2. Public screening take-ups							
stomach cancer (Year 1-2)	0.133	1.000	1.000	0.000			
stomach cancer (Year 1-6)	0.657	2.085	2.042	0.438	10.7	2009.9	384.8
general health (Year 1-2)	0.352	1.073	0.965	0.291	64.8	1075.1	363.5
general health (Year 1-6)	1.235	2.544	2.380	1.107	33.6	837.7	245.0
Panel B3. Mortality							
Cumulative stomach cancer detection (6 year)	0.0052	0.0049	0.0046	0.0050	1.0	0.7	1.1
Cumulative stomach cancer mortality (6 year)	0.0034	0.0011	0.0006	0.0037	3.7	34.5	20.9
Cumulative all-cause mortality (6 year)	0.0525	0.0140	0.0113	0.0569	5.2	148.4	75.8
Panel C. Female, breast cancer screening	Total	Compliers	Always Takers	Never Takers	(2)=(3)	T-stat (2)=(4)	(3)=(4)
Panel C1. Proportion	1.000	0.118	0.067	0.815			
Panel C2. Public screening take-ups							
breast cancer (Year 1-2)	0.136	1.020	1.010	0.000			
breast cancer (Year 1-6)	0.690	2.119	2.109	0.468	2.6	1952.6	390.2
general health (Year 1-2)	0.352	1.076	0.967	0.289	61.7	1094.4	351.7
general health (Year 1-6)	1.235	2.546	2.397	1.103	29.9	846.0	245.4
Panel C3. Mortality							
Cumulative breast cancer detection (6 year)	0.0052	0.0057	0.0064	0.0051	1.8	6.3	3.1
Cumulative breast cancer mortality (6 year)	0.0022	0.0001	0.0002	0.0026	1.4	40.0	19.2
Cumulative all-cause mortality (6 year)	0.0525	0.0140	0.0123	0.0571	3.8	148.8	81.2

Note: This table presents the mean characteristics of the entire sample for bandwidth [-0.3,0.3] (Column 1), compliers (Column 2), always takers in bandwidth [0,0.3] (Column 3), and never takers in bandwidth [-0.3,0] (Column 4). The mean characteristics of compliers are estimated from Equation (1.3). Columns 5 to 7 present t-statistics from two sample t-test comparing compliers and always takers, compliers and never takers, and always takers and never takers, respectively.

Table 1.11: Comparing Compliers with Always Takers and Never Takers

Dependent Variable : 6-Year Cumulative Cancer Mortality

	(1)	(2)	(3)	(4)
Panel A. Male stomach cancer				
Sample	Screening takers		Screening non-takers	
	Compliers vs. Always takers		Compliers vs. Never takers	
Bandwidth	0.3	IK(0.46)	0.3	IK(0.30)
	0.00070	0.00093+	-0.00053	-0.00059
	(0.00058)	(0.00052)	(0.00035)	(0.00035)
N	130,413	198,231	1,130,316	1,156,970
Panel B. Female stomach cancer				
Sample	Screening takers		Screening non-takers	
	Compliers vs. Always takers		Compliers vs. Never takers	
Bandwidth	0.3	IK(0.31)	0.3	IK(0.32)
	0.00027	0.00027	0.00022	0.00034
	(0.00023)	(0.00023)	(0.00018)	(0.00021)
N	185,371	185,371	1,210,710	1,321,841
Panel C. Female breast cancer				
Sample	Screening takers		Screening non-takers	
	Compliers vs. Always takers		Compliers vs. Never takers	
Bandwidth	0.3	IK(0.36)	0.3	IK(0.33)
	0.00014	-0.00007	0.00029+	0.00040*
	(0.00025)	(0.00020)	(0.00015)	(0.00015)
N	186,922	227,356	1,209,159	1,385,665

Note: The dependent variable is 6-Year cumulative cancer mortality. I restrict sample to screening takers in Columns (1) and (2), and screening non-takers in Columns (3) and (4). Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.12: Effect on Other Behavioral Responses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Male								
Dependent variable	Health screening				Medical expenditure ($)			
	Year 3-6		Year 1-2		Year 3-5		Year 1-5	
Bandwidth	0.3	IK (0.15)	0.3	IK(0.24)	0.3	IK(0.25)	0.3	IK(0.25)
	-0.0410+	-0.0593+	-1.1	0.6	2.1	2.9	11.9	14.0
	(0.021)	(0.033)	(12.5)	(13.9)	(27.8)	(33.2)	(34.3)	(41.1)
N	1,260,729	667,548	970,661	820,817	905,003	765,586	841,751	711,935
Panel B. Female								
Dependent variable	Health screening				Medical expenditure ($)			
	round 2 & 3		Year 1-2		Year 3-5		Year 1-5	
Bandwidth	0.3	IK (0.18)	0.3	IK(0.50)	0.3	IK(0.28)	0.3	IK(0.50)
	-0.0131	-0.0184	12.9	13.5+	-18.0	-13.8	-8.5	-11.7
	(0.013)	(0.015)	(8.8)	(7.2)	(20.9)	(20.7)	(27.6)	(26.4)
N	1,396,081	823,434	1,219,294	2,008,286	1,164,312	1,120,102	1,123,043	1,850,072

Note: Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.13: Effect on Intermediate Health Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Male								
Dependent variable	Blood sugar		BMI		Obesity		Cholesterol	
	R2	R3	R2	R3	R2	R3	R2	R3
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
	0.0858	0.4815	0.0089	0.0131	-0.0033	-0.0021	-0.2125	-0.2866
	(0.229)	(0.364)	(0.011)	(0.010)	(0.002)	(0.002)	(0.326)	(0.415)
N	385,222	340,704	384,997	340,497	384,997	398,406	384,736	339,665
Panel B. Female								
Dependent variable	Blood sugar		BMI		Obesity		Cholesterol	
	R2	R3	R2	R3	R2	R3	R2	R3
Bandwidth	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
	0.7122**	0.0908	0.0297**	0.0328**	0.0088**	-0.0017	-0.9765**	-0.3523
	(0.231)	(0.198)	(0.009)	(0.007)	(0.002)	(0.002)	(0.338)	(0.425)
N	273,933	268,214	273,675	267,961	273,675	353,559	273,417	267,015

Note: Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.14: Effect on Cumulative Mortality (6 Year)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Male										
Dependant variable	Stomach cancer mortality		All-cancer mortality		Non-cancer mortality		All-cause mortality			
Bandwidth	0.3	IK(0.32)	0.3	IK(0.44)	0.3	IK(0.31)	0.3	IK(0.27)		
	-0.00062+	-0.00052	-0.00147+	-0.00086	-0.00202	-0.00167	-0.00349	-0.00336		
	(0.00034)	(0.00035)	(0.00078)	(0.00073)	(0.00155)	(0.00156)	(0.00210)	(0.00216)		
N	1,260,729	1,441,721	1,260,729	1,931,589	1,260,729	1,373,890	1,260,729	1,218,000		
Panel B. Female										
Dependant variable	Stomach cancer mortality		Breast cancer mortality		All-cancer mortality		Non-cancer mortality		All-cause mortality	
Bandwidth	0.3	IK(0.25)	0.3	IK(0.28)	0.3	IK(0.32)	0.3	IK(0.24)	0.3	IK(0.33)
	-0.00011	-0.00007	0.00001	0.00003	-0.00014	0.00004	0.00079	0.00088	0.00065	0.00176
	(0.00015)	(0.00017)	(0.00012)	(0.00012)	(0.00042)	(0.00037)	(0.00101)	(0.00104)	(0.00126)	(0.00112)
N	1,396,081	1,026,628	1,396,081	1,343,027	1,396,081	1,524,119	1,396,081	1,026,628	1,396,081	1,178,589

Note: Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Table 1.15: Behavioral Responses by Cancer Screening Result

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Gender	Male					Female					
Dependant variable	Future screening take-ups		Medical Expenditure			Future screening take-ups			Medical Expenditure		
	Health	Stomach ca	Year 1-2	Year 3-5	Year 1-5	Health	Stomach ca	Breast ca	Year 1-2	Year 3-5	Year 1-5
Eligibility	-0.0825**	-0.0464**	4.0	-12.9	1.1	-0.0770**	-0.0696**	-0.0688**	20.2*	-25.7	-10.4
	(0.020)	(0.010)	(13.0)	(27.8)	(35.3)	(0.010)	(0.011)	(0.011)	(9.7)	(21.9)	(28.9)
Eligibility * Stomach_Normal	0.0260*	-0.0258	16.7	20.4	27.6	-0.0358	-0.0129+	0.0102	6.9	-24.3	-32.0
	(0.012)	(0.020)	(24.8)	(53.2)	(64.9)	(0.024)	(0.007)	(0.009)	(27.2)	(60.9)	(83.4)
Eligibility * Stomach_Cancer	0.1344	0.0038	761.5	58.0	377.0	-0.0026	-0.0390	-0.0885	588.7	944.5	603.2
	(0.093)	(0.100)	(812.3)	(934.5)	(1546.1)	(0.154)	(0.093)	(0.168)	(761.9)	(1365.2)	(1806.1)
Eligibility * Stomach_False(+)	0.0786**	0.0591*	52.3	403.4	367.6	0.0053	0.0185	0.0350	-70.3	169.5	113.1
	(0.023)	(0.022)	(131.1)	(251.7)	(377.1)	(0.044)	(0.037)	(0.039)	(70.8)	(132.7)	(171.5)
Eligibility * Stomach_Other	0.0399**	-0.0290+	-38.2	-79.8	-129.9	-0.0311	-0.0202**	0.0127	-31.6	-51.1	-87.8
	(0.012)	(0.017)	(24.0)	(65.5)	(80.8)	(0.026)	(0.007)	(0.011)	(23.4)	(55.6)	(72.2)
Eligibility * Breast_Normal						-0.0694**	-0.0230	-0.0801**	53.8*	138.8**	192.9**
						(0.019)	(0.021)	(0.026)	(23.1)	(39.6)	(55.8)
Eligibility * Breast_Cancer						0.1759	0.0994	0.0887	-1106.1	-1638.4	-1378.2
						(0.169)	(0.181)	(0.165)	(1864.6)	(974.2)	(1555.7)
Eligibility * Breast_False(+)						0.0009	0.1047*	0.1326*	210.9	481.1	651.2
						(0.070)	(0.049)	(0.052)	(192.9)	(518.8)	(622.8)
Eligibility * Breast_Other						-0.0600**	-0.0089	-0.0620**	31.3	231.5**	253.1**
						(0.018)	(0.019)	(0.022)	(21.2)	(36.8)	(50.8)
Constant	0.9711**	0.3727**	342.2**	1,492.6**	1,785.8**	0.7282**	0.4343**	0.4561**	287.3**	1,414.7**	1,681.3**
	(0.012)	(0.009)	(12.710)	(26.096)	(32.255)	(0.008)	(0.011)	(0.010)	(8.2)	(17.5)	(25.0)
N	1,260,729	1,260,729	970,661	905,003	841,751	1,396,081	1,396,081	1,396,081	1,219,294	1,164,312	1,123,043

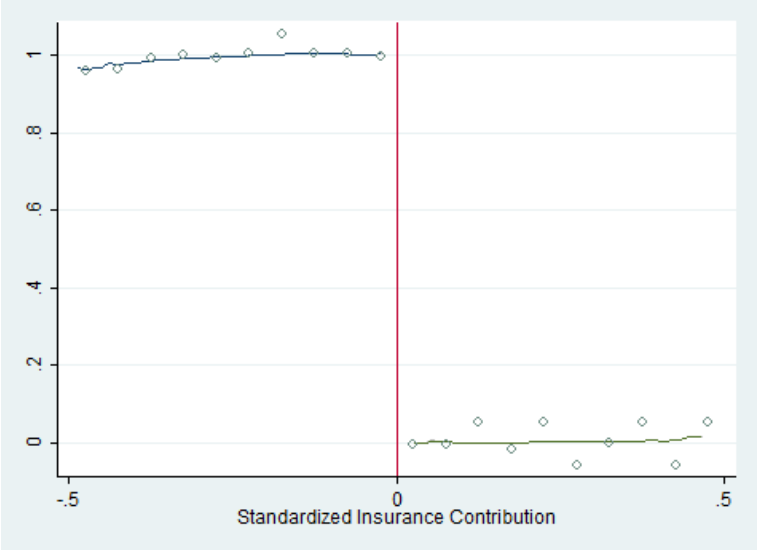
Note: The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively. N is number of observations

Table 1.16: Cost for cancer detection

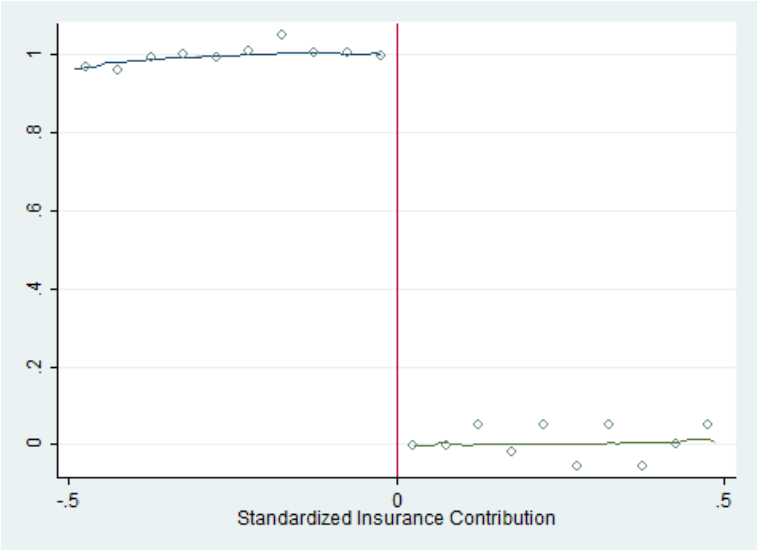
	(1)	(2)	(3)	(4)
Panel A. Male, Stomach cancer	N	Unit cost ($)	Total cost ($)	Proportion
Panel A1. Direct screening cost				
Administration	180.0	3.8	679	0.05
UGI	117.0	33.8	3,959	0.26
EGD	63.0	33.9	2,137	0.14
Biopsy	63.2	22.9	1,448	0.10
Conscious sedation	18.9	30.0	567	0.04
Subtotal			8,790	0.58
Panel A2. Transportation cost				
1st visit	180.0	17.9	3,225	0.21
Follow-up visit	8.0	17.9	143	0.01
Subtotal			3,367	0.22
Panel A3. Opportunity cost				
Loss of labor productivity	94.0	31.0	2,916	0.19
Total			15,073	1.00
Panel B. Female, Stomach cancer				
Panel B1. Direct screening cost				
Administration	808.9	3.8	3,049	0.05
UGI	547.9	33.8	18,541	0.31
EGD	261.1	33.9	8,855	0.15
Biopsy	295.9	22.9	6,778	0.11
Conscious sedation	78.3	30.0	2,349	0.04
Subtotal			39,573	0.66
Panel B2. Transportation cost				
1st visit	808.9	14.5	11,714	0.20
Follow-up visit	18.5	14.5	268	0.00
Subtotal			11,982	0.20
Panel B3. Opportunity cost				
Loss of labor productivity	259.2	31.0	8,035	0.13
Total			59,590	1.00
Panel C. Female, Breast cancer				
Panel C1. Direct screening cost				
Administration	646.0	3.8	2,435	0.04
Mammography	646.0	16.2	10,457	0.16
Biopsy	228.0	27.6	6,297	0.10
Ultrasonography	228.0	100.0	22,796	0.36
Subtotal			41,985	0.66
Panel C2. Transportation cost				
1st visit	832.7	14.5	12,058	0.19
Follow-up visit	60.0	14.5	869	0.01
Subtotal			12,927	0.20
Panel C3. Opportunity cost				
Loss of labor productivity	287.0	31.0	8,898	0.14
Total			63,811	1.00

Figure 1.1: Eligibility for Free Public Cancer Screening

Panel A. Eligibility in males



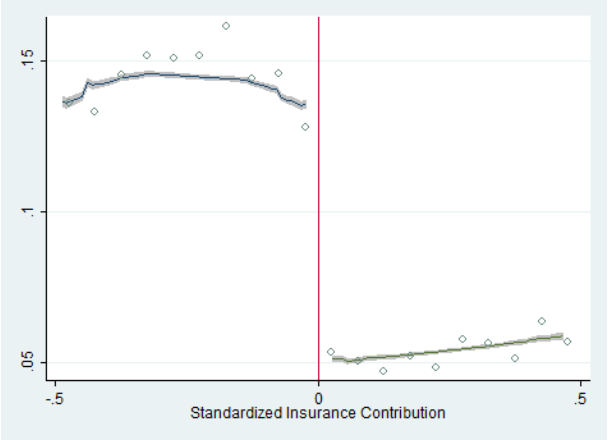
Panel B. Eligibility in females



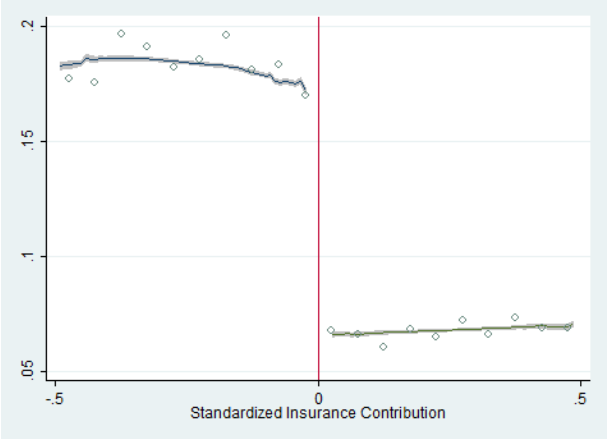
Note: The dependent variables in Panels A and B are eligibility for free public cancer screening in males and females, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. Y-axis is based on residuals from a regression (1) with a standard set of control variables. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals. Dependant variable is an indicator for eligibility.

Figure 1.2: Effect of Cost-Sharing on Cumulative Public Cancer Screening Take up, up to 2 years

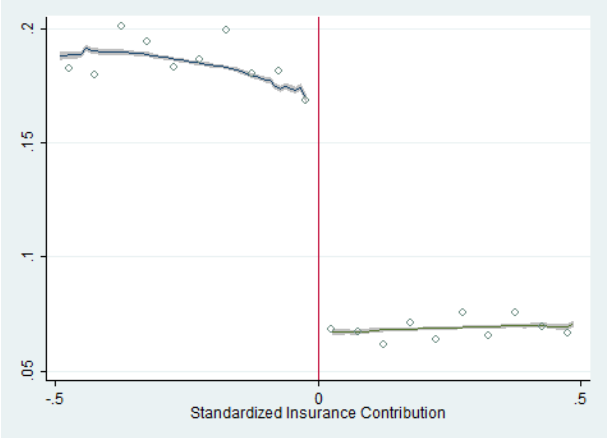
Panel A. Stomach cancer take up, Male



Panel B. Stomach cancer take up, Female



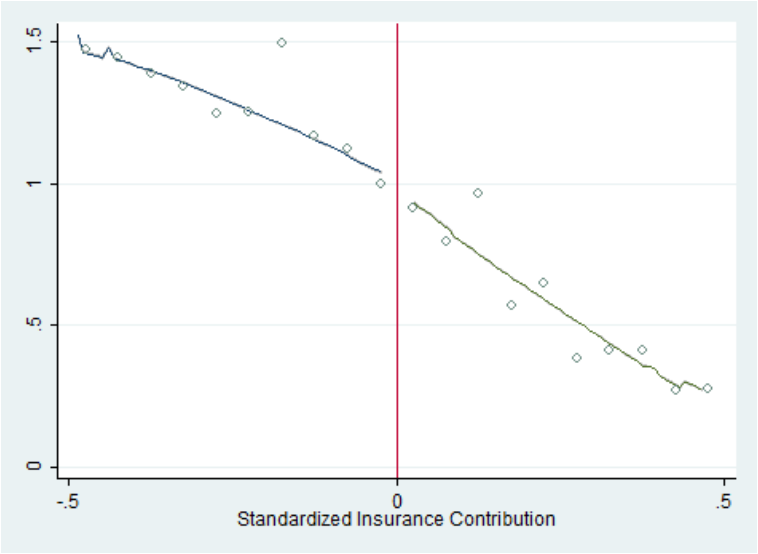
Panel C. Breast cancer take up, Female



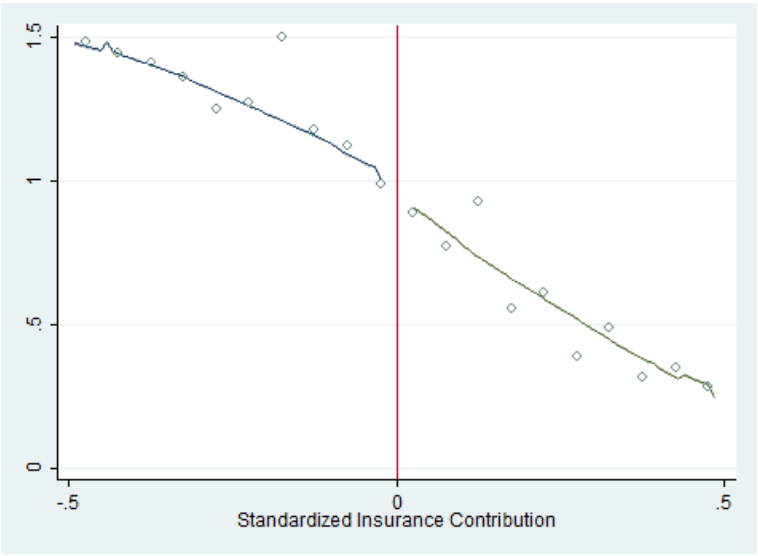
Note: The dependent variables in Panels A, B, and C are male stomach, female stomach, and female breast cancer take ups, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.3: Effect on Future Eligibility for Free Public Cancer Screening, 3 to 6 Years after Screening Offer

Panel A. Future eligibility in males



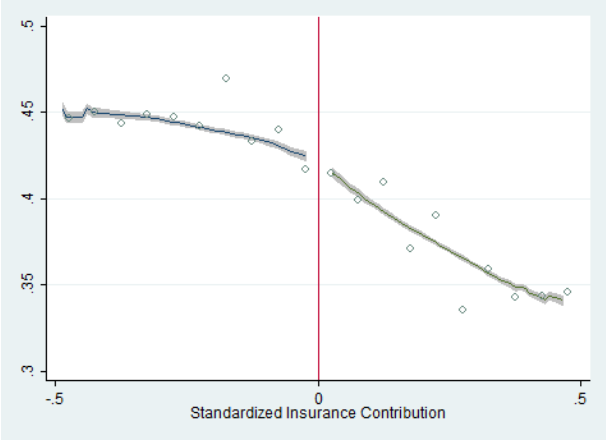
Panel B. Future eligibility in females



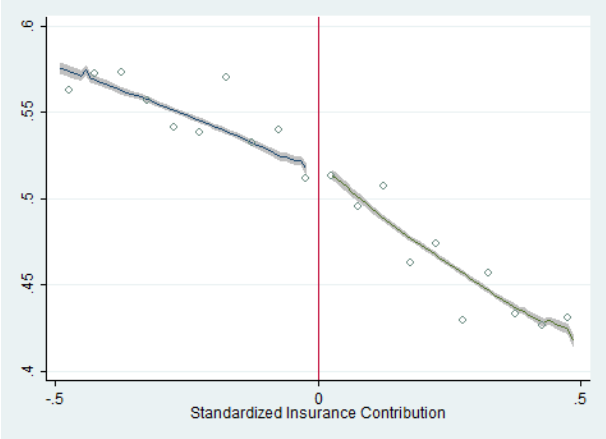
Note: The dependent variables in Panels A and B are future eligibility for free public cancer screening in males and females, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. Y-axis is based on residuals from a regression (1) with a standard set of control variables. The shaded regions are 95 percent confidence intervals. Dependant variable is an indicator for eligibility.

Figure 1.4: Effect on Future Cancer Screening Take up, 3 to 6 Years after Screening Offer

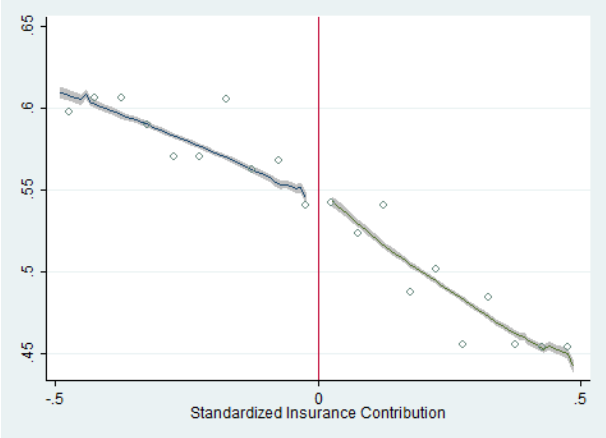
Panel A. Future stomach cancer take up, Male



Panel B. Future stomach cancer take up, Female



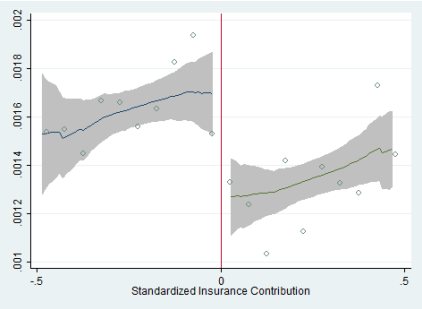
Panel C. Future breast cancer take up, Female



Note: The dependent variables in Panels A, B, and C are future male stomach, female stomach, and female breast cancer take ups, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

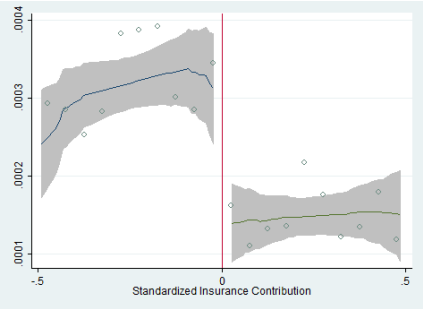
Figure 1.5: Effect on Cumulative Cancer Detections, up to 2 years

Panel A. Cancer detection by public mass cancer screening



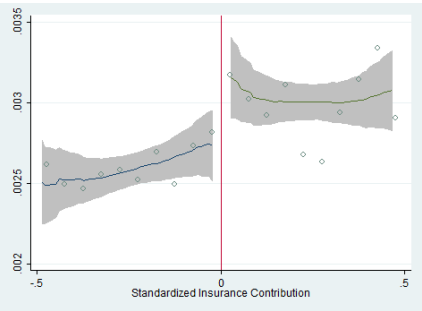
Stomach cancer, Male

Stomach cancer, Female

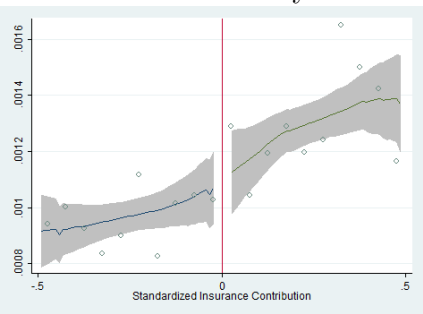


Breast cancer, Female

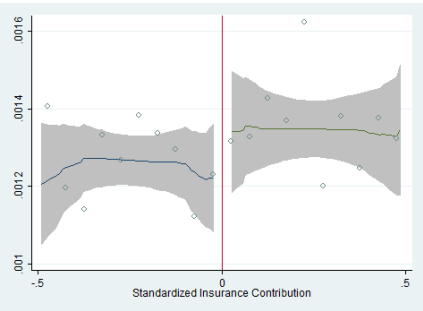
Panel B. Cancer detection by other channels



Stomach cancer, Male

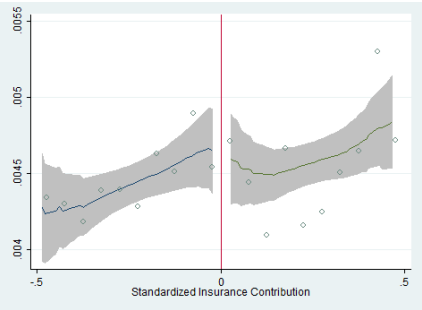


Stomach cancer, Female

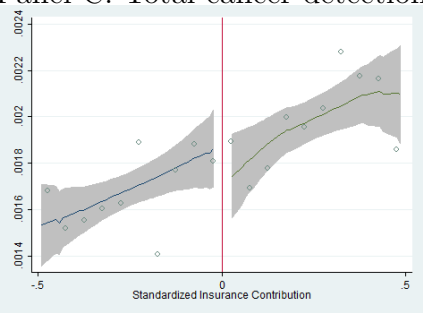


Breast cancer, Female

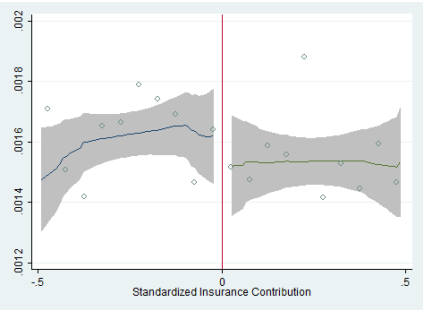
Panel C. Total cancer detection



Stomach cancer, Male



Stomach cancer, Female

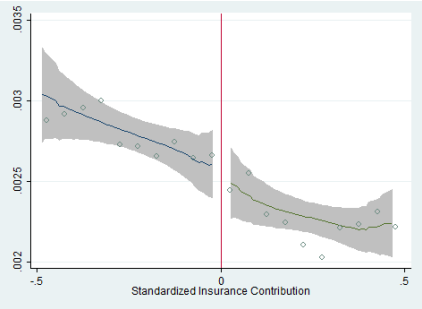


Breast cancer, Female

Note: The dependent variables in Panels A, B, and C are 2-year cumulative cancer detections by public cancer screening, by other channels, and overall cancer detections, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

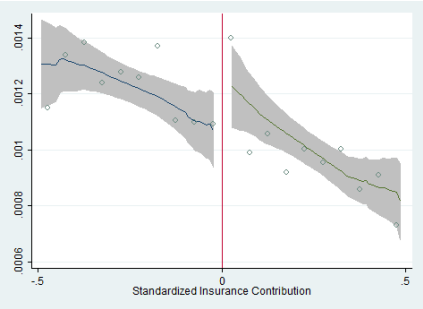
Figure 1.6: Effect on Future Cancer Detections, 3 to 6 Years after Screening Offer

Panel A. Cancer detection by public mass cancer screening



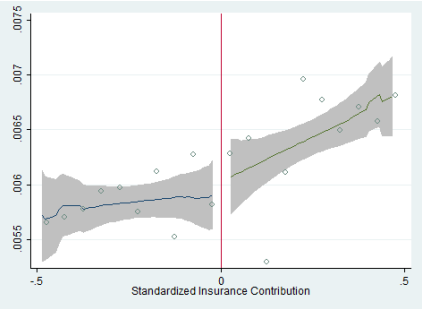
Stomach cancer, Male

Stomach cancer, Female

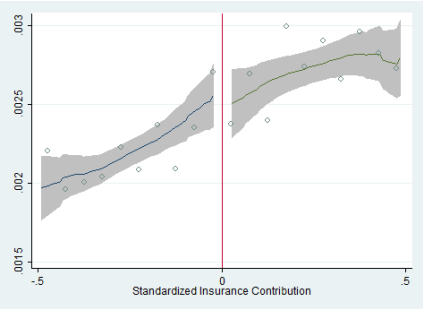


Breast cancer, Female

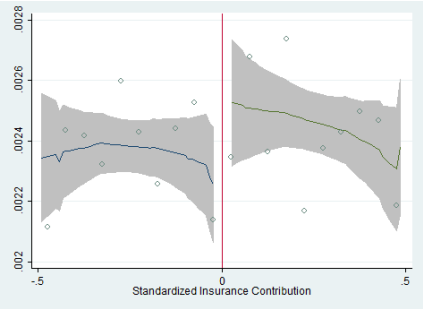
Panel B. Cancer detection by other channels



Stomach cancer, Male

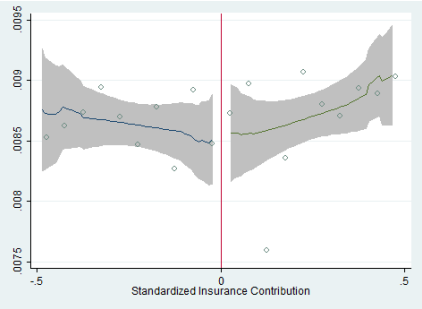


Stomach cancer, Female

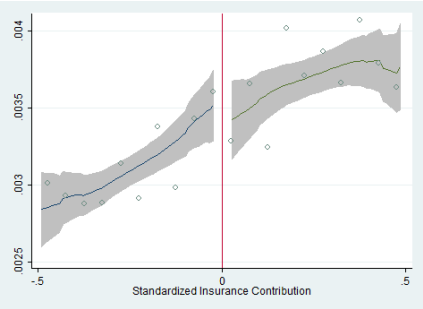


Breast cancer, Female

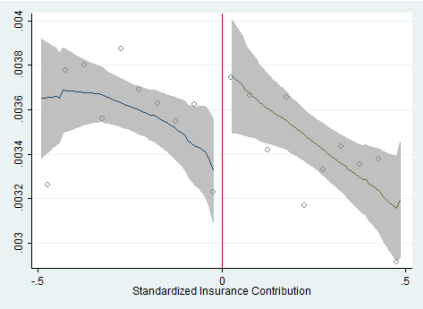
Panel C. Overall cancer detection



Stomach cancer, Male



Stomach cancer, Female

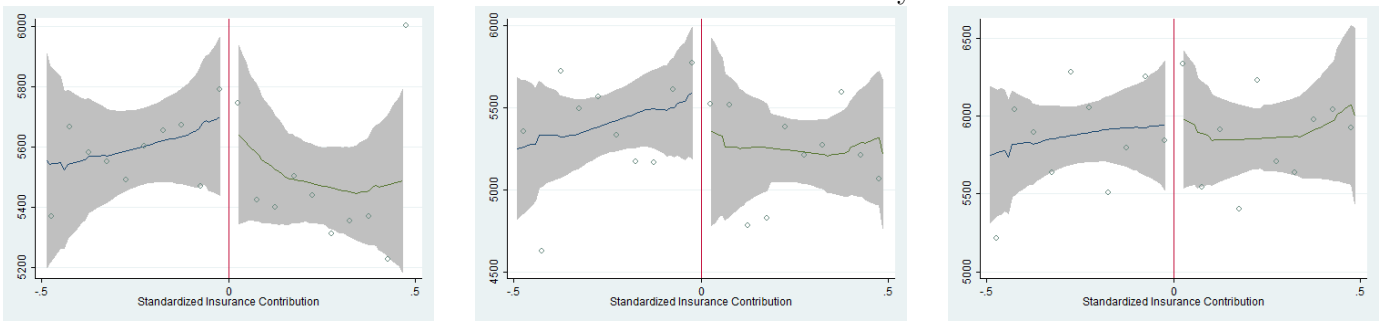


Breast cancer, Female

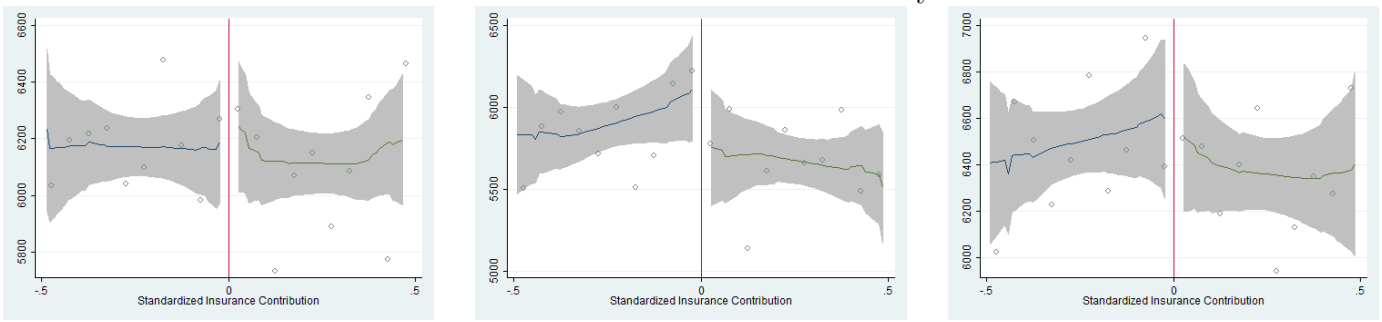
Note: The dependent variables in Panels A, B, and C are cumulative cancer detections between year 3 and 6 by public cancer screening, by other channels, and overall cancer detections, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.7: Effect on Medical Expenditure in the First Year of Cancer Detection (Early Detection)

Panel A. Cancer detected within 2 years



Panel B. Cancer detected after 3 to 6 years

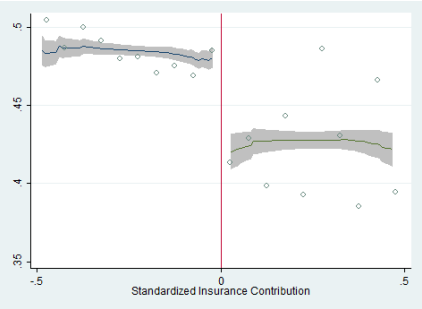


Stomach cancer, Male | Stomach cancer, Female | Breast cancer, Female

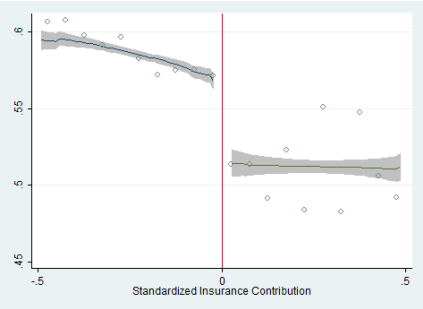
Note: The dependent variable in Panels A and B are the medical expenditure in the first year of cancer detection in year 1-2 and 3-6, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.8: Selection Effect by Cost-sharing (Screening Results): Among Screening-takers

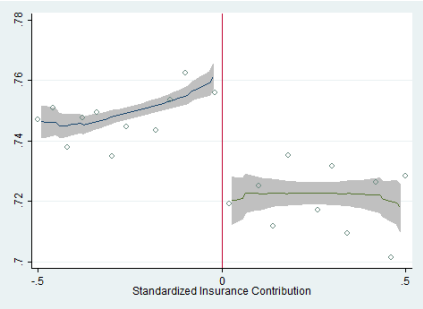
Panel A. Probability of being normal



Stomach cancer, Male

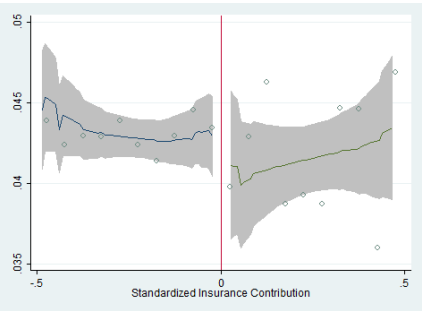


Stomach cancer, Female

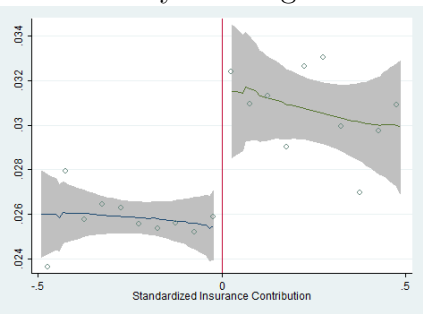


Breast cancer, Female

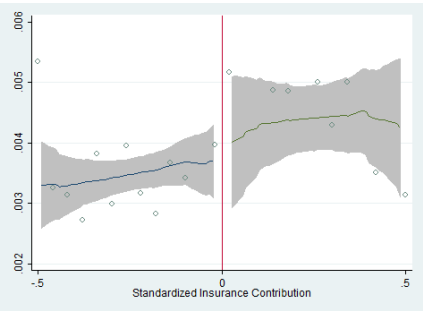
Panel B. Probability of being cancer suspicion



Stomach cancer, Male

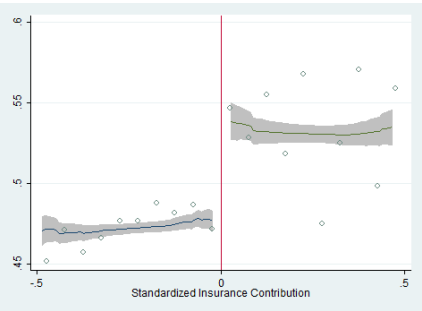


Stomach cancer, Female



Breast cancer, Female

Panel C. Probability of detecting other disease



Stomach cancer, Male

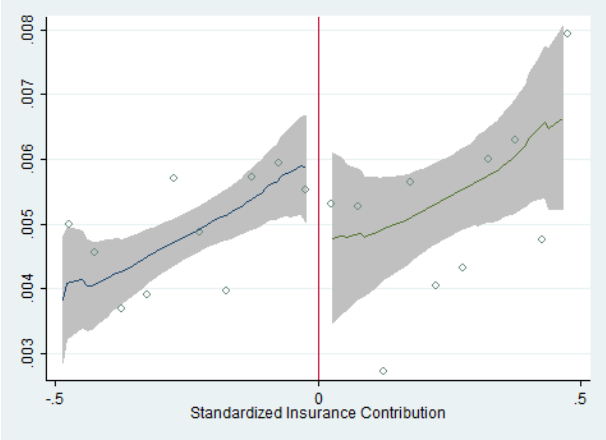
Stomach cancer, Female

Breast cancer, Female

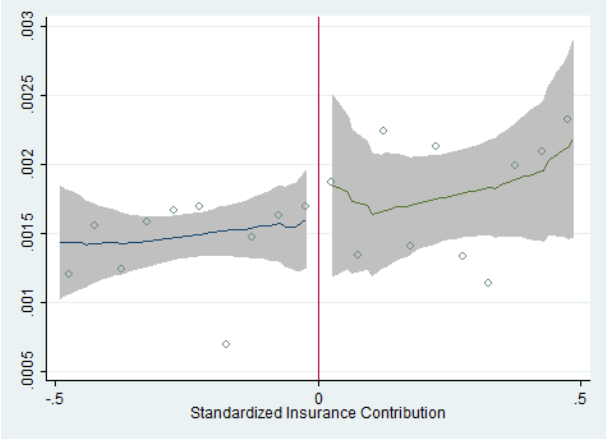
Note: The sample is restricted to screening takers. Dependent variables in Panels A, B, and C are probability of being normal, being cancer suspicion, and detecting other disease, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.9: Selection Effect by Cost-sharing (Cancer Detection): Among Screening-takers

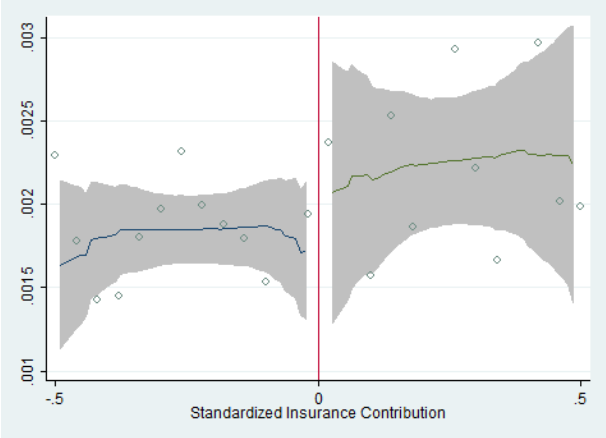
Panel A. Stomach cancer detection, Male



Panel B. Stomach cancer detection, Female



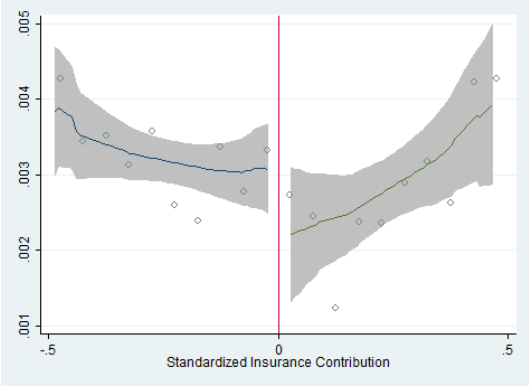
Panel C. Breast cancer detection, Female



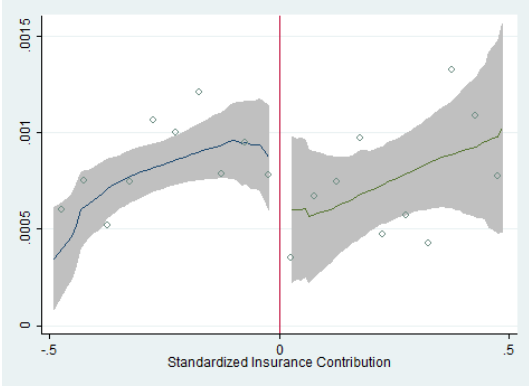
Note: The sample is restricted to screening takers. Dependent variables in Panels A, B, and C are male stomach, female stomach, and female breast cancer detections, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.10: Compliers vs. Always takers vs. Never takers: 6-Year Cumulative Cancer Mortality

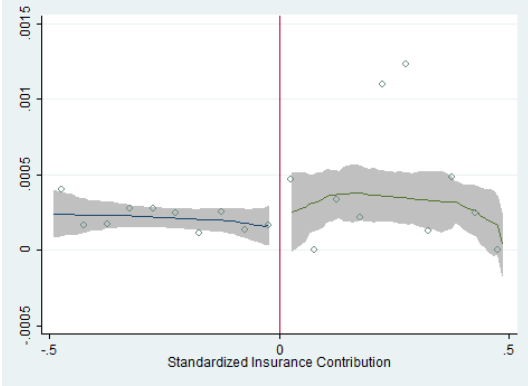
Panel A. Compliers vs. Always takers, Among screening takers



Male, Stomach cancer

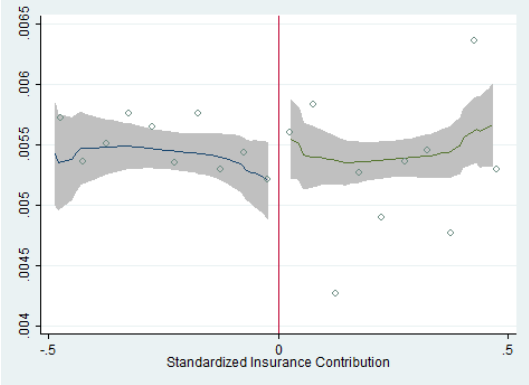


Female, Stomach cancer

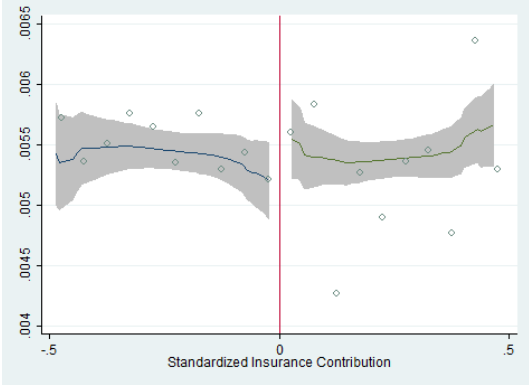


Female, Breast cancer

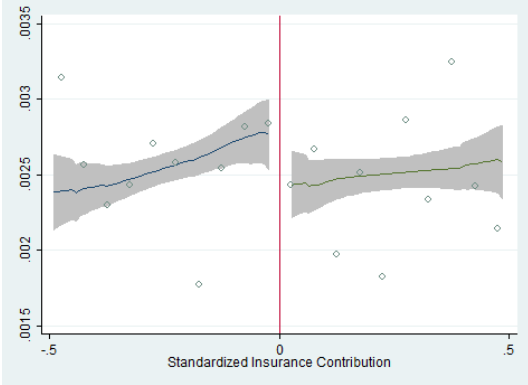
Panel B. Never takers vs. Compliers, Among screening non-takers



Male, Stomach cancer



Female, Stomach cancer



Female, Breast cancer

Note: The samples in Panel A and B are restricted to screening takers and screening non-takers, respectively. Dependent variable is 6-year cumulative cancer mortality. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It is estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.11: Effect on Future General Health Screening Take ups

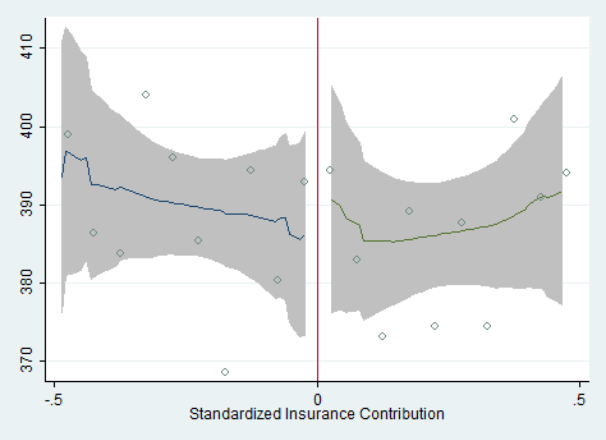
Panel A. Male

Panel B. Female

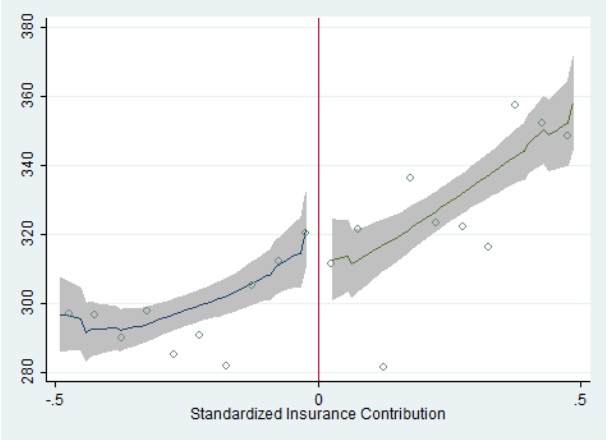
Note: The dependent variables in Panels A and B are the number of general health screening take ups in year 3-5 in males and females, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.12: Effect on Medical Expenditure

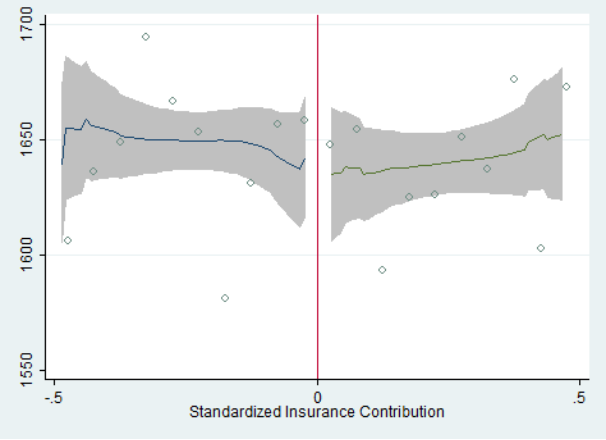
Panel A. In year 1-2, Male



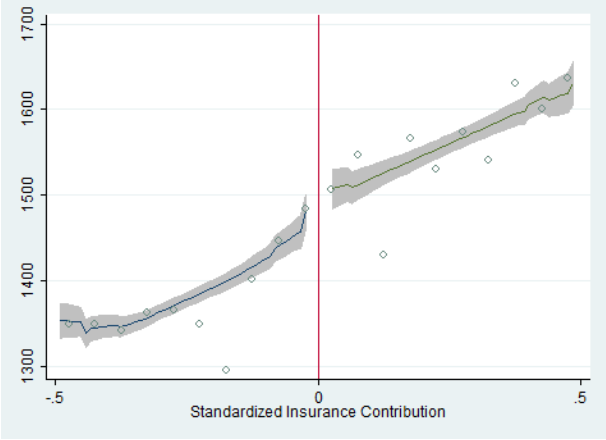
Panel B. In year 1-2, Female



Panel C. In year 3-5, Male



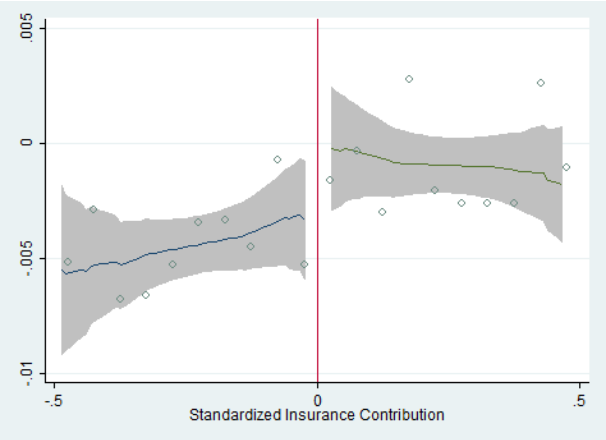
Panel D. In year 3-5, Female



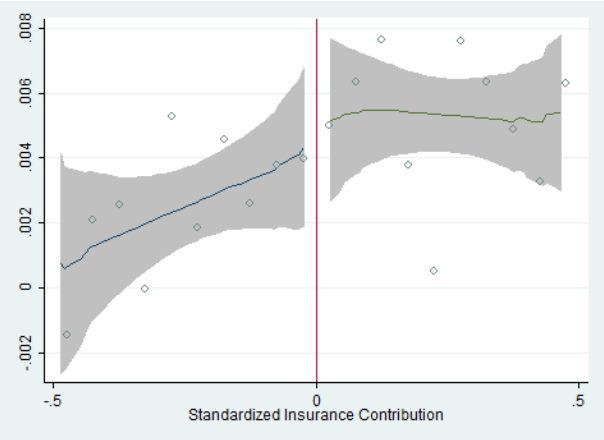
Note: The dependent variables in Panels A and B are medical expenditure in year 1-2 for males and females, respectively. The dependent variables in Panels C and D are medical expenditure in year 3-5 for males and females, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.13: Effect on Intermediate Health Outcomes, Male

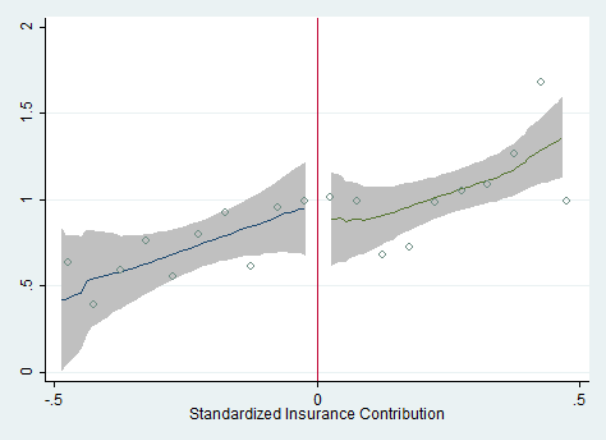
Panel A. Obesity, Round 2



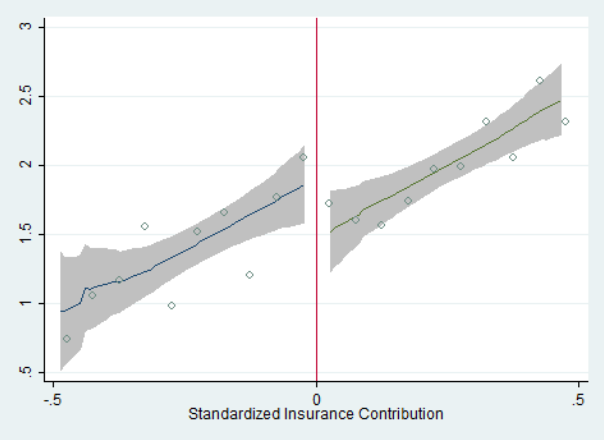
Panel B. Obesity, Round 3



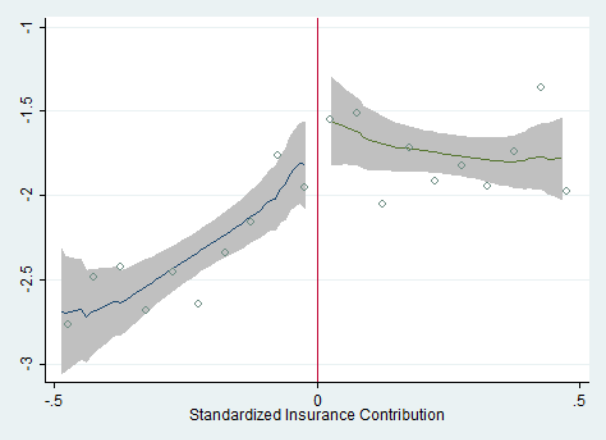
Panel C. Blood Sugar, Round 2



Panel D. Blood Sugar, Round 3



Panel E. Cholesterol level, Round 2

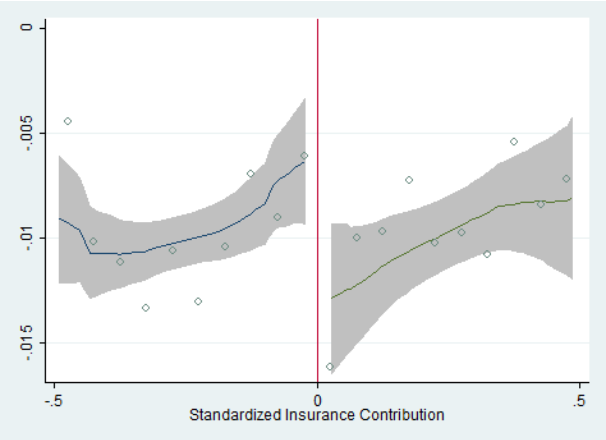


Panel F. Cholesterol level, Round 3

Note: The dependent variables are probability of being obese (BMI≥25), blood sugar level, and cholesterol level, respectively, in males. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

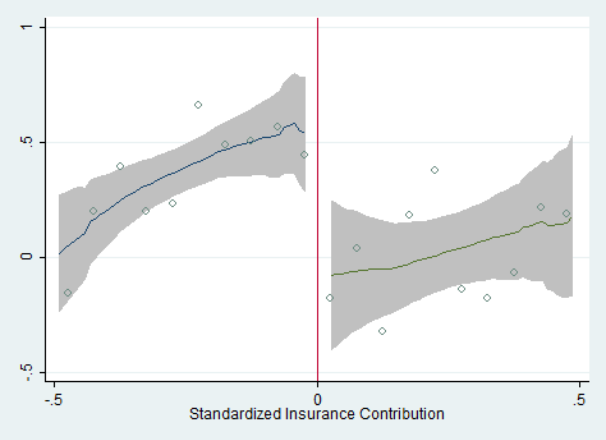
Figure 1.14: Effect on Intermediate Health Outcomes, Female

Panel A. Obesity, Round 2

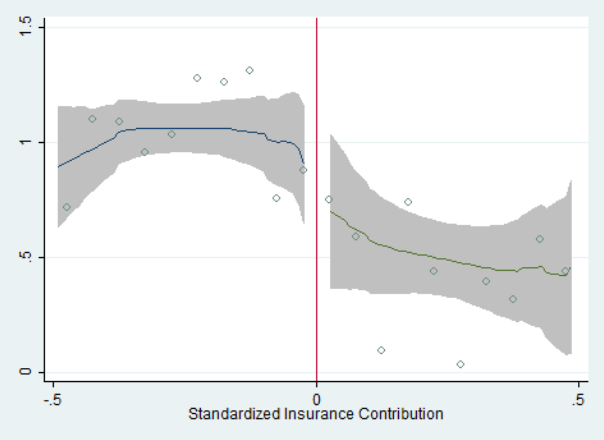


Panel B. Obesity, Round 3

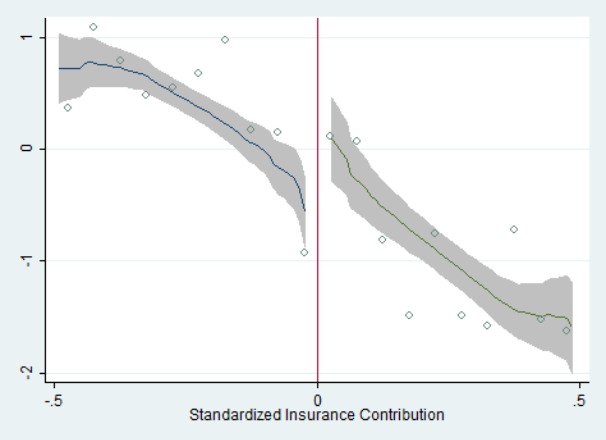
Panel C. Blood Sugar, Round 2



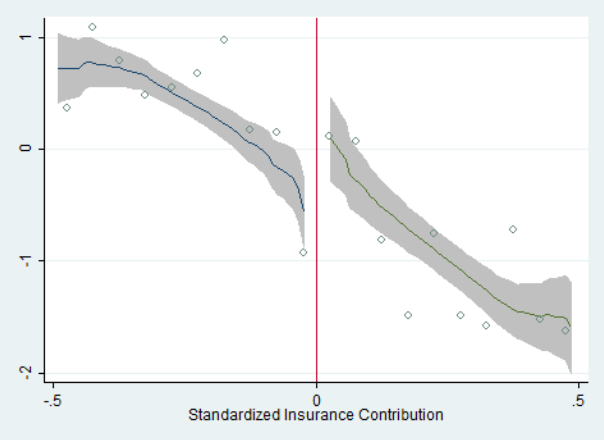
Panel D. Blood Sugar, Round 3



Panel E. Cholesterol level, Round 2



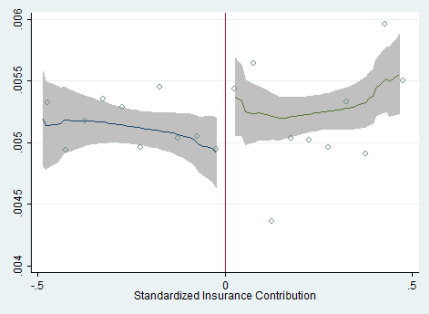
Panel F. Cholesterol level, Round 3



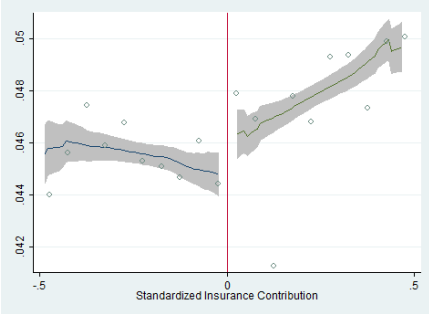
Note: The dependent variables are probability of being obese (BMI≥25), blood sugar level, and cholesterol level, respectively, in females. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.15: Effect on 6 year Cumulative Mortality

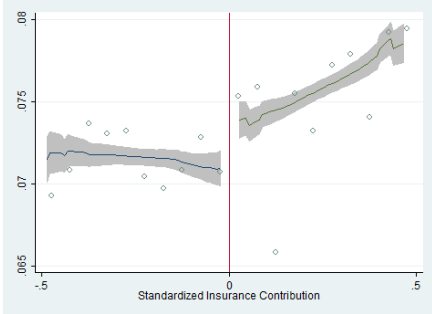
Panel A. Male



Stomach cancer mortality

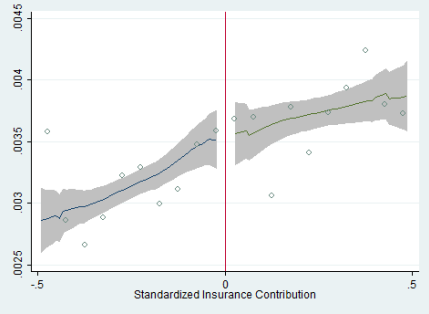


Non-cancer mortality

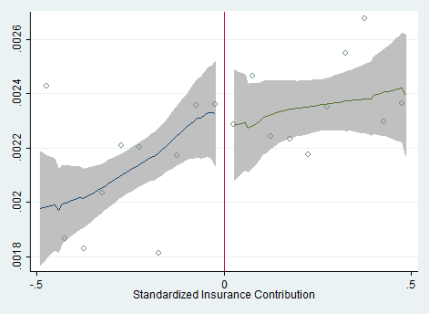


All-cause mortality

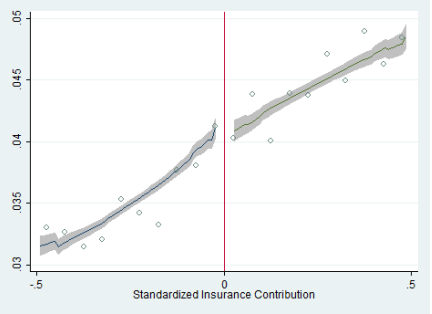
Panel B. Female



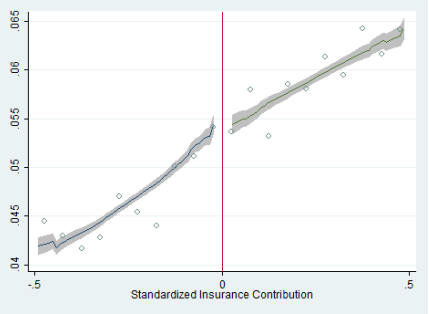
Stomach cancer mortality



Breast cancer mortality



Non-cancer mortality

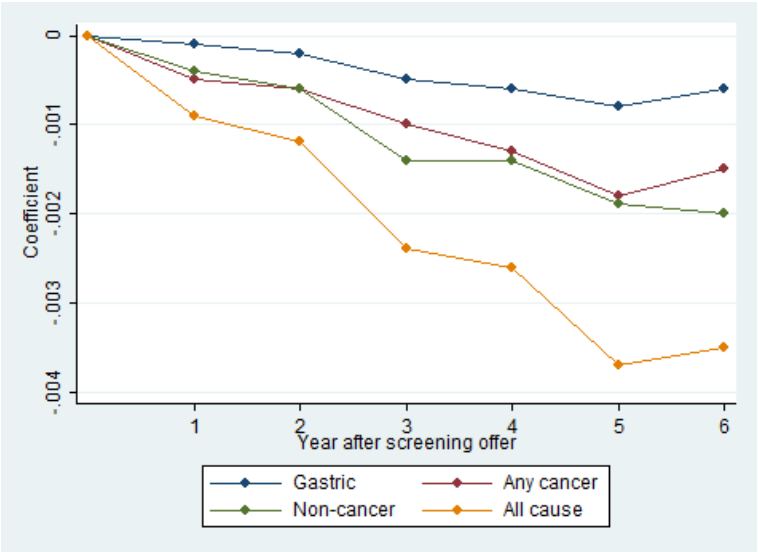


All-cause mortality

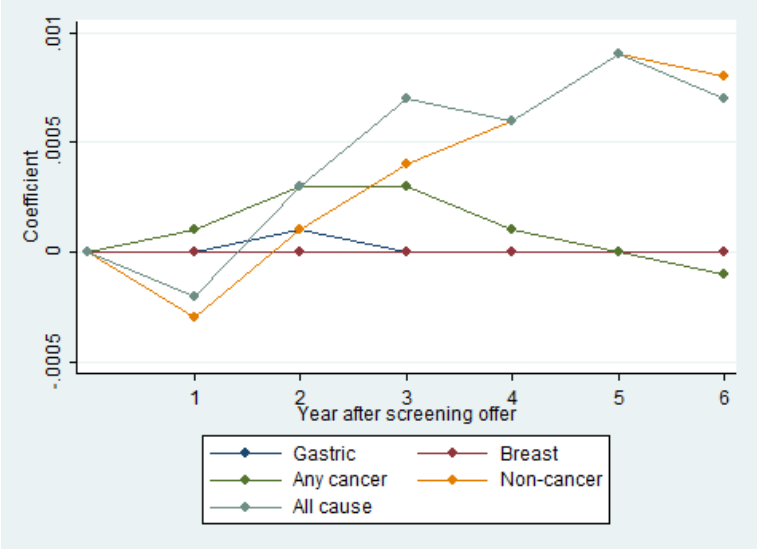
Note: The dependent variables in Panels A and B are 6-year cumulative cancer, non-cancer and all-cause mortalities in males and females, respectively. The running variable is the standardized insurance contribution. Open circles plot the mean of the dependent variable within 0.05 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 0.3 points. It estimated separately on each side of the cutoff. The shaded regions are 95 percent confidence intervals.

Figure 1.16: Effect on Mortality (trend)

Panel A. Stomach cancer, Male



Panel B. Stomach cancer, Female



Note: This figure illustrates a changing trend mortalities over time. Each dot represents a coefficient β from different local linear regression of equation (1.1). None of the estimates is statistically significant.

Chapter 2

Long-Term Care insurance, Informal Care, and Medical Expenditure

with Wilfredo Lim

2.1 Introduction

As countries face rapidly aging populations and rising healthcare costs, policies affecting long-term care—services targeting health or personal needs for people with chronic illness or disability—become increasingly important. For example, the share of those age 65 and over in the United States is expected to increase from 13.0% in 2010 to 20.2% in 2050. For Korea, the corresponding shares are 16.5% and 38.2%.[1] Moreover, the shares of those age 80 and over, for whom the need for long-term care is highest, are expected to double from 3.7% to 7.4% in the United States and increase severalfold from 1.9% to 14.5% in Korea.[2] At the same time, societal changes such as declining family size and rising female labor force participation are likely to reduce the availability of family caregivers. In terms of costs, public and private spending on long-term care in the U.S. totaled $183 billion in 2003, or 1.6% of GDP (GAO (2005)). Moreover, a third of Medicaid spending in 2006 went towards long-term care (CBO (2007)).

Much of long-term care is provided informally. As needs expand and costs rise, understanding the role of informal care in meeting this escalating demand becomes increasingly important. This paper aims to shed light on an important aspect—the substitutability of formal for informal care. For example, if formal long-term care services directly substitute for—rather than supplement—informal care, the cost of provision will rise without necessarily increasing the total care received by disabled persons. This could have welfare consequences for the caregivers in terms of their participation in the labor force as well as on intergenerational household bargaining. Thus, understanding the welfare impacts will require understanding under what situations and through which services formal care substitutes for informal care. Additionally, as governments develop and refine long-term care policies, implications for economic efficiency will be substantial. Informed policies will need to assess the costs and benefits of subsidizing various types of care—in particular, home versus facility—measured both by direct costs of subsidization as well as potential costs or savings from other medical expenses.

In this paper, we study subsidies for formal home and facility care and their corresponding

[1] Data are from http://dx.doi.org/10.1787/888932400893, in Colombo et al. (2011).

[2] Data are from http://dx.doi.org/10.1787/888932400874, in Colombo et al. (2011).

impacts on informal caregiving and medical expenditures in Korea. This study has a number of advantages that allow us to address this topic and improve upon the existing literature. First, we account for endogeneity in the choice of long-term care by using plausibly exogenous variation induced by a regression discontinuity design. Specifically, long-term care benefits in Korea are assigned based on an assessment score that is very difficult to precisely control. Second, these benefits vary at multiple cutoffs which allow us to separate the impact of home and institutional care benefits. Specifically, the first set of thresholds isolates the impact of just home care benefits; the second set of thresholds allows us to compare home or institutional care benefits versus just home care benefits; and the third set of thresholds allows us to look at the impact of an increase in the relative price of institutional care. Third, our analysis benefits from unique administrative data on formal home and institutional care, informal care, and medical expenditures.

Our main finding is that the benefits of home and facility care are heterogeneous across physical function level and therefore setting policy appropriately has the potential to dramatically reduce medical expenses. Specifically, substantial reductions in medical expenses arise from incentivizing transitions from home to facility care among people who are partially dependent for several activities of daily living, while incentivizing transitions from facility to home care among people who are completely dependent for several ADLs. This finding is not likely culturally or context specific and is consistent with programs in the U.S. such as Money Follows the Person that seeks to transition people with Medicaid from institutions to the community. We also find that formal long-term care is not a strong substitute for informal long-term care at the extensive margin, but find evidence that it does so at the intensive margin. Indeed, given that Korea is a "strong family ties" country, we argue that our results constitute a lower bound for similar effects in the U.S.

Specifically, we find that among individuals who are partially dependent for some activities of daily living[3], government subsidies for formal home care lead to an increase in its utilization, with no impact on informal caregiving at the extensive margin, as measured by child caregiving and independent living. We do find evidence for a reduction at the intensive margin, measured by the

[3]Partially dependent for some ADL's roughly corresponds to needing assistance moving around; partially dependent for several ADL's roughly corresponds to being unable to move on one's own; completely dependent for several ADLs roughly corresponds to being bedridden. See Table 2.1.

use of short-term respite care, which provides temporary relief for the receipient's caregiver. We also find no impact on medical expenses. Among individuals who are partially dependent for several activities of daily living, reimbursement of institutional care leads to an increase in its utilization with corresponding reductions in informal caregiving and medical expenses. Among individuals who are completely dependent for several activities of daily living, we find that an increase in the relative price of institutional care leads to substitution of home care for institutional care. We find no impact on informal caregiving, but we find substantial decreases in medical spending, largely accounted for by a reduction in hospital expenses. From a policy perspective, these findings suggest that among more able individuals, home care may be reduced with minimal detriment to their health; and that among the less able, incentives to transition from facility to home care may improve quality of life and reduce program costs overall.

We explore additional mechanisms for explaining our findings. First, we determine whether crowd out explains our lack of findings on informal care. While we find that public long-term care insurance leads to partial crowd out of private spending on long-term care, long-term care utilization still increases overall. Thus, crowd out is not likely the sole reason for our limited findings on informal care. We also assess the impact of long-term care insurance on short run-mortality, as this measure is important in and of itself and in order to rule out differential mortality in affecting our estimates. We find no statistically significant differences in mortality across all thresholds. Lastly, we show that our results are robust to various checks and specifications of our estimation strategy.

The remainder of this paper is structured as follows. Section 2.2 provides a brief discussion of the literature and our contribution. Section 2.3 explains Korea's Long-Term Care Insurance program and motivates our empirical strategy. Section 2.4 describes the data. Section 2.5 provides a conceptual framework for considering the impacts of subsidies for long-term care. Sections 2.6 and 2.7 present the empirical framework and results, followed by additional robustness checks in Section 2.8. Section 2.9 provides a brief discussion and Section 2.10 concludes.

2.2 Literature Review

Much of the empirical work on understanding the substitutability of formal for informal care is limited in scope and suffers from endogeneity concerns (papers that do not account for endogeneity include Soldo (1985), Wolinsky, Mosely, and Coe (1986), and Bass and Noelker (1987)). The concern with endogeneity is that absent an exogenous source of variation, confounding unobserved characteristics as well as the joint nature of the formal versus informal care decision may lead to misleading findings. For example, to the extent that formal and informal care are positively correlated with unobserved negative health shocks, a naive analysis would find them to be complements even if they were substitutes. Also problematic is that they may be substitutes in some instances and complements in others. For example, an individual may rely on a child caregiver for assistance with basic activities of daily living but may seek formal assistance for more skill-intensive needs such as physical therapy. This highlights the importance of being able to address endogeneity as well as account for different types of formal care and informal care.

One way to address the issue of endogeneity is through the use of instrumental variables. Using number of adult children and presence of a daughter who has no child at home as instruments, Lo Sasso and Johnson (2002) find that frequent help from children for basic personal care reduces the likelihood of future nursing home use. Using number of children and whether the eldest child is a daughter as instruments, Van Houtven and Norton (2004) find that informal care reduces home health care and nursing home use. Using children's gender, marital status, and distance as instruments, Charles and Sevak (2005) find that receipt of informal home care reduces the probability of future nursing home use. However, it is unclear whether the necessary exclusion restrictions would be satisfied, given the complexity of fertility decisions and bargaining over intergenerational transfers. Thus, it is useful to assess the robustness of these results through studies based on more plausibly exogenous sources of variation.

The Balanced Budget Act of 1997 induced such a source of variation. This act led to regional variation in overall decreases in Medicare reimbursement for home care services. Using this source of variation, McKnight (2006) finds resulting reductions in home care utilization that were not offset by increases in institutional care or other medical care. She also finds no adverse health

consequences as a result of the policy. Using the same source of variation, Orsini (2010) and Engelhardt and Greenhalgh-Stanley (2010) find reductions in independent living, and Golberstein et al. (2009) find increases in the probability of the use of informal caregiving. A significant limitation these papers share, however, is that due to data limitations and their source of identification, they focus primarily on the provision of home care.

The Channeling demonstration in the U.S. provides another opportunity to assess the relationship between informal and formal care, through randomized evaluation. This experiment sought to substitute a system of home and community care for institutional care. Christianson (1988) and Pezzin, Kemper, and Reschovsky (1996) assess the impact of public home care provision and find limited reductions in the care provided by informal caregivers. However, the latter paper does find a significant increase in the probability that unmarried persons live independently. This highlights the importance of considering both informal caregiving directly and independent living. The results of the Channeling demonstration are limited, however, in that the sample population was particularly frail and the scope was inherently limited to the provision of home care, not institutional care.

Regarding impacts on other medical expenditures, McKnight (2006) finds that reductions in home health care reimbursement and utilization did not lead to increases in other medical care and were not associated with adverse health consequences. Evaluating the impact of Channeling on other medical expenses, Wooldridge and Schore (1988) find large reductions in nursing home use among those who were already in a nursing home at baseline but no impact on the use of hospital, physician, and non-physician medical services.

Our view of the literature is that evidence on the substitutability of institutional and informal care is very limited and is based mostly on observational studies. Moreover, even though understanding the impact of institutional care on health and other medical expenses is necessary for cost-benefit analyses, very little is known at this point.[4] In addition, while there is more work on the substitutability of home and informal care, this evidence is limited in accounting for institutional care and in being generalizable to a broader population of the elderly. This study attempts

[4]In a review paper, Ward et al. (2008) conclude "there is insufficient evidence to compare the effects of care home environments versus hospital environments or own home environments on older persons rehabilitation outcomes."

to fill these gaps directly. By using longitudinal administrative data with measures of home care, institutional care, informal care, and medical expenditures, and a unique policy affecting a broader population of the elderly, we are able to account for the complex interrelationship among informal, home, and institutional care, as well as evaluate the corresponding impacts on health and medical expenses.

Lastly, much of the literature is based on findings in the United States and other Western countries. Other studies outside the U.S. include Stabile et al. (2006) for Canada and Bolin et al. (2008) for Europe. This paper contributes to this literature by providing evidence from an Asian country.

2.3 Background

Korea implemented universal health coverage in 1989. Individuals are covered either by National Health Insurance (NHI) or Medical Care Assistance (MCA), though both programs are overseen by the National Health Insurance Corporation (NHIC). The primary distinction between NHI and MCA is that the latter serves poor individuals. While health insurance coverage includes outpatient care, inpatient care, and prescription pharmaceuticals, no coverage for long-term care is included. In response to this, and due to the demographic and cultural changes affecting the need and provision of long-term care, National Long-Term Care Insurance was implemented in July 2008. This provides coverage for individuals age 65 and over and those with age-related needs such as dementia and Parkinson's disease.

2.3.1 Benefits

Long-term care insurance covers two categories of service benefits: home care and institutional care.[5] Home care includes services provided at the beneficiary's residence. This includes home help where a caregiver provides support for physical activities or housework, home bathing where a caregiver assists the beneficiary in bathing, and home nursing where a nurse provides assistance

[5]In exceptional cases (e.g. for individuals who live in remote regions with no access to long-term care services), cash benefits are provided. However, this represents less than 0.2% of cases.

with such things as medication and dental hygiene. Also included within home care benefits is short-term respite care which covers a short-term stay in a facility to allow the caregiver relief from caregiving activities. Lastly, equipment for the support of daily tasks and physical activities (e.g. a wheelchair) is also included in home care benefits. Institutional care benefits cover long-term residence in a facility where meals, care, and other necessities required for daily function are provided. See Table B.3 for more details. As in the case for general health care, the delivery of long-term care is primarily administered through private providers.

2.3.2 Eligibility

To receive long-term care benefits, individuals must apply, submit a doctor's referral, and be evaluated by an assessment team from the NHIC. Benefits are determined based on an adjusted score, which is the sum of two components, a preliminary score and committee points. The preliminary score is a complex, highly nonlinear function of the responses to 52 evaluation questions, encompassing physical and cognitive function, behavior, nursing assistance, and rehabilitation.[6] Then a local assessment committee, following guidelines determined at the national level, is able to add or subtract up to five points to this score, based on the assessment questions and the doctor's referral.[7]

The adjusted score is used to determine benefits, as depicted in Table 2.1. Individuals who score below 55 are not eligible for long-term care benefits. Individuals who score 55 or above (Grade 3) are eligible for reimbursement of formal home care services up to 740 USD per month, which corresponds to approximately two hours of home help care per day.[8] Individuals who score 75 or above (Grade 2) become eligible for reimbursement of institutional care or a home care benefit maximum of 880 USD per month.[9] Individuals who score 95 or above (Grade 1) continue to be

[6] An example of a physical function question is whether the individual is fully independent, partially dependent, or fully dependent for bathing. For more details, including calculation of the preliminary score, see Appendix B.

[7] Committee members are trained annually and when the guidelines are changed.

[8] See Table 2.1 for general descriptions of individuals falling into each category. All amounts in this paper are converted to USD at the rate of 1100 KRW : 1 USD.

[9] If one were to use both types of care in the same month, the home care benefit would be prorated based on the number of facility days used. However, home and facility care are inherently incompatible with each other (in our

eligible for reimbursement of institutional care or a home care benefit maximum of 1040 USD per month. The price of institutional care is 40 USD per day and 45 USD per day for individuals in Grades 2 and 1, respectively. To the extent that there is a copayment, this implies that the cost of institutional care for an individual scoring 95 is discretely higher than the cost for an individual scoring 94.9. As a result, the increased cost of facility care along with the more generous home care benefit incentivizes individuals to transition from institutional to home care at the margin.

Applicants are notified of their classification, not their score. They are reevaluated when major changes to their physical or mental status occur, for the renewal of benefits, or if they appeal for a reevaluation.[10] Benefits must be renewed every twelve months, with the exception of those with significantly high scores (> 100) who may have up to eighteen months.

Figure 2.1 illustrates the committee component of the score in relation to the preliminary score. Note first that most activity occurs within 5 points of the actual thresholds (55, 75, 95).[11] Focusing on preliminary scores in the range [50,55) we see that some individuals are given enough points so that their adjusted scores exceed 55, leading to eligibility for Grade 3 home only benefits. It appears that points are rarely added or subtracted unless doing so changes the eligibility status. Focusing on scores just above 55, the number of instances where points are deducted is negligible. Focusing on scores below 50, we see that the number of instances where points are added is negligible, reflecting the fact that any additional points less than 5 would not be enough to become eligible. We find similar patterns in committee action around the remaining thresholds, except we see more instances of subtracted points.

Figure 2.2a illustrates from another perspective how the committee component of the score influences eligibility around the 55 threshold. It also highlights the source of identification in our research design. The probability that the adjusted (post-committee) score exceeds the 55 point threshold is plotted against the preliminary (pre-committee) score.[12] When the preliminary score

---

data, only 3% of individuals utilize both types of benefits in the same year). Thus, the use of both types of services in the same month is likely due to changes in health status as opposed to joint use.

[10] They are able to appeal indefinitely, though this process typically takes longer than one month.

[11] In practice, scores outside of five points from a threshold are less likely to be reviewed by the committee.

[12] See Section 2.6 for a discussion of the specification used to generate the figures.

is below 50, the probability that the adjusted score exceeds the 55 point threshold is effectively zero, consistent with the guideline that the maximum number of points that can be added is five. When the preliminary score is above 55, the probability that the adjusted score exceeds the 55 point threshold is effectively one, reflecting the rarity with which the committees subtract points around this threshold. Between 50 and 55, enough points are added to the preliminary scores of a fraction of individuals so that their adjusted scores exceed 55. Note that this illustration suggests an implicit threshold at 50 (and similarly at 70 and 90). That is, scores above the explicit threshold of 55 virtually guarantee eligibility; scores below the implicit threshold of 50 virtually exclude the possibility of eligibility.

Correspondingly, this figure illustrates the source of identification for our analysis—namely, comparing similar individuals who have different probabilities of treatment.[13] For instance, those with preliminary scores just below 50 have a probability of eligibility for home care benefits of zero. Those with preliminary scores just above 50 have a probability of about 8 percent. This allows us to use variation in the probability of eligibility in order to look at the impact of eligibility on reimbursed formal long-term care utilization and relevant outcomes, including independent living, informal caregiving, and medical expenditures. Moreover, the different grades of benefits afford us the possibility of studying several aspects of long-term care utilization. The 50 and 55 thresholds isolate the impact of home care benefits, while the 70 and 75 thresholds isolate the impact of home and institutional care benefits versus just home care benefits. The 90 and 95 thresholds allow us to look at the impact of an increase in the price of institutional care along with an increase in the maximum benefit for home care.

2.3.3 Financing

Long-term care insurance is financed by the government (20%), copayments (up to 20%), and insurance contributions. Insurance contributions were 0.21%, 0.24%, and 0.35% of wages in 2008, 2009, and 2010, respectively. Employers paid 50% of this amount. The copayment for home care services is 15% while that of institutional care is 20%, but the poor (MCA individuals) are exempt

[13] We discuss our empirical strategy more formally in Section 2.6.

from copayments, and individuals with certain conditions faced reduced copayments.[14]

2.4 Data

This study uses a merged dataset combining NHIC administrative data for National Long-Term Care Insurance (NLTCI) and National Health Insurance (NHI). The sample consists of 171,373 individuals who were assessed in 2008 and 2009. The NLTCI data spans 2009 and the first half of 2010 and contains information on gender, age, living and caregiving arrangements, preliminary and adjusted scores from the first eligibility assessment, and reimbursed long-term care utilization.[15] The NHI data spans 2007 through 2009 and contains annual total, hospital, outpatient, and pharmacy expenditures. Our main explanatory variable is the 2009 preliminary score. Our main measures of formal care are reimbursed home care expenditures and number of institutional care days. We measure home care in expenditures as an aggregate measure to capture the variety of home care services that are used. Our main measures of informal care are an indicator of whether a child is the primary caregiver and an indicator of whether the individual lives alone or with a spouse. The latter measure is our measure of independent living, consistent with the previous literature. Our main measures of medical expenditures are total medical and hospital expenses.[16]

Table 2.2 displays summary statistics by grade. All measures are at baseline, except for long-term care facility days and home care expenditures. ADL Index is a composite score based on activities of daily living questions from the assessment, with a higher number indicating less function. Individuals with lower grades are sicker as measured by the ADL Index, medical expenditures, and hospital days, and tend to have more resources as measured by insurance contribution and MCA percentage. Finally, sicker individuals are less likely to have a child caregiver and live independently.

---

[14]Individuals who face reduced copayments include the disabled, people with rare and incurable diseases, and the marginally poor.

[15]Because we only observe NLTCI data for the first half of 2010, our sample is reduced by approximately half when looking at informal care outcomes. Analysis of predetermined variables shows covariates are balanced in the reduced sample.

[16]These amounts are inherently exclusive of long-term care expenses.

2.5 Conceptual Framework

2.5.1 Household Responses to Public Long-Term Care Reimbursement

We adapt the model developed by Stabile, Laporte, and Coyte (2006) in order to determine what implications arise from public reimbursement for long-term care.

Consider a two person household consisting of an elderly care recipient and an informal caregiver (e.g. a child). Let household utility be

$$U(X, L, A)$$

where X represents market goods and services, L the leisure time of the caregiver, and A the care recipient's functional ability. The care recipient's ability is defined by the technology

$$A = A(C, H, F)$$

where C is time spent delivering informal care, H is formal home care, and F is institutional (facility) care. Time and financial constraints are satisfied if

$$P_X X + P_H(1 - s_H)H + P_F(1 - s_F)F + WC = V + W(T - L)$$

where P_X is the cost of market goods and services, P_H is the cost of formal home care, P_F is the cost of facility care, s is the relevant government subsidy (in other words, 1-copay), V is non-wage income, W is the cost of the caregiver's time, and T is the total time for leisure, caregiving, and labor market work. The household selects performance ability A so that the marginal benefit of greater ability is equal to the marginal cost of its production. The household cost-effectively selects H, F, and C in order to achieve ability A. L is selected so that the marginal benefit of leisure equals the marginal cost of foregone market goods and services.

We now illustrate the relevant intuition and predictions derived from the model (see Stabile, Laporte, and Coyte (2006) for a more extensive treatment). When an individual is ineligible for reimbursed benefits, she may pay out-of-pocket for H at price P_H. Grade 3 benefits provide a subsidy for H, reducing its effective price to $P_H(1 - s_H)$ up to the maximum level of benefits m_H.

This is depicted in Figure 2.4, where the isocost line rotates out as the price of H falls from P_H to $P_H(1 - s_H)$, up to the point where $H = m_H$. After this point, the price returns to P_H. Through an income effect, these benefits will increase the optimal level of A and lead to corresponding increases in C and H if these are normal inputs to its production. Because H is cheaper relative to C, the substitution effect will lead to increases in H but decreases in C. Thus, while Grade 3 benefits are predicted to lead to increases in A and H, the net impact on C is unclear.

Grade 2 benefits lead to both an increase in the maximum level of home benefits, m_H, as well as provide a subsidy for facility benefits, s_F. Note that home and facility care are effectively perfect substitutes in the production of A as they are inherently incompatible with each other.[17] This is reflected in Figure 2.5, where the isocost line rotates out as the effective price of F falls, and the individual chooses to utilize only F. To the extent that F and C are substitutes, this should lead to an increase in F and decreases in C and H. If the individual decides not to utilize F, then the impact of m_H on H would depend on the amount used with only Grade 3 benefits, as in shown in Figure 2.6. If the individual were using less than the maximum beforehand, there would be no impact on A, C, or H. If the individual were using the maximum, this would lead to a pure price effect, resulting in an increase in A and H, but a decrease in C. Therefore, we expect A and F to increase, but C to decrease. The impact on H is uncertain.

Grade 1 benefits lead to both an increase in the maximum level of home benefits, m_H, as well as an effective increase in the price of facility benefits, P_H, as discussed in Section 2.3.2. Thus, the impact of these benefits is a combination of the figures for previous benefits. We expect the increase in the relative price of F to entice some people to switch from F to H (reverse of Figure 2.5). Combined with an increase in m_H (Figure 2.6) we expect a decrease in F and increase in H. The impact on A is ambiguous, however, as the impact of the relative increase in P_F may not be offset by the increase in m_H. The impact on C is also ambiguous and depends again on whether H and C are substitutes or complements.

In summary, the model yields the following predictions for government reimbursement of long-term care:

[17] In our data, only 3% of individuals utilize both home and facility benefits in the same year, and this is likely due to changes in health status as opposed to joint use.

1. Grade 3 benefits lead to an effective price decrease in home care. As a result, we expect increases in ability and home care. The impact on informal caregiving will depend on whether home care and informal caregiving are substitutes or complements.

2. Grade 2 benefits lead to an effective price decrease in facility care and an increase in the maximum level of home care benefits. Thus, we expect increases in ability and facility care, and a decrease in informal caregiving. The impact on home care is ambiguous.

3. Grade 1 benefits lead to an effective price increase in facility care and an increase in the maximum level of home care benefits. Thus, we expect an increase in home care and a decrease in facility care. The impacts on ability and informal care is ambiguous.

2.6 Empirical Framework

We conduct a regression discontinuity analysis at the thresholds 50, 55, 70, 75, 90, and 95 of the preliminary score that exploit the discontinuous probabilities of eligibility resulting from the committee adjustment portion of the score. Specifically, the aim is to compare outcomes across individuals with similar characteristics but differing probabilities of eligibility for benefits.

The corresponding regression model we estimate is:

$$\text{outcome} = \beta \mathbb{1}\{S \geq \tau\} + f(S) + \gamma X + \epsilon, \tag{2.1}$$

where S is the preliminary score, $f(S)$ is a function of the score, τ is the relevant cutoff, and X is a set of control variables—age, gender, insurance dummies, region type dummies, and health insurance contribution (a proxy for income)—which serve to improve precision of the estimates.

In implementing the regression discontinuity design, an important consideration is the modeling of $f(S)$. One approach is to model it parametrically through linear, quadratic, or higher order polynomials that are allowed to differ on each side of the cutoff. The other approach, which we follow here, is to estimate the discontinuity nonparametrically, which we implement by local linear

regression with a rectangular kernel.[18] Our preferred estimates are based on a bandwidth of 2.5 points, in order to reduce bias by staying close to the cutoff while still maintaining enough precision. To assess the sensitivity of our results, we also present results from the optimal bandwidth determined by the procedure in Imbens and Kalyanaraman (2009), hereafter abbreviated IK. We also evaluate the sensitivity of our results to other bandwidths and higher order polynomials in Section 2.8.3.

A critical assumption to our identification strategy is that individuals just below a threshold are indeed comparable to individuals just above a threshold. One potential threat to this assumption is whether individuals are able to precisely sort around the threshold (Lee (2008)). If this assumption holds, then one implication is that the density of scores should be continuous around the threshold. Figure 2.3 displays the density of scores, in 0.1-point bins, in our sample around each threshold. With the exception of 75, we see no indication that the density is discontinuous around the threshold. Figure B.2a displays a smoother density of scores, in 1-point bins, which suggests a discontinuity in the density at 55. To address concerns of possible sorting, Figure B.2b displays the density of scores for those who were assessed in April of 2008, the first opportunity for eligibility evaluations and two months before the actual launch of the program. To the extent that the patterns in the 2009 density are due to sorting, we would not expect to see them in the April 2008 density, when individuals have no experience with how responses are mapped into scores. A comparison of Figures B.2a and B.2b indicates that the distribution of scores around the thresholds is strikingly similar for both periods.

Figure B.3 provides evidence for the complexity of the score function and the amount of variation inherent in the score. We take the set of individuals who responded "fully independent" for changing position and changed their response to "needs partial support." We recalculate their score and then plot this against their original score. Highlighting how highly interactive the score function is, note how the change in the response may lead to a change in the score ranging from a few points to more than ten points. This example indicates three things. First, it is difficult to pre-

[18] As noted in Lee and Lemieux (2010), the choice of kernel typically has little impact and while a triangular kernel is boundary optimal, a more transparent way of putting more weight on observations close to the cutoff is to reestimate a rectangular kernel based model using a smaller bandwidth.

cisely control the score. Second, there is a large degree of randomness within a few points. Third, it is possible that a response that indicates a sicker individual may actually lead to a *reduction* in points. This results from the highly interactive nature of the way the score is calculated.[19]

To the extent that there is no sorting and that the observed distribution of scores is due to the score function, individuals on each side of the threshold may still be comparable. As discussed in Urquiola and Verhoogen (2009), stacking alone may not violate the regression discontinuity assumptions since violation arises from the interaction of the stacking and the endogenous sorting of individuals. Thus, the more fundamental question for our identification strategy is whether the distribution of predetermined characteristics is identical on each side of the threshold. We show in Section 2.8.1 that with the exception of the 75 threshold, predetermined characteristics appear balanced around each threshold.

2.7 Results

We begin with our main results on the impact of eligibility on reimbursed utilization of formal long-term care, informal caregiving, and medical expenditures in Section 2.7.1. In Section 2.7.2, we address crowd out of private spending on formal-long term care and other potential explanations for our findings. In Section 2.7.3, we assess the cost-effectiveness of the LTCI program by comparing reimbursed long-term care expenses to medical expenditures.

2.7.1 Findings on Reimbursed Formal LTC, Informal Caregiving, and Medical Expenditures

Grade 3 (Home Care Only) Benefits

Figure 2.2a displays the probability of eligibility for Grade 3 benefits (i.e. home care only) as a function of the preliminary score, and Table 2.3a the estimated increases in probability at 50 and 55. Scoring just above 50 leads to an 8 percentage point increase in the probability of eligibility for home care benefits while scoring just above 55 leads to a 17 percentage point increase. To address

[19]We conducted this exercise for all questions and responses. This example is representative of our findings.

the impact of eligibility on utilization, Figure 2.7a displays reimbursed home care expenditures as a function of the preliminary score. Note that the pattern of expenditures corresponds well with the pattern of eligibility. In particular, as the score increases from 50 to 55, home care expenditures increase with the probability of eligibility for home care benefits. Moreover, there are discrete increases in expenditures at 50 and 55 corresponding to the discrete increases in the probability of eligibility for home care benefits at those points. Panel A of Table 2.4 contains estimates of the increases in reimbursed home care expenditures at 50 and 55. The increase in eligibility at 50 leads to a $300 increase in reimbursed home care expenditures while the increase in eligibility at 55 leads to a $850 increase. Regarding institutional care, Figure 2.7b displays reimbursed facility care days as a function of the preliminary score and Panel B of Table 2.4 contains estimates of the corresponding increases at 50 and 55. Consistent with no change in facility care benefits, the increases in eligibility for Grade 3 benefits at 50 and 55 do not lead to a statistically significant increase in facility care use.

We now assess the corresponding impacts of these changes in reimbursed formal care utilization on informal care. Figure 2.8 displays the one year changes in the probabilities of living independently (living alone or with one's spouse) and having a child caregiver as functions of the preliminary score. Figure 2.8a shows that the probability of living independently over time falls across all scores as individuals get sicker on average. Moreover, the decrease is larger for those who were not eligible for Grade 3 benefits relative to those who were. In particular, the pattern corresponds to the pattern of reimbursed home care utilization. Despite the overall patterns, however, the increased utilization of reimbursed home care at the thresholds does not translate to a statistically significant change in the probability of living independently as estimated in Panel D of Table 2.4. We find similar results for child caregiving. As seen in Figure 2.8b, the change in child caregiving is positive across all scores as individuals age and become sicker over time. However, it increases trivially among those eligible for Grade 3 benefits, suggesting that formal home care is able to avert the use of informal care. Moreover, the use of child caregiving increases among those who were not eligible for Grade 3 benefits. Again, however, despite the overall patterns, the increased utilization at the thresholds is not associated with a statistically significant change in

child caregiving as estimated in Panel C of Table 2.4.

There are several possible explanations for the limited impact on informal care. One potential explanation is that individuals who are ineligible for home care benefits may be able to finance these services privately, so that the probability of living independently (having a child caregiver) would fall (increase) less than in the absence of such an option. Another potential explanation is that formal home care allows a partial reduction, as opposed to complete elimination, of informal care. In other words, while there is no estimated impact on the extensive margin, there may still be an impact on the intensive margin. We address these potential explanations in Section 2.7.2.

Lastly, we assess the impact of increased home care utilization on medical expenditures and hospital utilization. Figure 2.9 displays the one year changes in these measures as functions of the preliminary score. We find no evidence that home care use impacts these outcomes, both across scores and treatment regimes as well as at the thresholds. The latter estimates are confirmed in Panels E and F of Table 2.4. The finding of no impact on medical expenditures is perhaps unsurprising given that the primary purpose of long-term care is not so much to restore or maintain health as it is to increase the general quality of life for the individual. We discuss these findings further in Section 2.7.3.

In summary, we find that eligibility for reimbursed home care benefits leads to the utilization of reimbursed formal home care. However, the use of reimbursed formal home care has no statistically significant impact on the use of informal care at the extensive margin nor on other medical utilization. There are various possible explanations for explaining the lack of an impact on informal care, which we address in Section 2.7.2.

Grade 2 (Home or Institutional Care) Benefits

We now assess the impact of Grade 2 benefits (i.e. where individuals can choose between home and institutional care benefits) on our outcomes of interest. Figure 2.2b displays the probability of eligibility for Grade 2 benefits as a function of the preliminary score, and Table 2.3b the estimated increases in probability at 70 and 75. Scoring just above 70 leads to a 4 percentage point increase in the probability of eligibility for home and institutional care benefits while scoring just above

75 leads to a 37 percentage point increase. To address the impact of eligibility on utilization, Figure 2.10 displays reimbursed home care expenditures and facility care days as a function of the preliminary score. We see that the pattern of reimbursed institutional care days corresponds well with the pattern of eligibility for those benefits. Consequently, reimbursed home care expenditures decrease as individuals substitute facility care for home care. Moreover, there are discrete increases (decreases) in facility (home) care use corresponding to the discrete increases in the probability of eligibility for institutional care at 70 and 75. Panels A and B of Table 2.4 contains estimates of the increases in reimbursed formal care expenditures at 70 and 75. The increase in eligibility at 70 leads to a 24 day increase in reimbursed facility use and a $400 decrease in home care expenditures. The increase in eligibility at 75 leads to a 23 day increase in reimbursed facility use and a $550 decrease in home care expenditures.

We next assess corresponding changes in informal care. Figure 2.11 displays the one year change in the probabilities of living independently and having a child caregiver as functions of the preliminary score. Again, we see that the change in the probability of living independently is negative across all scores as individuals get sicker over time, with the reduction slightly stronger for individuals eligible for facility benefits. However, there is no statistically significant change in independent living corresponding to the change in long term care utilization at 70 and 75 as estimated in Panels D of Table 2.4. For child caregiving, we see that it falls with the onset of facility care benefits, mimicking the pattern of eligibility for Grade 2 benefits. There is also suggestive evidence that the increased utilization of facility care benefits over home care benefits at 70 translates to a reduction in child caregiving, consistent with estimates in Panel C of Table 2.4. Estimates at our preferred bandwidth suggest that Grade 2 benefits lead to a statistically significant decrease in the probability of child caregiving of 3 percentage points. Estimates at more stringent bandwidths, including the IK, suggest similarly negative impacts, but these estimates are not precise enough to be statistically significant. Similarly for 75, estimates suggest negative, but not statistically significant, impacts on child caregiving.

There are several possible explanations for these findings. That there is no impact on independent living may not be a surprise. While facility care substitutes for home care, they both are

linked to dependent living situations. Although we do not find impacts of home care on the use of child caregiving, we do find suggestive impacts of facility care on the use of child caregiving. This is consistent with the fact that formal home care may reduce but not completely eliminate child caregiving. It is less likely that significant child caregiving would continue while the care recipient resides in a facility. We address these considerations more carefully in Section 2.7.2.

Lastly, we look at the impact of increased facility care and decreased home care utilization on medical expenditures and hospital utilization. Figure 2.12 displays the one year changes in these measures as functions of the preliminary score. We find no evidence that the substitution of facility care for home care at 70 impacts these outcomes. However, there is suggestive evidence at 75 that the substitution of facility care for home care leads to reductions in medical expenses and that this is largely accounted for by a reduction in hospital expenses. The estimates are shown in Panels E and F of Table 2.4. One explanation for this finding is that these individuals in this setting are less likely to experience costly accidents. Another explanation is that patients are able to transition sooner out of more expensive hospital care and into less expensive facility care. We discuss these findings further in Section 2.7.3.

In summary, we find that eligibility for facility care benefits leads to the substitution of facility care for home care. There is no impact on independent living, but there is suggestive evidence that this leads to a reduction in child caregiving at the extensive margin. There is also evidence for a corresponding reduction in medical utilization. As in our analysis of Grade 3 benefits, it will be important to take into account the ability of individuals to pay for formal long-term care services out of pocket, which we address in Section 2.7.2.

Grade 1 (Increased Maximum for Home Care, Increased Price for Institutional Care) Benefits

We now assess the impact of Grade 1 benefits on our outcomes of interest. Recall that these benefits are effectively an increase in the maximum benefit for home care combined with a discontinuous increase in the cost of facility care at the threshold. Figure 2.2c displays the probability of eligibility for Grade 1 benefits as a function of the preliminary score, and Table 2.3c the estimated increases

in probability at 90 and 95. Scoring just above 90 does not lead to a statistically significant increase in eligibility for Grade 1 benefits. Thus, assessments at this threshold serve as placebo tests for this design. As expected, we find no statistically significant impacts on reimbursed home expenditures and facility days, child caregiving and living independently, and medical and hospital expenses (see Figures 2.13 to 2.15 and the fifth row of Table 2.4).

A preliminary score just above 95 leads to an 83 percentage point increase in the probability of eligibility for Grade 1 benefits. To address the impact of eligibility on utilization, Figure 2.13 displays reimbursed home care expenditures and facility care days as functions of the preliminary score, and Panels A and B of Table 2.4 corresponding estimates of the discontinuities. Due to how Grade 1 benefits lead to a relative price increase in facility care, Grade 1 benefits at 95 lead to a 30 day decrease in the number of facility days used and a $930 increase in reimbursed home expenditures. As shown in Figure 2.14, with corresponding estimates in Panels C and D of Table 2.4, this shift in formal long-term care mix is not statistically significantly associated with changes in informal care, as measured by child caregiving and independent living. However, as shown in Figure 2.15 and Panels E and F of Table 2.4 we do find a statistically significant decrease in medical expenses of almost $700, coupled with a decrease in hospital expenditures of nearly the same amount. The fact that we find an impact of home care on medical expenditures in this case but not for Grade 3 may be due to the fact that individuals who receive Grade 1 benefits are more frail and susceptible to health shocks that can be ameliorated by formal care. We discuss our findings on medical expenditures further in Section 2.7.3.

In summary, we find that a relative increase in the price of facility care leads to increased utilization of formal home care. This shift in formal long-term care services has no impact on informal care but has a substantial impact on medical expenses, largely due to decreased hospital expenditures.

2.7.2 Crowd Out and Informal Care Intensity

The analysis of Grade 3 benefits in Section 2.7.1 indicated that an increase in reimbursed home care expenditures had little impact on informal care as measured by independent living and child

caregiving. One possible explanation for this finding is that public financing simply crowds out private expenditures for home care. Another possible explanation is that publicly financed home care enables individuals to reduce informal caregiving at the intensive margin. Unfortunately, our data does not provide measures of private spending on home care, nor does it contain measures of the amount of caregiving. Instead we focus on a subpopulation of individuals—those in the MCA program and thus are poor—for whom the likelihood of out-of-pocket spending is expected to be low.

The first column of Table 2.5 indicates estimates of the increase in home care utilization at 50 and 55 for the subset of MCA individuals. As in the overall sample, Grade 3 benefits lead to an increase in home care expenditures at 50 and 55 for MCA individuals. Columns two and three indicate estimates of the change in informal care at 50 and 55. As in the overall population, there is no statistically significant impact of Grade 3 benefits on informal care at the extensive margin for MCA individuals. Given that MCA individuals are unlikely to pay for home care out of pocket, these results suggest that the lack of an observable impact on informal care is not likely to be solely due to crowd out of private spending on formal care by public reimbursement.

A remaining explanation for why public reimbursement has no impact on informal care at the extensive margin is that the impact is on the intensive margin. To shed light on this possibility, we look at the impact of Grade 3 benefits on the use of a particular home care service, short-term respite care. Short-term respite care is short-term (i.e. a few days) facility care used to provide temporary relief for the regular caregiver. Thus, use of this type of home care is a strong indication for reduction in informal caregiving at the intensive margin. Indeed, as shown in Table 2.6, which shows estimates for several home care services, we find Grade 3 benefits lead to a statistically significant increase in the use of short-term respite care at 55.

As in the case for home care, we only observe publicly financed facility care. To measure the extent of crowd out we need a measure of all facility care, regardless of whether it is financed publicly or privately. To accomplish this we use an indirect measure of all facility utilization: medical spending occurring in a long-term care facility (i.e. regardless of financing). If the probability of having medical spending occurring in a long-term care facility is a fixed percentage of those

who attend a long-term care facility (at the threshold) then changes in the probability of having medical spending occurring in a long-term care facility will capture changes in the probability of attending a long-term care facility. In other words, if $\frac{\text{\# w/Medical Spending in LTC Facility}}{\text{\# in LTC Facility}}$ is fixed, then a percentage increase in the denominator will be tied to a percentage increase in the numerator of the same magnitude.[20] Table 2.7 presents estimates of the impact at 70 and 75 of the probability of using a publicly financed long-term care facility and the probability of having medical spending occurring in a long-term care facility. Scoring just above 70 is associated with a 25% increase (6.5 percentage points on a base of 25.7%) in the probability of using publicly financed facility care. However, using the probability of medical spending occurring in a long-term care facility as a proxy for all facility care shows that the probability of using facility care, regardless of financing, increases only 18.4% (2.9 percentage points on a base of 15.7%) at 70. This suggests that about a quarter of publicly financed care is used to substitute for out of pocket expenditures. The corresponding measure of crowd out at 75 is 46.7%. The fact that crowd out is higher at 75 than 70 is not surprising, given that individuals at 75 have more need for long-term care and thus are more likely to privately finance facility care in the absence of LTCI. While these measures of crowd out are substantial, they also suggest that crowd out is not complete, and therefore cannot fully explain our lack of findings for informal care.

2.7.3 LTC Expenditures and Reductions in Medical Expenses

In light of the previous results showing decreases in medical expenditure, a useful metric for assessing the cost-effectiveness of this policy and its costs to the government is to compare the reimbursed long-term care expenses to the changes in medical expenses. Recall that with the administrative data we use, we are able to measure the both the universe of medical expenditures and the universe of reimbursed long-term care expenditures. The first set of columns of Table 2.9 display the estimated impacts of all thresholds on reimbursed long-term care expenditures. For

[20]It is possible that those who spend out of pocket (i.e. those below the threshold) are likely to be sicker and thus have a higher probability of medical spending occurring in a facility. To the extent that this is the case, we will find a smaller change in the probability of having medical spending occurring in a facility and an over (upper bound) estimate of crowdout.

convenience, the second set of columns redisplay the impacts on medical expenditures. The third set of columns indicate the medical expenditures saved per additional dollar of long-term care expenditure reimbursed.

A preliminary score above 50 and 55 leads to a $208 and $931 increase in total reimbursed long-term care expenditures, respectively. As seen earlier, however, this results in little, if any, savings in medical expenditures. Focusing on Grade 2, we see that additional benefits for facility care lead to an additional $500 in expenditures as individuals substitute more expensive facility care in place of home care. However, corresponding to this increase in expenditure we find a decrease in medical expenditures of more than $300, for a medical expenditure savings of $0.6 per dollar of long-term care reimbursed. The fact that there is no apparent savings at 70 may be due to heterogeneous impacts of the policy or possible bias at 75. Focusing on Grade 1 at 95 (recall that there is little effective change in eligibility at 90), we see that additional benefits for Grade 1 lead to only small changes in expenditures as individuals tend to use more home care and less facility care. However, this substitution leads to large impacts on medical expenditures—nearly a $700 reduction. Thus, Grade 1 benefits lead to a medical expenditures savings of more than $650 per dollar of long-term care reimbursed. Clearly, the amount of long-term care reimbursed is not a complete measure of the costs of the program as it does not include the administrative expenses, for example. Moreover, medical expenses are not a complete measure of the potential cost savings of the program as impacts on labor outcomes could have impacts on government revenue.[21] However, the large impact we measure here highlights the importance of considering the potential program savings from reduced medical expenditures.

[21] Our limited findings on informal care at the extensive margin suggest that these labor market impacts are likely small.

2.8 Robustness

2.8.1 Balance of Covariates

As discussed in Section 2.3.2, an important assumption for our identification strategy is that individuals on each side of the thresholds are comparable. A test of this assumption is to check the balance of observable characteristics across the thresholds. Table 2.10 contains estimates of the discontinuities around the thresholds for predetermined variables that are likely to be correlated with our dependent variables of interest. With the exception of the 75 threshold, most of the variables appear to be continuous around the thresholds at our preferred bandwidth.

Because we are testing numerous variables and thresholds, some discontinuities will be statistically significant by random chance. As a result, we conduct two tests which account for this, with results presented in the last columns of Table 2.10. First, we look at a summary measure—the predicted medical expenditures from a regression of medical expenditures on the other predetermined variables. Again, with the exception of the 75 threshold, there appear to be no discontinuities in predicted medical expenditures at our preferred bandwidth. Second, we test whether the discontinuities are jointly significant by seemingly unrelated regression, as described in Lee and Lemieux (2010). Consistent with the first exercise, the only threshold for which the discontinuities are jointly significant at the preferred bandwidth is 75. This leads us to believe that our results are not impacted by unobserved confounders at the other thresholds. Nonetheless, we controlled for the few instances of significance occurring in our variables of interest by estimating differences in our dependent variables in our regressions.

2.8.2 Differential Mortality

Another relevant outcome is whether these benefits had any impact on mortality. This measure is important in and of itself, and is useful because it is objective and well-defined. Moreover, it is important to address the concern that differential mortality around the thresholds could account for our findings. For example, if individuals just below the threshold were more likely to die as a result of not receiving treatment, relatively healthy individuals would remain in the sample,

minimizing any estimated impacts. We assess this by looking at mortality by 2010 around the thresholds. Table 2.8 displays estimates of Equation 2.1 with mortality by 2010 as the outcome. We find no statistically significant differences in mortality at all thresholds. Thus, the increase in long-term care utilization at the thresholds has no impact on mortality in the short-run.

2.8.3 Other Specifications

A consequential decision in estimating Equation 2.1 is the choice of bandwidth. Although we have shown that our results are qualitatively consistent at both our preferred bandwidth and the IK bandwidth, it is useful to know how sensitive our findings are to bandwidth choice. To do so, we reestimate Equation 2.1 for our main outcomes of interest at several bandwidths—from 1 to 5, in increments of 0.5. Figures **??** to **??** plot the estimated coefficients with 95% confidence bands against the bandwidth. There are two things worth highlighting. First, coefficients are less precisely estimated and more variable at very small bandwidths. Second, the coefficient estimate at our preferred bandwidth falls within the 95% confidence bands of the estimates at other bandwidths in general, indicating that our findings are not too sensitive to bandwidth selection.

On the specification of $f(S)$, our approach in this paper follows Hahn, Todd, and van der Klaauw (2001) by using local linear regressions to estimate the discontinuity at the threshold. As shown in the previous section, our findings are consistent even at very small bandwidths. Moreover, visual inspection suggests the relationship between eligibility (as well as our outcomes of interest) and the preliminary score is fairly linear even at relatively large distances from the thresholds. Nonetheless, in Figures **??** to **??** we explore how sensitive our findings are to higher order specifications of $f(S)$ at our preferred bandwidth. For the most part, the coefficient estimate based on a linear specification of $f(S)$ falls within the 95% confidence bands of estimates for higher order specifications. However, the variance of the higher order specifications grows quite large, which lends support for the use of linear splines.

2.8.4 Differences-in-Differences Estimation

Our research design takes advantage of a setting with a continuous measure of long-term care needs (i.e. the preliminary score) and thresholds that lead to "as good as random" variation in the probabilities of benefits. One limitation of this design, however, is the reduced precision from relying primarily on observations around the threshold. In this section, we estimate a differences-in-differences model that relies on stronger assumptions, but has potentially improved precision. Specifically, we compare three groups of individuals: individuals who are treated based solely on the preliminary score (for Grade 3, we consider individuals with preliminary scores in [55,60)), individuals who are treated based on committee guidelines (for Grade 3, these are individuals with preliminary scores in [50,55)), and individuals who are not treated (for Grade 3, these are individuals with preliminary scores in [45,50)). For $\tau \in \{55, 75, 95\}$, we define $\text{commit}_\tau \equiv \mathbb{1}\{\tau - 5 \leq S < \tau\}$ and $\text{treat}_\tau \equiv \mathbb{1}\{\tau \leq S < \tau + 5\}$, where S is the 2009 preliminary score. With the untreated individuals (i.e. $\{S : \tau - 10 \leq S < \tau - 5\}$) as our reference group, we estimate the following saturated differences-in-differences model for an individual i in a one point bin b at time t:

$$\text{outcome}_{ibt} = \sum_{t \neq 0} (\beta_t^C \text{commit}_\tau \cdot \phi_t + \beta_t^T \text{treat}_\tau \cdot \phi_t) + \gamma_i + \phi_t + \eta_b \cdot t + \epsilon_{ibt}, \tag{2.2}$$

where γ_i is an individual-specific effect, ϕ_t a year-fixed effect, $\eta_b \cdot t$ is a bin-specific linear time trend, and the baseline year is set to $t = 0$.[22] The key assumption underlying this estimation method is that there are no unobserved factors that affect the three groups differentially within a year.

Table 2.11 presents estimates of β_1^C and β_1^T from Equation 2.2. Grade 3 expenditures lead to a statistically significant decrease in child caregiving, but have no statistically significant impact on independent living. There is no statistically significant impact on medical expenditures or hospital expenses. Additional long-term care expenditures resulting from Grade 2 benefits are also associated with a statistically significant decrease in child caregiving, but not independent living. The use of Grade 2 benefits leads to a decrease in other medical expenses, accounted for

[22]For our outcome measures (medical expenditures and informal care), $t \in \{-1, 0, 1\}$. For our long-term care utilization measures, $t \in \{0, 1, 2\}$. Table 2.11 presents estimates of β_1^C and β_1^T from Equation 2.2.

largely by hospital expenses. These results translate into a medical dollars saved per additional dollar of reimbursed long-term care expenditure of 0.2–0.3. Grade 1 benefits are associated with a statistically significant increase in child caregiving, but not independent living. This is consistent with the increased use of home care among these individuals that was found earlier. The use of Grade 1 benefits leads to a decrease in other medical expenses, largely accounted for by hospital expenses. In this case, the medical dollars saved per additional dollar of reimbursed long-term care expenditure is more than one, suggesting strong program savings.

The findings from this analysis are fairly consistent with our findings from the regression discontinuity analysis. Even though the differences-in-differences analysis suggests statistically significant impacts on child caregiving while RD estimates do not, this could be due to lack of statistical precision. Moreover, estimates of medical expenditures saved per dollar of reimbursed long-term care are similar across both estimation strategies.

Lastly, this estimation strategy allows us to compare the committee affected group to the automatically treated group. This is particularly relevant given that assigning treatment based solely on the preliminary score may not be optimal and that leaving room for discretionary assignment of treatment may improve efficiency. In this analysis, there do not appear to be any striking differences in performance between the two groups among Grades 3 and 2 individuals. However, it appears that the committee affected group has a more substantial impact among Grade 1 affected individuals. While this suggests the possibility that a more discretionary decision-making procedure for determining treatment may be more effective than a hard rules-based criteria, we caution that this measure (vs. quality of life, for example) may not be the primary objective to optimize from the standpoint of the committee.

2.9 Discussion

In this paper, we find that publicly financed LTCI leads to small, if any, impacts on informal care at the extensive margin. We determine that this is not solely due to crowdout, but partly explained by the fact that informal care is reduced at the intensive margin. That we find limited impacts on informal care stands in contrast to some of the previous literature, but is not surprising given

that South Korea is a strong family ties country. That is, due to family obligations, Koreans may find it more difficult to give up completely the responsibility of taking care of their elderly parents. That we still find reductions in the intensive margin indicate that our results constitute a lower bound for the effect in the U.S. in general, and may be directly indicative of population subgroups in the U.S. such as Asians and Hispanics.

Interestingly, we find that among people who are partially dependent for several activities of daily living, transitioning from home to facility care results in decreased medical expenditures. This may come as a surprise at first, given that the purpose of long-term care is not so much to restore or maintain health as it is to increase the general welfare of the individual by facilitating activities of daily living. Indeed, we find no impacts on health as captured by mortality. However, a plausible explanation is that the increased attention one receives in a facility may prevent costly accidents like falling and breaking one's hip. Another possibility is that patients are able to transition sooner out of more expensive hospital care and into less expensive facility care. Surprisingly, among individuals who are completely dependent for several activities of daily living, the opposite transition leads to substantially lower medical expenses. This may be mediated by the fact that the presence of medical professionals in a facility may lead to additional or more costly care than if one were being cared for at the home, and that, among this population of individuals, this effect predominates the previously mentioned effects. In fact, that transitioning people from institutions to the community may be beneficial is consistent with the objectives of programs such as Money Follows the Person in the U.S. This supports the more general point that our findings on medical expenses are not culturally or context specific, and that understanding the relationship between long-term care expenses and medical expenses may be a fruitful avenue to contain health care costs.

2.10 Conclusion

Results from this paper provide insight into the welfare impacts of government reimbursement of long-term care on care recipients, caregivers, and taxpayers, as well as suggestions for the design of optimal long-term care policy. Our main finding is that the benefits of home and facility care

are heterogeneous across physical function level and therefore setting policy accordingly has the potential to dramatically reduce medical expenses. We also find that formal long-term care is not a strong substitute for informal long-term care at the extensive margin.

Among individuals who are partially dependent for some activities of daily living we find that government subsidies for formal home care lead to an overall increase in its utilization, even accounting for crowd out, with no impact on informal caregiving at the extensive margin, medical expenses, or mortality. While we find evidence for a reduction in informal caregiving at the intensive margin, this suggests that if the policy objective is to increase the labor supply of individuals caring for this population, subsidies for home care may have limited impact. Moreover, the converse of our findings on medical expenses and mortality suggest that home care reimbursement may be reduced without significant detriment to the health of the care recipient.

Among individuals who are partially dependent for several activities of daily living, additional reimbursement of institutional care leads to the crowd out of privately financed institutional care of up to 47%. Institutional care does increase overall, leading to reductions in informal caregiving and medical expenses. From a policy perspective, the latter finding suggests that while substitution of institutional care for less expensive home care may lead to increased costs, this may be partially offset by reductions in medical expenses. Moreover, our finding on informal caregiving suggests that this policy may lead to increased labor supply of individuals caring for this population. In this case, optimal policy depends on the objective function of the policymaker in balancing the tradeoff between increased taxpayer costs, reduced informal caregiving, and improved quality of life for the care recipient.

Among individuals who are completely dependent for several activities of daily living, we find that an increase in the price of institutional care combined with an increase in the benefit maximum for home care leads to substitution of home care for institutional care. While we find no impact on informal caregiving, we find substantial decreases in medical spending. From a policy perspective, this suggests that increased incentives for the use of home care may lead to an improvement in the welfare of care recipients while limiting or even reducing costs to taxpayers.

Table 2.1: Overview of Grades of Benefits

Classification	Description	Criteria	Home Care Max Benefit (USD)	Institutional Care Max Benefit (USD)
No Benefits		score < 55	none	none
Grade 3	Need assistance moving around, partially dependent for some ADLs	55 ≤ score < 75	750 / month	none
Grade 2	Unable to move on own, partially dependent for several ADLs	75 ≤ score < 95	900 / month	40 / day
Grade 1	Bedridden, completely dependent for several ADLs	95 ≤ score	1100 / month	45 / day

1 USD ≈ 1100 KRW

Table 2.2: Summary Statistics by Grade

Adjusted Score	No Benefits [45,55)	Grade 3 [55,75)	Grade 2 [75,95)	Grade 1 [95+]
# Obs	*35,580*	*43,615*	*76,170*	*12,090*
Age	76.06	76.87	78.13	77.01
	(8.16)	(8.85)	(8.74)	(9.68)
Female	0.77	0.73	0.74	0.73
	(0.42)	(0.44)	(0.44)	(0.44)
Urban	0.73	0.76	0.78	0.77
	(0.45)	(0.43)	(0.41)	(0.42)
Insurance Contribution	41.27	54.32	62.49	64.19
	(63.92)	(73.04)	(74.51)	(79.54)
MCA	0.43	0.31	0.23	0.25
	(0.50)	(0.46)	(0.42)	(0.43)
ADL Index	17.42	20.07	24.96	30.05
	(3.78)	(4.48)	(5.63)	(5.99)
Medical Expenditures	2,255	2,850	4,165	5,080
	(4,312)	(5,190)	(6,719)	(8,060)
Hospital Days	12.53	19.56	42.48	55.37
	(42.53)	(57.86)	(89.82)	(105.59)
Child Caregiver	0.26	0.30	0.23	0.19
	(0.44)	(0.46)	(0.42)	(0.39)
Live Independently	0.60	0.42	0.21	0.21
	(0.49)	(0.49)	(0.41)	(0.40)
LTC Facility Days	21.90	66.82	159.47	158.58
	(63.69)	(137.56)	(167.55)	(168.94)
Home Care Exp	2,885	5,061	3,442	3,384
	(3,037)	(3,836)	(4,263)	(4,682)

Notes: Sample consists of individuals who were assessed for long-term care insurance in 2008 and 2009. Grade categorization is based on the 2009 adjusted score. All measures are at baseline, except for long-term care facility days and home care expenditures. See text for definitions of variables.

Table 2.3: Effect of Thresholds on Changes in Eligibility

(a)

Grade 3 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score ≥ 50	0.08**	0.09**	1.3
	(0.01)	(0.02)	
Score ≥ 55	0.17**	0.10**	0.6
	(0.01)	(0.03)	

(b)

Grade 2 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score ≥ 70	0.04**	0.06**	0.8
	(0.01)	(0.01)	
Score ≥ 75	0.37**	0.39**	1.1
	(0.01)	(0.03)	

(c)

Grade 1 Eligibility			
Bandwidth	**2.5**	**IK**	**IK B/W**
Score ≥ 90	0.01+	0.01*	0.9
	(0.00)	(0.00)	
Score ≥ 95	0.83**	0.83**	1.1
	(0.01)	(0.02)	

Notes: The first two columns of each panel report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.4: Main Results on LTC Utilization, Informal Care, and Medical Expenditures

	Reimbursed Formal LTC Utilization						Informal Care						Medical Utilization					
	Panel A			Panel B			Panel C			Panel D			Panel E			Panel F		
	LTC Home Expenditures			LTC Facility Days			Pr(Child Caregiver)			Pr(Live Independently)			Medical Expenses			Hospital Expenses		
Bandwidth	2.5	IK	IK B/W	2.5	IK	IK B/W	2.5	IK	IK B/W	2.5	IK	IK B/W	2.5	IK	IK B/W	2.5	IK	IK B/W
Score ≥ 50	311*	480**	2.0	-2.3	-4.5	1.0	-0.03	0.13*	1.4	-0.02	-0.14**	1.1	-97	-198	1.8	-177	-241	1.9
	(157)	(181)		(3.7)	(6.5)		(0.04)	(0.06)		(0.04)	(0.05)		(174)	(217)		(174)	(213)	
Score ≥ 55	850**	748**	1.7	0.2	24.2**	1.0	0.01	0.00	1.1	-0.02	-0.02	1.3	59	-67	1.9	60	2	3.1
	(134)	(163)		(4.0)	(6.6)		(0.02)	(0.03)		(0.02)	(0.02)		(146)	(171)		(141)	(127)	
Score ≥ 70	-392**	-561**	1.3	23.8**	37.8**	1.2	-0.03*	-0.02	1.2	0.00	0.00	1.2	101	86	2.6	145	93	2.7
	(145)	(208)		(5.3)	(8.2)		(0.01)	(0.02)		(0.02)	(0.02)		(173)	(170)		(176)	(170)	
Score ≥ 75	-554**	-571**	2.1	22.5**	24.7**	1.0	-0.01	-0.03	1.2	-0.02	0.00	1.1	-327+	-405*	2.1	-370*	-125	1.5
	(141)	(155)		(5.4)	(9.1)		(0.01)	(0.02)		(0.02)	(0.02)		(178)	(200)		(183)	(242)	
Score ≥ 90	25	-21	2.2	3.7	2.6	2.5	0.02	-0.01	1.3	0.02	0.02	1.4	-344	-249	2.7	-245	-149	2.9
	(219)	(231)		(8.5)	(8.4)		(0.02)	(0.03)		(0.02)	(0.02)		(295)	(283)		(302)	(289)	
Score ≥ 95	926**	586+	1.6	-29.5**	-27.9*	1.5	0.02	0.04	1.5	0.00	0.00	1.5	-691*	-757*	2.3	-666+	-727*	2.3
	(242)	(320)		(9.0)	(11.9)		(0.02)	(0.03)		(0.03)	(0.03)		(319)	(333)		(342)	(356)	

Notes: The first two columns of each panel in this table report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.5: Utilization and Informal Care for MCA Individuals

	LTC Home Expenditures			Pr(Child Caregiver)			Pr(Live Independently)		
Bandwidth	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**
Score ≥ 50	476+	577	1.5	0.06	0.05	1.9	-0.06	-0.06	1.9
	(247)	(355)		(0.06)	(0.06)		(0.06)	(0.06)	
Score ≥ 55	930**	-358	1.0	0.03	0.04	1.6	-0.04	-0.04	1.6
	(232)	(371)		(0.02)	(0.03)		(0.03)	(0.03)	

Notes: The first two columns of each panel in this table report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Sample consists of individuals in the MCA program. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.6: Detailed Home Care Utilization

	Home Help		Home Bath		Home Nursing		Day / Evening Care		Respite Care		Equipment	
Bandwidth	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**
Score ≥ 50	11.42*	11.44	0.04	0.15	0.06	-0.06	-2.43	-0.56	-0.74	-1.76	2.85	1.93
	(4.76)	(7.44)	(0.45)	(0.45)	(0.20)	(0.25)	(1.81)	(1.67)	(1.31)	(1.45)	(2.07)	(1.97)
Score ≥ 55	16.02**	1.28	0.50	0.20	-0.33	-0.39	2.39	3.49	6.67**	6.76**	1.25	0.05
	(4.21)	(5.73)	(0.42)	(0.55)	(0.25)	(0.24)	(2.11)	(2.49)	(1.38)	(1.38)	(2.22)	(2.53)

Notes: Each cell reports estimates of β from a different local linear regression of Equation (2.1). Dependent variables are measured in # of visits. The running variable is the 2009 preliminary score. Rectangular kernel. Optimal bandwidths for the IK procedure are omitted for space considerations and are available from the authors upon request. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.7: Crowd Out of Facility Care

Bandwidth	Pr(Publicly Financed Facility Care) 2.5	Pr(Med Spending in LTC Facility) 2.5	Crowd Out (%)
Change at 70	0.065**	0.029*	
	(0.016)	(0.0135)	
Base at 70	0.257**	0.156**	
	(0.011)	(0.009)	
% Change at 70	25.4%	18.4%	**27.4%**
Change at 75	0.068**	0.016	
	(0.017)	(0.014)	
Base at 75	0.483**	0.214**	
	(0.015	(0.012)	
% Change at 75	14.2%	7.5%	**46.7%**

Notes: Columns 1 and 2 report coefficient estimates from Equation (2.1). Dependent variables are indicators for public reimbursement of facility care and medical spending in a LTC facility. The running variable is the 2009 preliminary score. Rectangular kernel. "Change at 'X' " is the estimate of β. "Base at 'X' " is the predicted value of the dependent variable at 'X' minus the "Change at 'X' ". ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.8: Effect of Eligibility on Mortality

	Mortality by 2010		
Bandwidth	**2.5**	**IK**	**IK B/W**
Score ≥ 50	0.00	-0.02	1.1
	(0.01)	(0.02)	
Score ≥ 55	0.00	0.00	1.2
	(0.01)	(0.01)	
Score ≥ 70	0.00	0.00	1.2
	(0.01)	(0.02)	
Score ≥ 75	0.00	0.02	0.9
	(0.01)	(0.02)	
Score ≥ 90	0.00	0.03	1.5
	(0.02)	(0.03)	
Score ≥ 95	-0.02	-0.04	1.5
	(0.02)	(0.03)	

Notes: The first two columns of this table report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. The dependent variable is mortality by 2010. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third column of each panel reports the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 2.9: LTC Expenses vs. Medical Care Savings

	LTC Expenditures			Medical Expenses			$ Med Exp Saved / $ LTC Spent	
Bandwidth	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**	**IK B/W**	**2.5**	**IK**
Score ≥ 50	208	32	1.8	-97	-198	1.8	0.5	6.1
	(169)	(209)		(174)	(217)			
Score ≥ 55	931**	1,090**	1.8	59	-67	1.9	-0.1	0.1
	(140)	(172)		(146)	(171)			
Score ≥ 70	524**	796**	1.9	101	86	2.6	-0.2	-0.1
	(156)	(183)		(173)	(170)			
Score ≥ 75	535**	711**	1.1	-327+	-405*	2.1	0.6	0.6
	(164)	(267)		(178)	(200)			
Score ≥ 90	155	10	2.4	-344	-249	2.7	2.2	23.9
	(259)	(268)		(295)	(283)			
Score ≥ 95	1	-379	1.2	-691*	-757*	2.3	668.3	-2.0
	(281)	(432)		(319)	(333)			

Notes: The first two columns of the first two panels in this table report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. The running variable is the 2009 preliminary score. Controls include age, gender, region type, insurance type, and insurance contribution. Rectangular kernel. The third columns report the optimal bandwidth determined by the IK procedure. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. The third panel equals the coefficient in the second panel divided by the coefficient in the first panel.

Table 2.10: Covariate Balance

	Med Exp		Med Exp in LTC		MCA Insurance		Insurance Contrib.		Child Caregiver		Live Independently		SUR	Predicted Med Exp	
Bandwidth	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**2.5**	**IK**	**p-value**	**2.5**	**IK**
Score ≥ 50	311	216	17	18	-0.03	-0.02	3.16	3.25	-0.04	-0.13*	0.05	0.22**	0.49	141	83
	(202)	(189)	(81)	(86)	(0.02)	(0.03)	(3.50)	(3.46)	(0.04)	(0.06)	(0.05)	(0.07)		(127)	(135)
Score ≥ 55	42	115	43	195**	-0.01	0.02	-1.17	-3.78	-0.03	-0.03	0.07**	-0.01	0.19	115	125
	(167)	(148)	(58)	(70)	(0.02)	(0.02)	(3.21)	(3.65)	(0.02)	(0.03)	(0.03)	(0.04)		(95)	(76)
Score ≥ 70	-28	-243	120	261+	-0.02	-0.05*	-3.94	-4.45+	0.00	0.01	-0.02	-0.02	0.34	183	340*
	(206)	(229)	(111)	(135)	(0.02)	(0.02)	(2.77)	(2.51)	(0.02)	(0.03)	(0.02)	(0.03)		(130)	(158)
Score ≥ 75	1,095**	591*	753**	274	-0.03+	-0.03	3.68	3.55	0.01	0.04	-0.01	-0.01	0.00	826**	509**
	(217)	(255)	(125)	(175)	(0.01)	(0.03)	(2.45)	(2.63)	(0.02)	(0.03)	(0.02)	(0.03)		(141)	(167)
Score ≥ 90	258	338	174	180	0.03	0.02	1.29	2.73	-0.02	-0.03	-0.07*	-0.11**	0.35	137	191
	(333)	(386)	(220)	(222)	(0.02)	(0.03)	(4.42)	(4.98)	(0.03)	(0.04)	(0.03)	(0.04)		(239)	(254)
Score ≥ 95	801*	410	297	303	-0.01	-0.03	6.65+	9.33+	-0.03	0.11*	0.03	0.05	0.23	465+	454+
	(396)	(330)	(249)	(285)	(0.02)	(0.04)	(3.91)	(4.88)	(0.03)	(0.05)	(0.03)	(0.04)		(274)	(275)

Notes: Columns 1-6 and 8 report estimates of β from local linear regression of Equation (2.1). Each cell represents a different regression. Dependent variables are 2008 measures. The running variable is the 2009 preliminary score. Rectangular kernel. Optimal bandwidths for the IK procedure are omitted for space considerations and are available from the authors upon request. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Column 7 reports the p-value from a joint test of the coefficients in each row from a SUR where the bandwidth is 2.5.

Table 2.11: Differences-in-Differences Estimates

Grade 3	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	982.1**	41.40	105.5	-0.0227	-0.0111	0.0
	(62.85)	(92.88)	(88.27)	(0.0162)	(0.0168)	
Treatment	3,141**	-11.50	22.99	-0.0604**	0.0182	0.0
	(51.97)	(79.67)	(74.13)	(0.0133)	(0.0139)	
Individuals	33,005	33,005	33,005	32,997	32,061	
R-squared	0.571	0.006	0.004	0.003	0.021	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [50,55), Treatment = Preliminary Score in [55,60), Omitted = Preliminary Score in [45,50). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

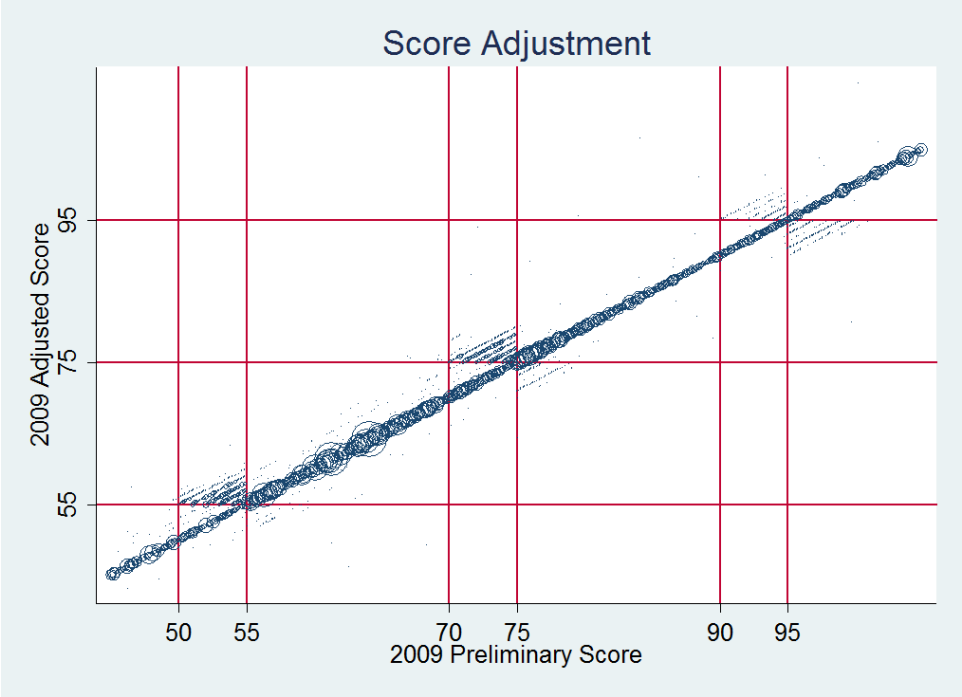
Grade 2	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	391.4**	-94.00	-49.03	-0.00143	-0.00664	0.2
	(49.91)	(86.28)	(85.78)	(0.00735)	(0.00810)	
Treatment	1,400**	-432.8**	-419.9**	-0.0321**	-0.00845	0.3
	(45.24)	(79.39)	(79.33)	(0.00655)	(0.00708)	
Individuals	49,930	49,930	49,930	49,923	48,194	
R-squared	0.623	0.012	0.010	0.029	0.057	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [70,75), Treatment = Preliminary Score in [75,80), Omitted = Preliminary Score in [65,70). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

Grade 1	LTC Exp	Med Exp	Hosp Exp	Child Caregiver	Live Independently	$ Med Exp Saved / $ LTC Spent
Committee	-93.11	-326.6*	-307.6*	0.0141	-0.00297	-3.5
	(93.09)	(153.7)	(156.7)	(0.0123)	(0.0125)	
Treatment	303.4**	-589.8**	-546.4**	0.0348**	0.00128	2.0
	(95.48)	(159.0)	(161.6)	(0.0121)	(0.0130)	
Individuals	17,490	17,490	17,490	17,481	16,742	
R-squared	0.631	0.013	0.012	0.049	0.064	

Post-period regression coefficients from differences-in-differences estimation with individual fixed effects, year fixed effects, and bin-specific linear time trends. Committee = Preliminary Score in [90,95), Treatment = Preliminary Score in [95,100), Omitted = Preliminary Score in [85,90). Robust standard errors in parentheses.
** p<0.01, * p<0.05, + p<0.1

Figure 2.1: Adjusted Scores vs. Preliminary Scores, 2009

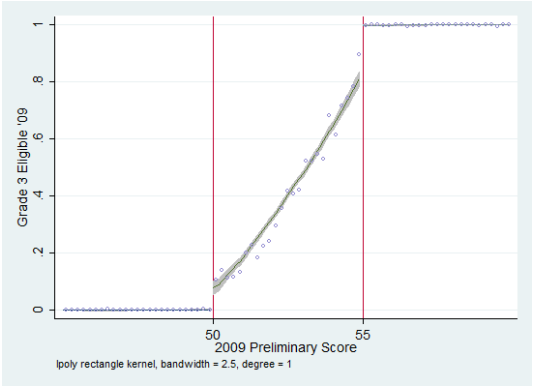


Notes: This figure plots the 2009 adjusted score against the 2009 preliminary score, for individuals whose preliminary scores fall between 45 and 105. Circle sizes correspond to the number of individuals with the associated adjusted/preliminary score combination.

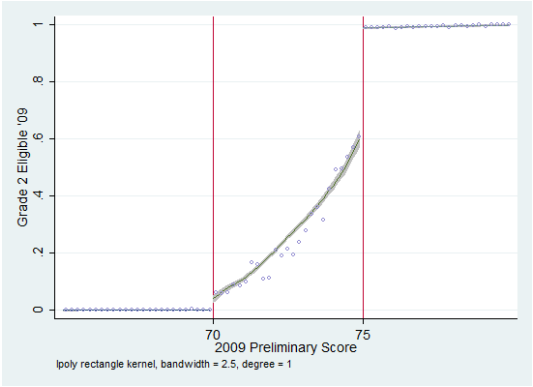
2009 adjusted score = 2009 preliminary score + committee points, where committee points $\in [-5, 5]$.

Figure 2.2: Probability of Eligibility vs. 2009 Preliminary Score

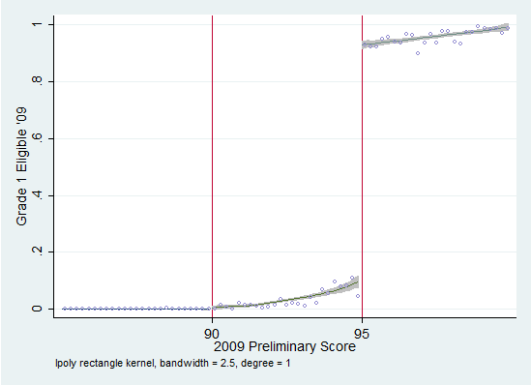
(a) Grade 3 Eligibility



(b) Grade 2 Eligibility



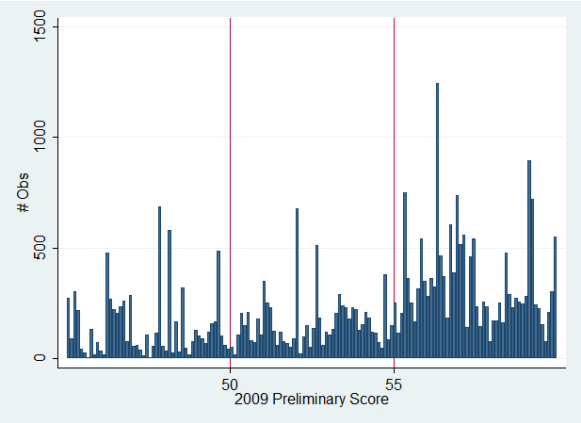
(c) Grade 1 Eligibility



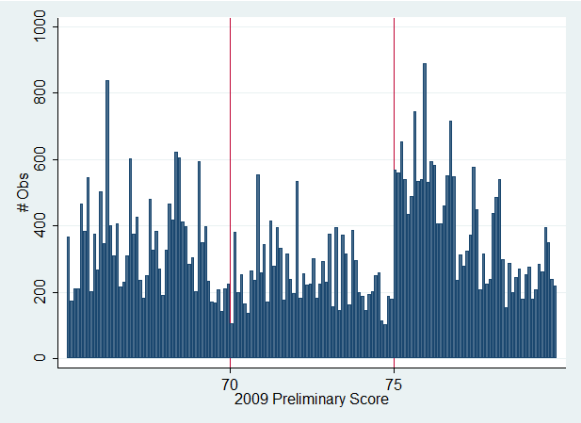
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.3: Histograms of Scores

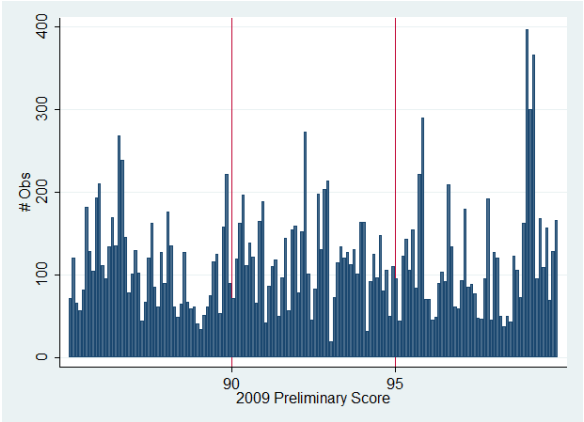
(a) Around Grade 3 Thresholds



(b) Around Grade 2 Thresholds



(c) Around Grade 1 Thresholds



Notes: 2009 preliminary score in 0.1 point bins.

Figure 2.4: Impact of $\uparrow s_H$ on A, C, and H

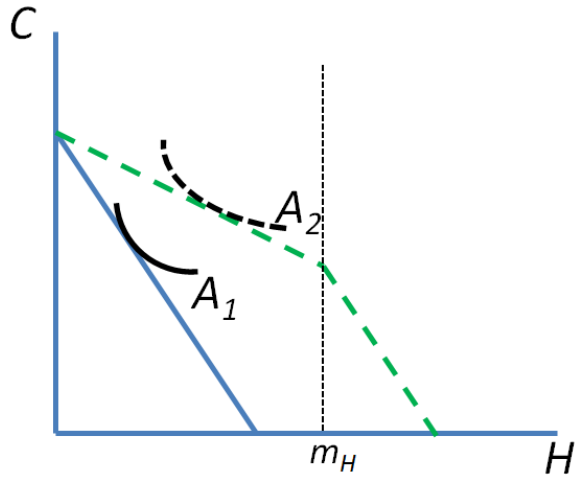


Figure 2.5: Impact of $\uparrow s_F$ on A, H, and F

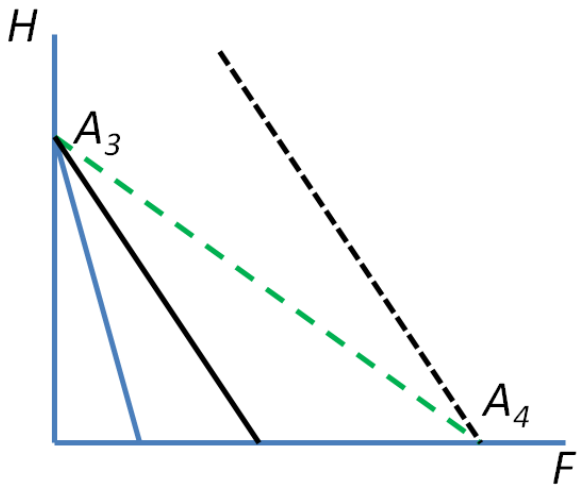
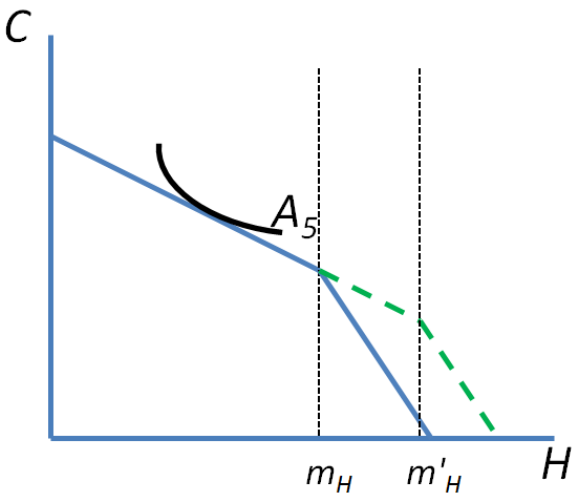


Figure 2.6: Impact of $\uparrow m_H$ on A, C, and H

(a) $H_1 < m_H$



(b) $H_1 = m_H$

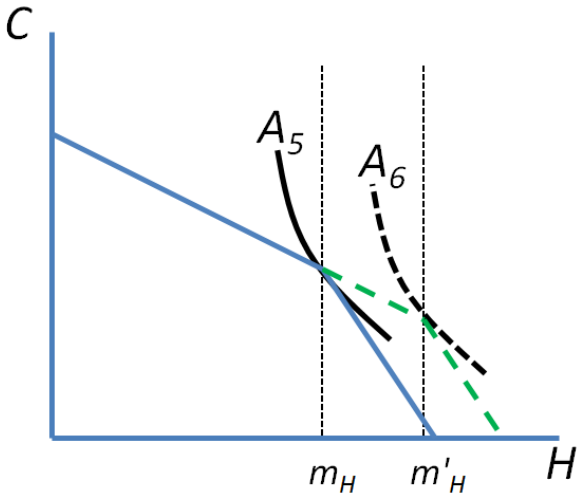
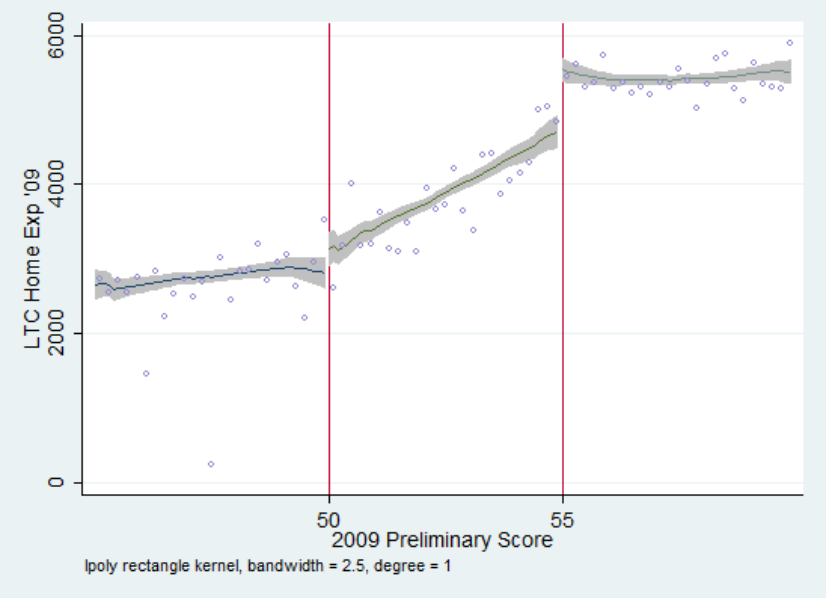
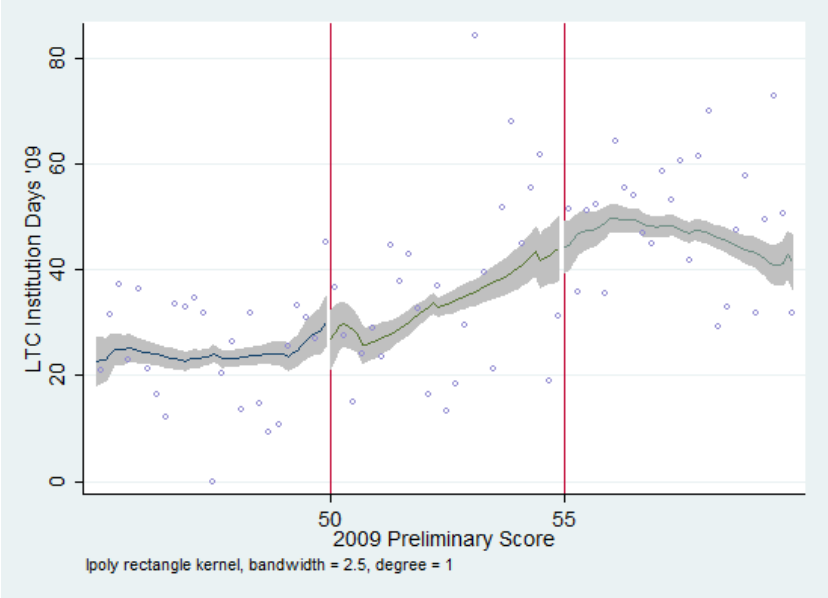


Figure 2.7: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 3

(a) Home Care Expenditures



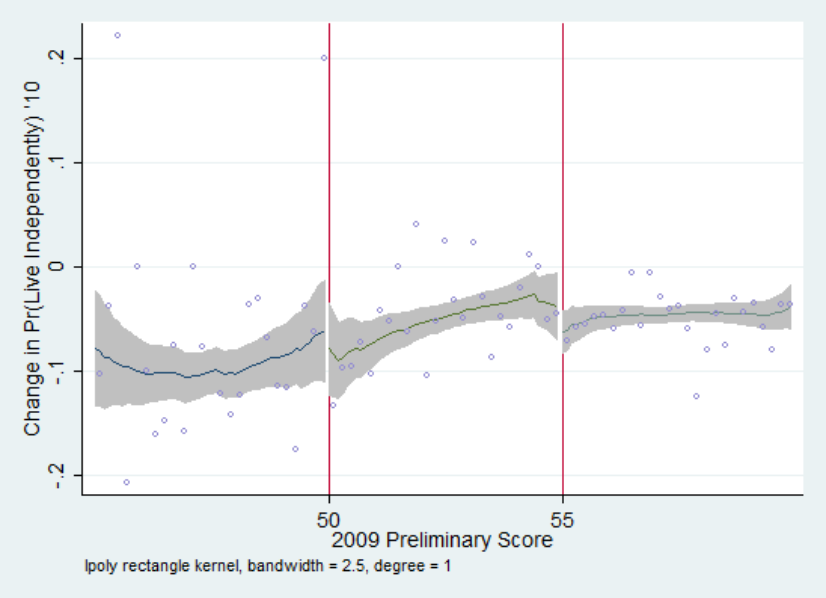
(b) Institutional Days



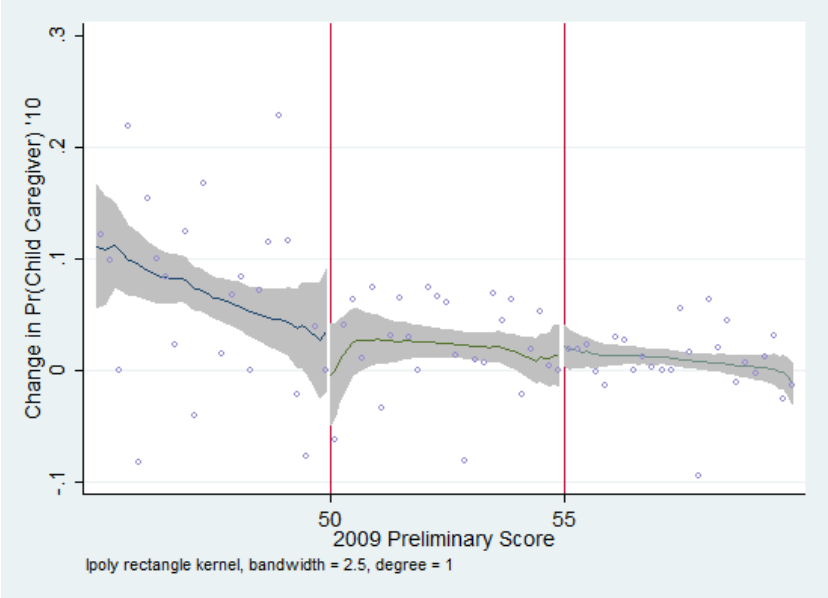
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.8: Change in Informal Care vs. Preliminary Score Around Grade 3

(a) Δ Pr(Live Independently)



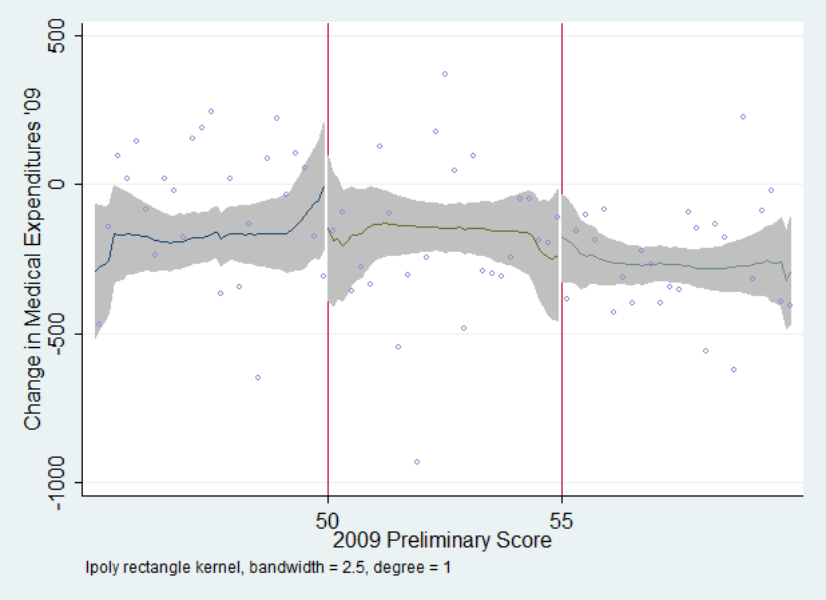
(b) Δ Pr(Child Caregiver)



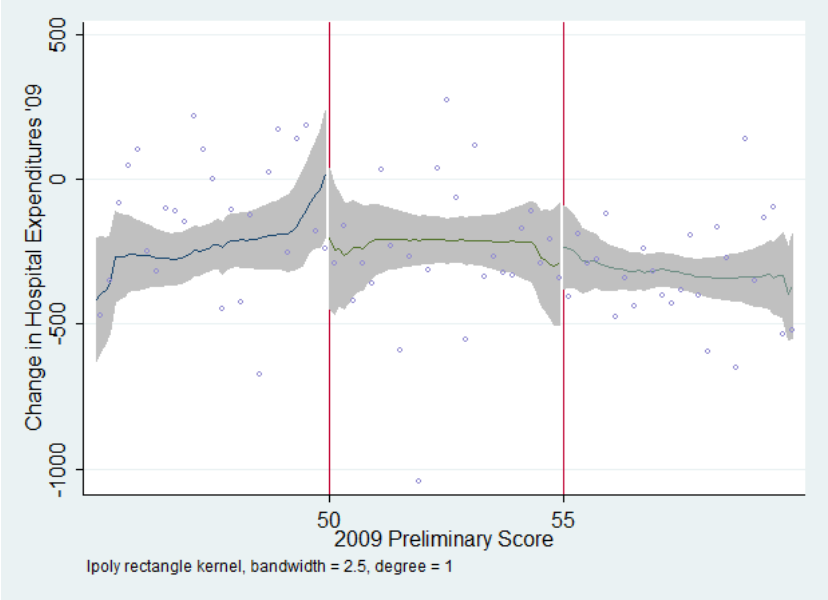
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.9: Change in Medical Utilization vs. Preliminary Score Around Grade 3

(a) Δ Medical Expenses



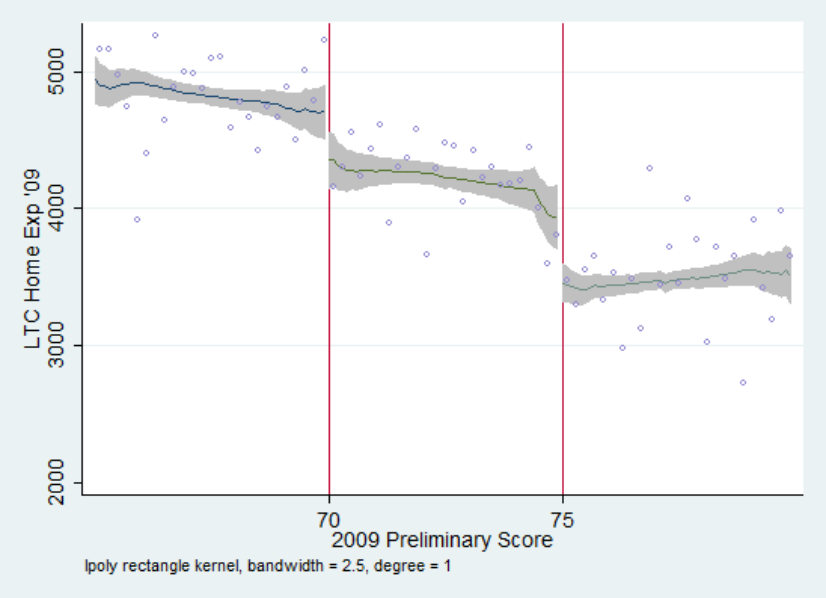
(b) Δ Hospital Expenses



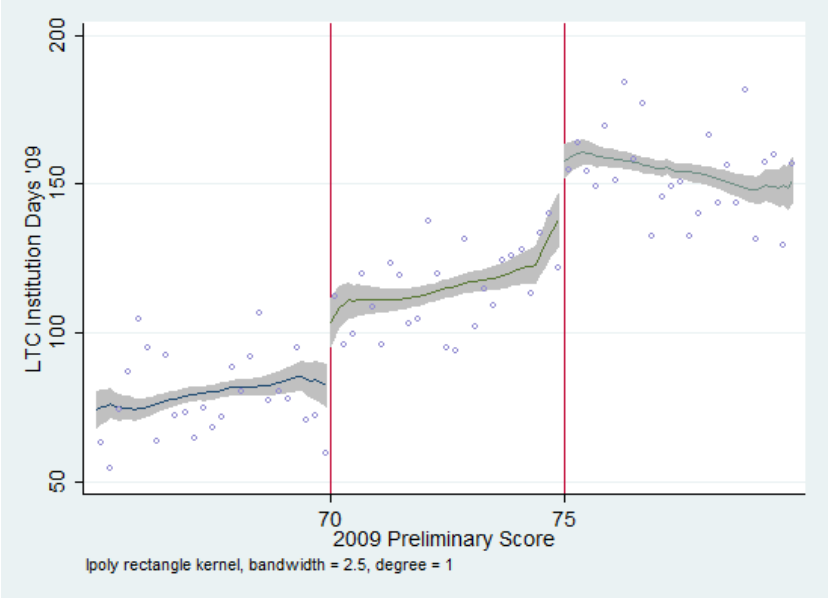
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.10: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 2

(a) Home Care Expenditures



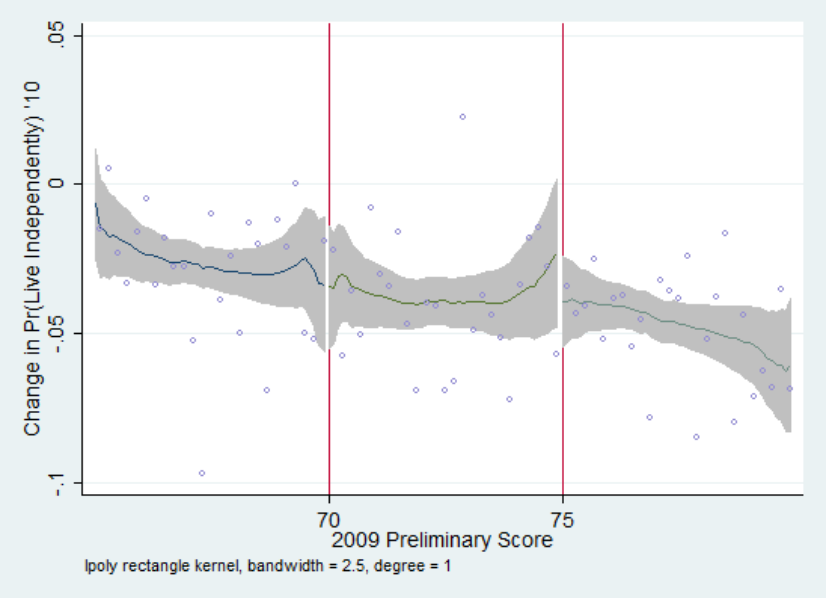
(b) Institutional Days



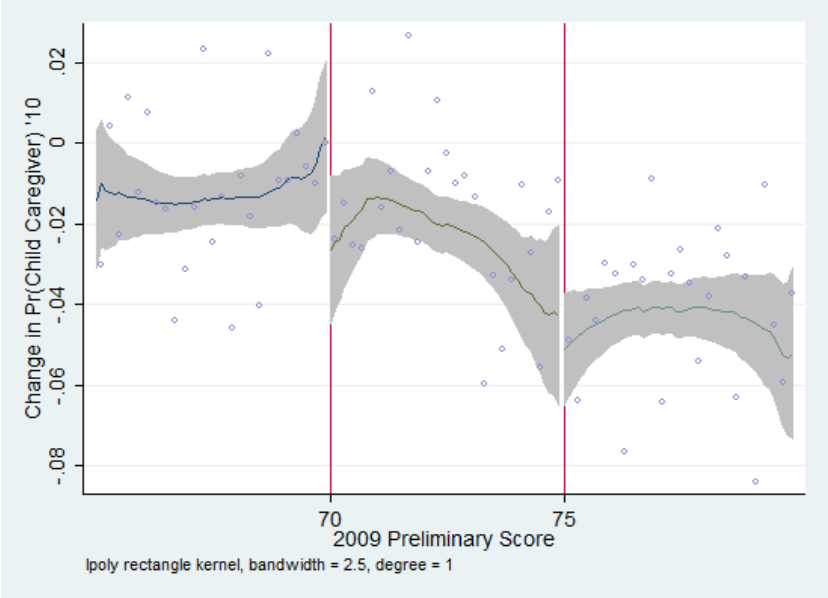
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.11: Change in Informal Care vs. Preliminary Score Around Grade 2

(a) Δ Pr(Live Independently)



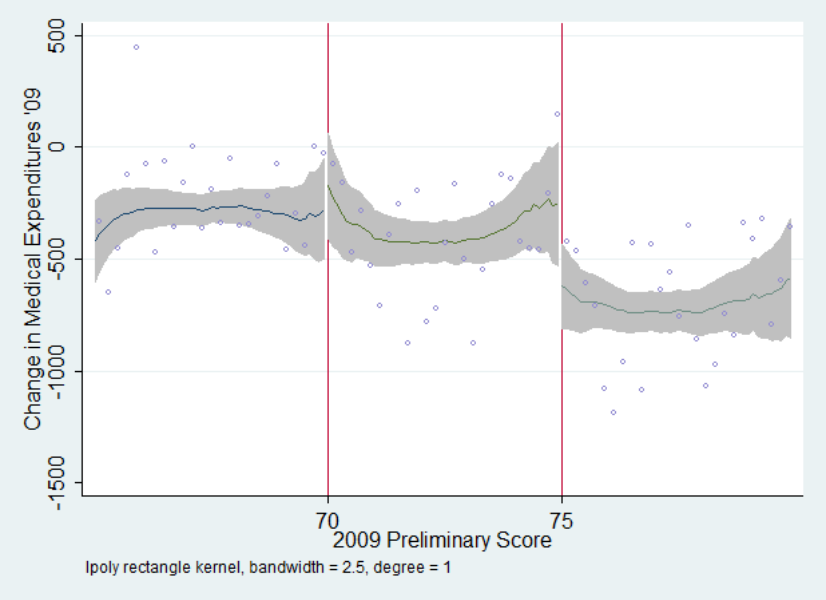
(b) Δ Pr(Child Caregiver)



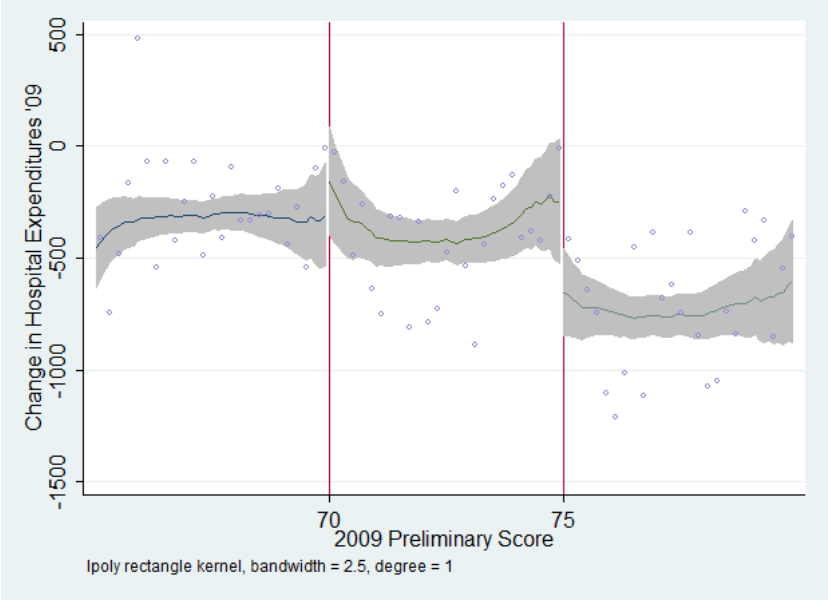
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.12: Change in Medical Utilization vs. Preliminary Score Around Grade 2

(a) Δ Medical Expenses



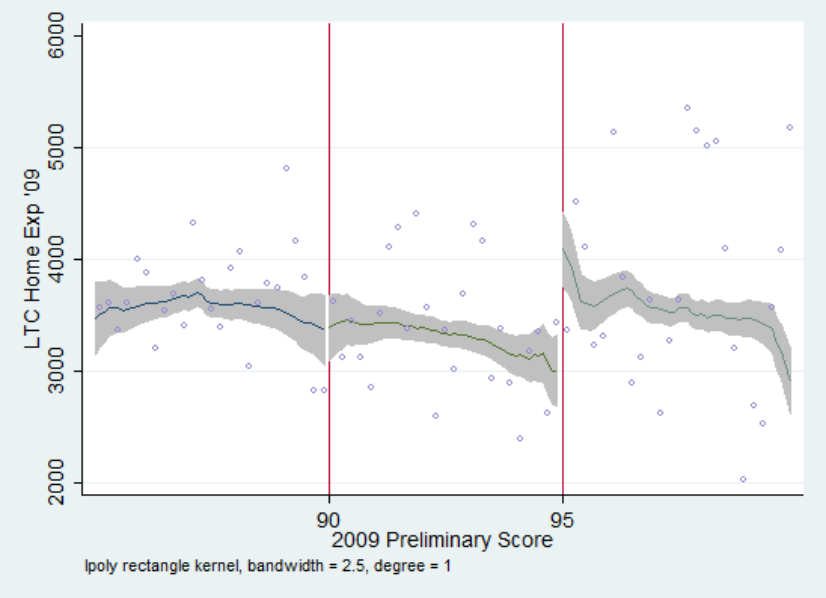
(b) Δ Hospital Expenses



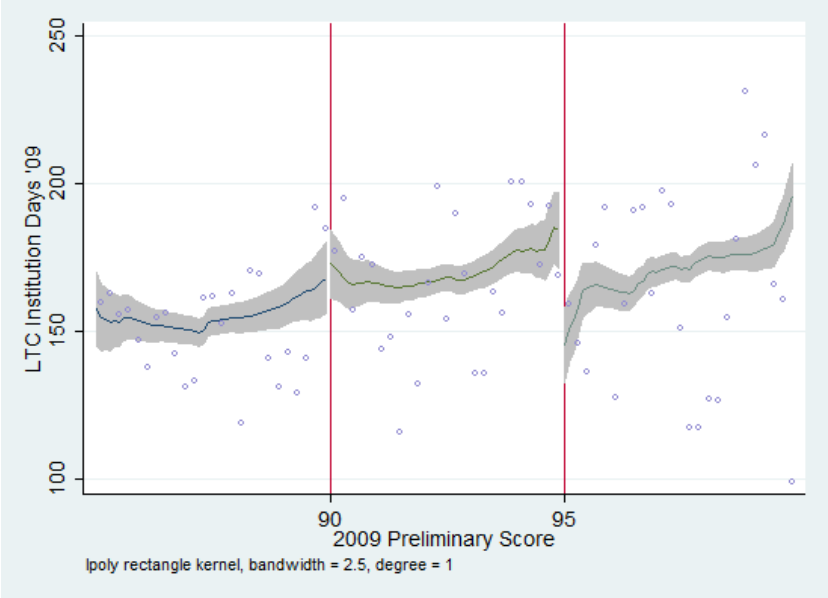
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.13: Reimbursed Formal Care Utilization vs. Preliminary Score Around Grade 1

(a) Home Care Expenditures



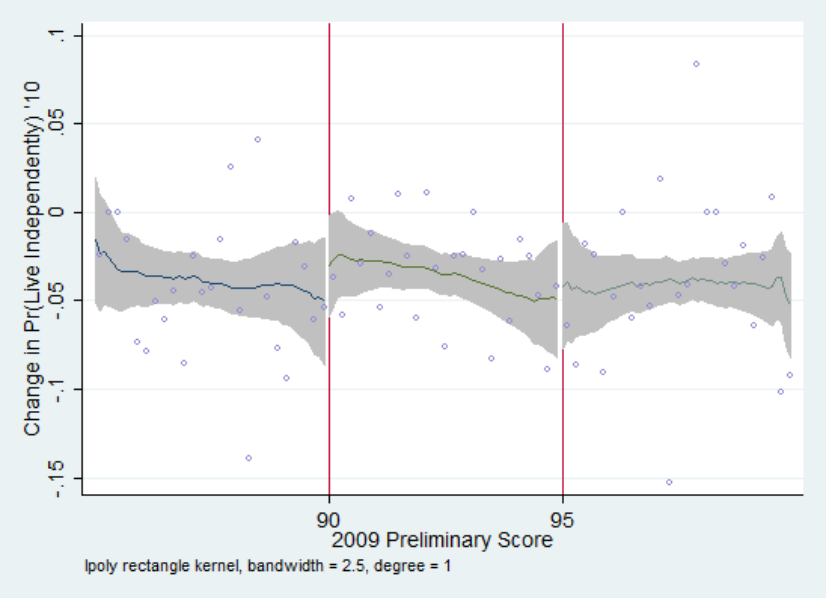
(b) Institutional Days



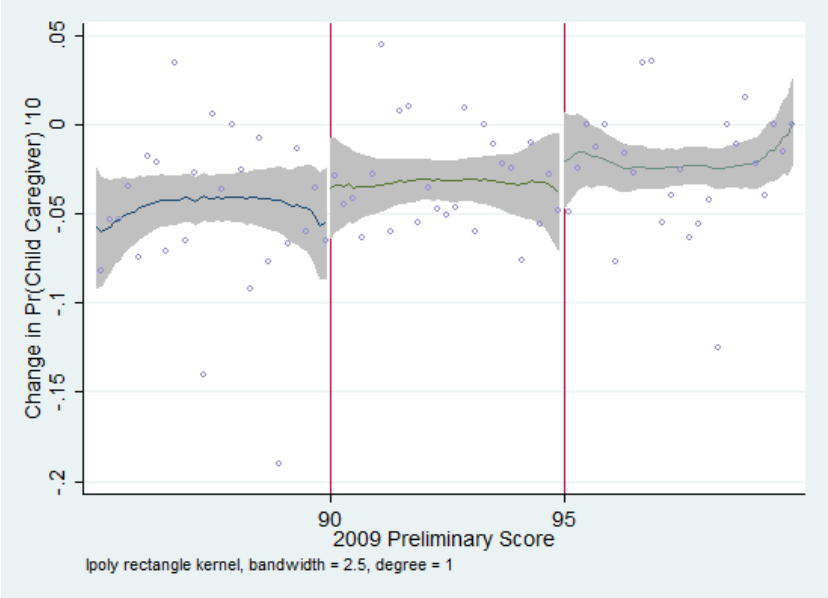
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.14: Change in Informal Care vs. Preliminary Score Around Grade 1

(a) Δ Pr(Live Independently)



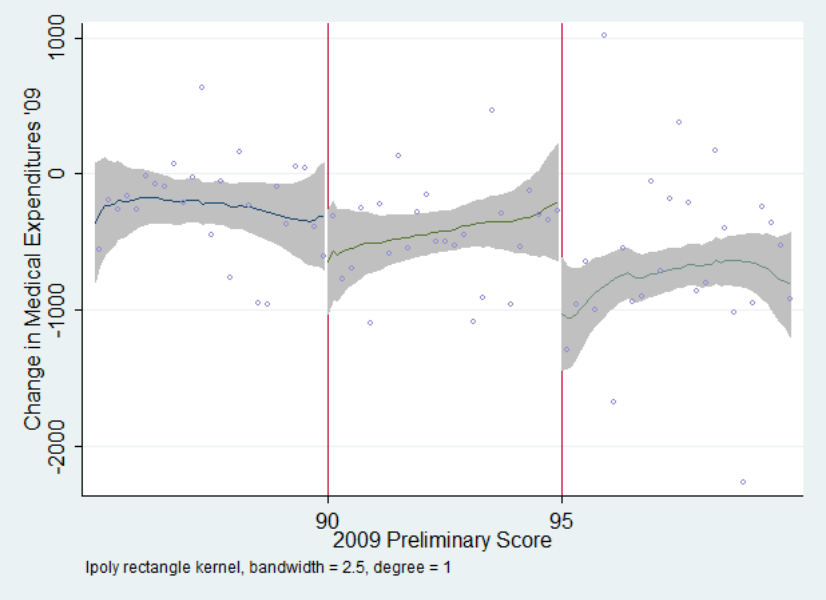
(b) Δ Pr(Child Caregiver)



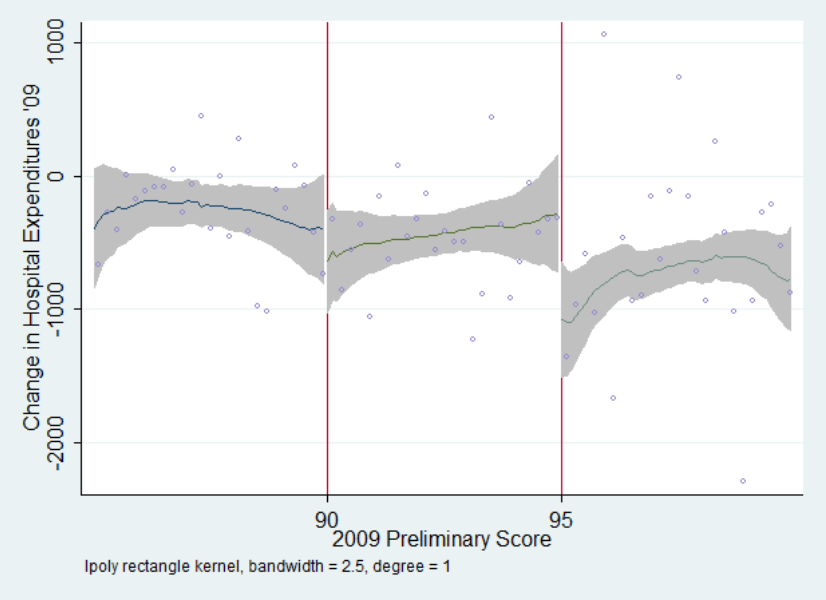
Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Figure 2.15: Change in Medical Utilization vs. Preliminary Score Around Grade 1

(a) Δ Medical Expenses



(b) Δ Hospital Expenses



Notes: The running variable is the 2009 preliminary score. The open circles plot the mean of the dependent variable within 0.2 point bins. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 2.5 points. The shaded regions are 95 percent confidence intervals.

Chapter 3

Is Knowing Half the Battle? The Case of Health Screenings

with Wilfredo Lim

3.1 Introduction

According to the Centers for Disease Control, diabetes is the largest and fastest growing chronic disease in the U.S. Diabetes is characterized by high blood sugar because the body is unable to make or use insulin, resulting in complications such as heart disease, kidney disease, amputations, and blindness. It is frequently not diagnosed until complications appear and as a result almost one-third of all people with diabetes may be undiagnosed (CDC (2011)). Moreover, there is evidence for the benefits of early treatment of diabetes or prediabetes diagnosed through usual clinical care. These facts suggest the possible benefits from screenings of asymptomatic individuals. This is the first study based on a quasi-experimental design to analyze the impact of screening for diabetes in the general population. While we do not assess the effectiveness of screening versus not screening, we do assess the impact of information obtained from screening.

The motivation behind screening in general is that early detection and intervention can increase the likelihood of reducing complications from disease or eliminating the disease altogether. However, the effectiveness of screening depends on many factors, including the reliability of the relevant diagnostic test, the effectiveness of treatments in terms of both efficacy and quality of life, and costs. An example of a disease for which the potential risks are thought to outweigh the benefits of screening is prostate cancer. This is due to false positives, benign cancers, and serious side effects (USPSTF (2008)). In the case of diabetes, the U.S. Preventive Services Task Force currently makes no recommendation for diabetes screening for individuals without high blood pressure, due to the lack of sufficient evidence.

In the case of diabetes, for which exercise and diet play a substantial role in the management of the disease and prevention of complications, an important component determining the benefit of screening is the degree to which results from the screening lead individuals to undergo changes in behavior. In light of this, whether and how individuals process the information from screenings is an important question. Much literature has been devoted to understanding the role of education in the production of health (Cutler and Lleras-Muney (2006), Grossman (2006), Lange (2011), Lleras-Muney (2005)). This paper is complementary to this literature by accounting for factors such as education that are correlated with both the attainment of information and their responses to that

information. In addition to assessing patients' responses and outcomes due to that information we also study whether such responses vary according to a proxy for education.

We address these questions by exploiting a unique program where individuals undergo screening for various health measures, including blood sugar for diabetes, and then receive notification of their health status classification—either "normal", "risk group", or "suspected disease". These classifications vary discontinuously at different blood sugar thresholds which enable us to assess individuals' responses to those classifications while controlling for unobservable factors correlated with both the attainment of information and responses to that information. We assess longer term outcomes which are expected to be affected by screening such as future blood sugar level and mortality. We also assess behavioral responses including follow-up clinic visit after screening, future screening take-ups, outpatient days, and total medical expenditures. Lastly, we assess whether responsiveness varies by education, in order to shed light on the role of education in processing information.

We find that individuals who are classified as "risk group" have no different behavioral response and health outcome within five years of screening than those who are classified as "normal". However, we find that "disease suspected" classification leads to increase of clinic visit for the secondary examinations and future screening take-ups, and decrease of outpatient days and medical expenditure, however no impact on future blood sugar level and mortality. When assessing differences in this outcome by insurance contribution (a proxy for education), we see that those in the highest quintile respond less than those in the other quintiles, which may be explained by the fact that more educated (higher insurance contribution) individuals respond less to the classification because they have already largely incorporated this information from the blood sugar measure itself.

The remainder of this paper is organized as follows. Section 3.2 describes the institutional context and the screening program which creates the setting for this analysis. Section 3.3 describes the data and Section 3.4 the empirical framework. Section 3.5 presents the results. Section 3.6 concludes.

3.2 Institutional Details

Korea provides universal health care. Individuals are covered either by National Health Insurance (NHI) or Medical Care Assistance (MCA), though both programs are overseen by the National Health Insurance Corporation (NHIC). The primary distinction between NHI and MCA is that the latter serves poor individuals.

NHI operates the National Health Screening Program (NHSP), which since 1995 has provided general screening services to people ages 40 and over free of charge every two years. People born in odd-numbered years are encouraged to undergo screening in odd-numbered years and vice versa. NHSP consists of the recording of medical history; measuring of height, weight, blood pressure, vision, and hearing; chest X-ray; urine sample; blood test, including hemoglobin, cholesterol, and rGTP; oral examination; and counseling.[1] Individuals are notified of the screening results by mail. In the report, patients are informed of their blood sugar level. They are also informed of "normal" and "at risk" levels of blood sugar, which are "under 111" and "111-120", respectively.[2] Other measures are reported similarly. In addition to the results of each individual measure, individuals are also informed of whether their overall results classify them as "normal", "risk group", or "disease suspected". If one's blood sugar falls between 111 and 120, this increases the unconditional probability of receiving a "risk group" notification relative to a blood sugar level less than 111. If blood sugar exceeds 120, this increases the unconditional probability of being classified as "disease suspected" (and inherently reduces the probability of being classified "risk group"). The actual notification depends on the individual's other measures.[3] Individuals classified as "disease suspected" are eligible for and encouraged to undergo a secondary examination.

---

[1]Screening is distinct from diagnostic testing, which is performed in response to symptoms or signs of disease. The purpose of screening is to identify disease in asymptomatic individuals.

[2]Blood sugar is measured in units of mg/dL. Levels of blood sugar exceeding 120 is associated with diabetes.

[3]Incorporating the other measures, such as blood pressure, is beyond the scope of this paper. However, future work will incorporate this information to isolate the impact of the diabetes specific risk vs. overall health risk. It would also allow comparison of responses to different diseases for which people may have varying degrees of responsiveness to information.

3.3 Data

This study uses a merged dataset combining administrative data from NHI and NHSP. The sample consists of males born in even-numbered years who were eligible for and participated in general screening in 2002. The data spans 2001-2006 and contains information on gender, age, insurance contribution, screening results for individual measures including blood sugar (measured in mg/dL), the overall health classification, whether an individual undertook a secondary examination, annual medical expenditures, annual outpatient days, and mortality.

Our main explanatory variable is the blood sugar level in 2002 (baseline). Our key outcomes of interest are cumulative mortality, medical expenditures and outpatient days through 2006. We are also interested in whether individuals undertook a secondary examination and whether individuals underwent diabetes screening in the next eligible periods. Table 3.1 displays summary statistics of the variables of interest by overall health classification.

3.4 Empirical Framework

The objective of this paper is to assess the impact of receiving different overall health classifications on the health and behavior of patients.[4] To do so, we conduct a regression discontinuity analysis at the 111 and 121 blood sugar levels where the probability of being notified of a particular health status is discontinuous. Specifically, the aim is to compare outcomes across individuals who are effectively identical but for receiving different health status notifications.[5]

The corresponding regression model we estimate is:

$$\text{outcome} = \beta \mathbb{1}\{S \geq \tau\} + f(S) + \epsilon, \tag{3.1}$$

where S is the blood sugar level, $f(S)$ is a function of the blood sugar level, and τ is the relevant cutoff (111 or 121).

[4]In the case of medically related actions, this is likely to be the joint behavior of doctors and their patients.

[5]Because of the nature of the empirical design and data limitations, we are *not* assessing the impact of screening vs. not screening. Rather, we are assessing the impact of the information obtained from screening.

In implementing the regression discontinuity design, an important consideration is the modeling of $f(S)$. One approach is to model it parametrically through linear, quadratic, or higher order polynomials that are allowed to differ on each side of the cutoff. The other approach, which we follow here, is to estimate the discontinuity nonparametrically, which we implement by local linear regression with a rectangular kernel.[6] Our preferred estimates are based on a bandwidth of 5 mg/dL, in order to reduce bias by staying close to the cutoff while still maintaining enough precision.[7]

A critical assumption to our identification strategy is that individuals just below a threshold are indeed comparable to individuals just above a threshold. One potential threat to this assumption is if individuals are able to precisely sort around the threshold (Lee (2008)). It is not likely that individuals are able to precisely manipulate their blood sugar level. However, other features such as measurement may be just as problematic if it led to nonrandom sorting of scores according to some unobservable characteristic if this characteristic were also correlated with our outcomes of interest. An example of this would be if hospitals that served certain types of patients recorded blood glucose levels above 120 mg/dL as 120 mg/dL.

If our original assumption holds, then an implication is that the density of scores should be continuous around the threshold (indeed, everywhere). Figure 3.1 displays the density of blood sugar level for the majority of our sample.[8] Note that there are striking discontinuities in the density at 110 and 120, just at the relevant thresholds. One possible explanation is that there was rounding down (e.g. a score of 114 being recorded as 110), but this would not fully explain the pattern observed (i.e. the reductions in density after 110 and 120).

As discussed in Urquiola and Verhoogen (2009), stacking alone may not violate the regression discontinuity assumptions since violation arises from the interaction of the stacking and the endogenous sorting of individuals. Thus, the more fundamental question for our identification

[6] As noted in Lee and Lemieux (2010), the choice of kernel typically has little impact and while a triangular kernel is boundary optimal, a more transparent way of putting more weight on observations close to the cutoff is to reestimate a rectangular kernel based model using a smaller bandwidth.

[7] Our results are not sensitive to this choice.

[8] Measures outside of [50,150] were removed for visual reasons.

strategy is whether the distribution of predetermined characteristics is identical on each side of the threshold. Figure 3.3 displays the baseline measures of our outcomes of interest (for which there is data) as a function of blood sugar level. For the most part, predetermined characteristics appear to be balanced around each threshold.

To partially address the stacking in the density, we conduct "donut" regression discontinuity analyses by omitting observations with blood sugar levels of 110 and 120. For transparency, in our figures we show the mean estimates of our outcomes at all blood sugar levels, including 110 and 120. In addition to this, I further explore the possibility of systematic reporting error by taking advantage of the fact that size of the extra density jump at the thresholds varies by county. Figure 3.2 displays the density of blood sugar level by the counties where the size of extra density jump is in bottom, middle and top tertiles as well as bottom 10%.

3.5 Results

3.5.1 Classification

Figure 3.4 displays the probability of being classified as "normal" as a function of blood sugar measured at baseline. There is a discrete drop in the probability of "normal" status at 111. The fuzziness arises from the other measures which can also affect status. There is minimal corresponding impact at 121. Figure 3.5 displays the probability of being classified as "risk group" as a function of baseline blood sugar. There is a discrete increase in the probability of being "risk group" at 111. At 121, there is a discrete decrease in the probability of being "risk group". This is due to the fact that these individuals are informed that they likely have the disease. This is shown explicitly in Figure 3.6 which displays the probability of "disease suspected" as a function of baseline blood sugar.

Table 3.2 presents the corresponding estimates and illustrates the change in information at each threshold. At the 111 threshold, there is a 11 percentage point drop in the probability of "normal" status and a complementary 11 percentage point increase in the probability of "risk group" status. At the 121 threshold, there is a 24 percentage point drop in the probability of "risk group" status

and complementary increase of the same magnitude in the probability of "disease suspected" status. Thus, the 111 threshold captures the marginal impact of "risk group" vs "normal" while the 121 threshold captures the marginal impact of "disease suspected" vs. "risk group". Changes in disease status classification are robust across size of extra jump at the cutoffs.

3.5.2 Behavioral Responses

This section illustrates that the changes in disease status classification led to behavior responses including retest, future screening take-ups, annual medical expenses and outpatient days. Individuals who were informed that they were "disease suspected" were encouraged to undergo a secondary examination, but not those with "risk group". Figure 3.7 depicts the probability of retest as a function of baseline blood sugar level. As expected, Table 3.3 indicates that more than half who were informed "disease suspected" followed up for a retest. "Disease suspected" classification also led increase of public health screening take up in the next eligible period (after two years) while no impact for those with "risk group". Consistent with these pattern, we also find that there are no statistically significant impacts of information status on future behavioral responses by "risk group" classification, however "disease suspected" decreases medical expenses and outpatient days.

3.5.3 Future Blood Sugar Level and Mortality

The motivation behind screening, particularly for diabetes, is in order for individuals with chronic conditions to manage their disease and limit the occurrence of future preventable negative health shocks and complications. We seek to assess the impact of screening information on blood sugar level in the future and five year cumulative mortality in this section. Figure 3.8 and Table 3.4 show that there is no apparent impact on these outcomes.

3.5.4 Outcomes by Insurance Contribution

In this section, we would like to explore whether responses to information differ by education. This is interesting for many reasons, but particularly because education may be an important factor in processing information. Because the data does not include information on education, we

use insurance contribution (which is directly proportional to income) as a proxy. The outcome we consider is revisits, mainly because this is the only behavioral response that was statistically significant in the previous section. We focus on the 121 threshold, which is the relevant threshold for this outcome. Table 3.5 summarizes our findings. While there is no particular pattern in the changes in information at 121, there is a fairly clear indication that the top quintile has a lower response to the information than the other quintiles, with only 42% responding with a revisit relative to the 60% of other quintiles. There are several possible explanations. One is that the more responsive among the richer are more educated individuals who are more likely to have already been diagnosed. This explanation is not likely given that there are no substantial differences in information/diagnosis between the top and other quintiles. Another potential explanation is that educated individuals respond less to the (redundant) diagnostic label because they have already largely incorporated this information from the blood sugar measure itself.

3.6 Conclusion

This paper studies the impact of information from screening on health outcomes and behavior. We find that "disease suspected" classification leads to increase clinic visit for the secondary examinations and future screening take-ups, and decrease of outpatient days and medical expenditure, however few impacts on health outcomes such as future blood sugar level and mortality. We also find that the responsiveness to the classifications among the highest income quintiles is lower than among the other quintiles, consistent with more educated individuals incorporating information directly from the blood sugar measure itself.

Table 3.1: Summary Statistics by Blood Sugar Level

Baseline Blood Sugar	**[101,109]**	**[111,119]**	**[121,129]**
# Obs	*86,565*	*38,303*	*13,880*
Normal	0.26	0.07	0.03
	(0.44)	(0.25)	(0.16)
Risk Group	0.33	0.47	0.12
	(0.47)	(0.50)	(0.33)
Disease Suspected	0.41	0.46	0.85
	(0.49)	(0.50)	(0.35)
Insurance	38,362	37,132	37,346
Contribution	(24,941)	(24,461)	(25,314)
Baseline Medical	477	516	590
Expenditures	(822)	(902)	(1,029)
Baseline	0.78	0.84	0.98
Hospital Days	(5.48)	(5.54)	(6.34)
Baseline	14.4	15.3	16.4
Outpatient Days	(17.2)	(18.2)	(19.6)
5 Year Cumulative	0.03	0.04	0.04
Mortality	(0.18)	(0.19)	(0.20)
Medical Expenditures	1,208	1,335	1,504
5 Years Later	(2,771)	(2,914)	(2,863)
Hospital Days	2.22	2.45	2.87
5 Years Later	(14.10)	(15.32)	(17.69)
Secondary Exam	0.22	0.26	0.48
	(0.42)	(0.44)	(0.50)
Outpatient Days	15.8	16.7	17.9
Screening Year	(18.6)	(19.5)	(20.4)
DM Screen	0.77	0.75	0.73
Next Period	(0.42)	(0.43)	(0.44)

Notes: Sample consists of male individuals born in even-numbered years and who participated in general screening in 2002. Categories are based on baseline blood sugar levels. Individuals with blood sugar levels of 110 and 120 were omitted from the analysis as discussed in the text. See text for definitions of variables.

Table 3.2: Impact of Baseline Blood Sugar Level on Notified Status

	Screening Outcome		
	Normal	At risk	Disease
Panel A. 110 Cutoff			
Whole sample	-0.106**	0.110**	-0.006
	(0.006)	(0.005)	(0.004)
Bottom Tertiles	-0.088**	0.094**	-0.006
	(0.014)	(0.009)	(0.022)
Middle Tertiles	-0.104**	0.136**	-0.024
	(0.010)	(0.019)	(0.025)
Top Tertiles	-0.122**	0.099**	0.013
	(0.008)	(0.006)	(0.009)
Bottom 10%	-0.117**	0.060**	0.024
	(0.015)	(0.008)	(0.030)
Panel B. 120 Cutoff			
Whole sample	-0.008*	-0.241**	0.239**
	(0.003)	(0.007)	(0.006)
Bottom Tertiles	-0.015+	-0.252**	0.258**
	(0.007)	(0.013)	(0.008)
Middle Tertiles	-0.014*	-0.207**	0.219**
	(0.004)	(0.014)	(0.018)
Top Tertiles	0.008*	-0.249**	0.233**
	(0.003)	(0.011)	(0.019)
Bottom 10%	-0.009	-0.274**	0.268**
	(0.007)	(0.039)	(0.038)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.3: Impact of Notified Status on Medical Behavior

	Screening take-up				Outpatient Days					Medical Expenditure				
	Retest	Retest Rate	After 2 years	After 4 Years	Year 0	Year 1	Year 2	Year 3	Year 4	Year 0	Year 1	Year 2	Year 3	Year 4
Panel A. 110 Cutoff														
Whole sample	0.002		-12.414	0.015**	0.004*	0.048	0.132	0.154	0.152	0.045	5.692	-0.525	-21.072	-21.464
	(0.005)		(24.864)	(0.004)	(0.002)	(0.291)	(0.299)	(0.363)	(0.276)	(0.281)	(4.798)	(10.920)	(19.960)	(24.885)
Bottom Tertiles	0.002		-37.046+	0.010	0.006+	0.442+	0.383	0.744*	0.698+	0.520	17.384	24.177*	12.864	20.341
	(0.008)		(16.136)	(0.006)	(0.003)	(0.199)	(0.426)	(0.295)	(0.331)	(0.409)	(11.596)	(9.351)	(38.815)	(37.376)
Middle Tertiles	-0.025**		-106.983+	0.031**	0.000	-0.655	-0.719	-0.743	-1.215+	-1.197*	1.951	3.791	-58.956	16.580
	(0.005)		(51.162)	(0.007)	(0.001)	(0.485)	(0.712)	(0.564)	(0.603)	(0.422)	(7.692)	(35.577)	(40.204)	(54.383)
Top Tertiles	0.026+		104.794	0.003	0.007**	0.419	0.796+	0.563	1.087**	0.960	-1.510	-26.316	-14.697	-96.353**
	(0.012)		(59.308)	(0.010)	(0.001)	(0.475)	(0.347)	(0.401)	(0.235)	(0.633)	(8.170)	(23.795)	(12.820)	(25.132)
Bottom 10%	0.059**		-81.550+	0.034	0.010**	1.050	0.545	0.727	0.406	0.453	4.259	42.798+	72.470	2.590
	(0.007)		(39.015)	(0.023)	(0.003)	(0.753)	(0.486)	(0.620)	(0.537)	(0.716)	(28.210)	(19.217)	(41.035)	(42.552)
Panel B. 120 Cutoff														
Whole sample	0.141**	59.0%	0.015*	0.002	-0.588	-0.837**	-1.113**	-1.725**	-1.596**	-31.653*	-41.966**	-70.996*	-103.021*	-117.785**
	(0.008)		(0.005)	(0.015)	(0.356)	(0.167)	(0.306)	(0.303)	(0.306)	(9.883)	(11.841)	(23.459)	(31.021)	(30.821)
Bottom Tertiles	0.170**	65.9%	0.016+	-0.032+	0.066	-0.961	-1.383*	-1.304	-0.078	1.017	-55.275	-19.485	-85.407	-55.781
	(0.011)		(0.008)	(0.014)	(1.016)	(0.620)	(0.525)	(0.907)	(0.481)	(24.318)	(37.183)	(25.048)	(52.437)	(54.190)
Middle Tertiles	0.114**	52.1%	0.028**	0.041	-0.966*	-1.604**	-1.922*	-1.998**	-2.389*	-15.762	-41.007**	-210.878**	-143.218+	-286.214**
	(0.011)		(0.004)	(0.024)	(0.316)	(0.460)	(0.786)	(0.567)	(0.992)	(9.285)	(10.443)	(53.569)	(68.340)	(45.273)
Top Tertiles	0.126**	54.1%	0.007	0.001	-0.976+	-0.179	-0.430	-1.824**	-2.377**	-78.171**	-37.401	-36.421	-87.079+	-57.376
	(0.013)		(0.005)	(0.017)	(0.504)	(0.589)	(0.379)	(0.493)	(0.441)	(14.517)	(27.383)	(37.758)	(40.006)	(37.149)
Bottom 10%	0.193**	72.0%	0.054+	-0.016	-1.338	-1.055	-0.295	-2.447	1.049	-106.111+	-101.326+	-152.164*	-323.855**	-150.227**
	(0.042)		(0.026)	(0.023)	(1.593)	(1.311)	(1.695)	(2.159)	(1.552)	(53.900)	(50.953)	(52.026)	(90.916)	(40.204)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.4: Impact of Notified Status on Outcomes 5 Years After Screening

	Blood sugar level		Mortality
Panel A. 110 Cutoff	After 2 years	After 4 Years	5 Year
Whole sample	0.003	1.055**	0.080
	(0.002)	(0.251)	(0.134)
Bottom Tertiles	0.002	0.728*	-0.928
	(0.006)	(0.245)	(0.640)
Middle Tertiles	-0.002	0.566	0.023
	(0.006)	(0.322)	(0.427)
Top Tertiles	0.010**	1.710+	1.074**
	(0.002)	(0.787)	(0.319)
Bottom 10%	0.015	1.586*	0.572
	(0.023)	(0.624)	(0.329)
Panel B. 120 Cutoff			
Whole sample	-0.324	1.215*	-0.004
	(0.879)	(0.424)	(0.005)
Bottom Tertiles	-2.421	0.634	0.002
	(1.672)	(0.902)	(0.007)
Middle Tertiles	-2.626*	3.469*	-0.021+
	(0.926)	(1.091)	(0.011)
Top Tertiles	2.770*	-0.437	0.004
	(0.843)	(0.486)	(0.007)
Bottom 10%	-4.229+	1.076	0.002
	(2.103)	(1.506)	(0.008)

Notes: Each cell presents estimates of β from local linear regression of Equation (3.1). The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Table 3.5: Revisit Percentage by Income

	Disease	Retest	Retest/Disease
Top Quintile	0.36**	0.15**	0.42**
	(0.03)	(0.03)	(0.08)
2nd Quintile	0.27**	0.17**	0.65**
	(0.03)	(0.03)	(0.11)
3rd Quintile	0.39**	0.24**	0.61**
	(0.04)	(0.04)	(0.07)
4th Quintile	0.34**	0.19**	0.57**
	(0.03)	(0.03)	(0.08)
Bottom Quintile	0.32**	0.19**	0.60**
	(0.03)	(0.03)	(0.09)

Notes: Columns 1 and 2 present estimates of β from local linear regression of Equation (3.1) at the 121 threshold. Column 3 is Column 2 divided by Column 1. Each row is a different income quintile. The running variable is baseline blood sugar level. Rectangular kernel. Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$.

Figure 3.1: Histogram of Baseline Blood Sugar

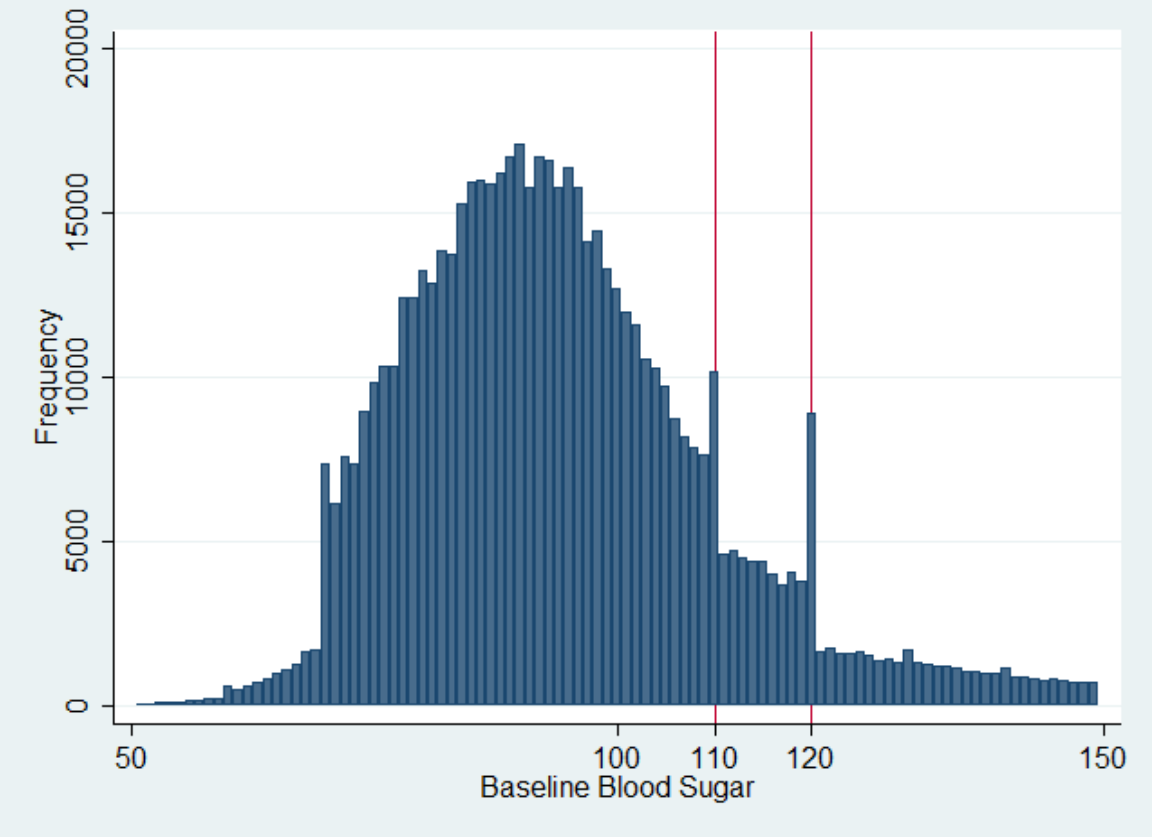
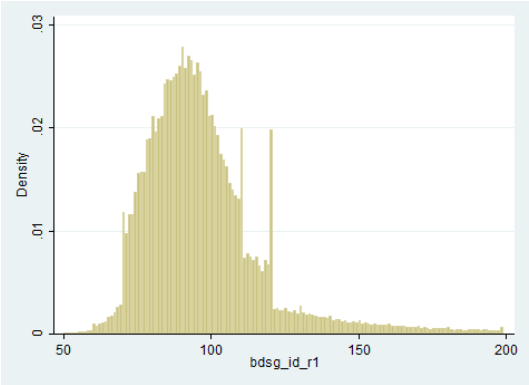
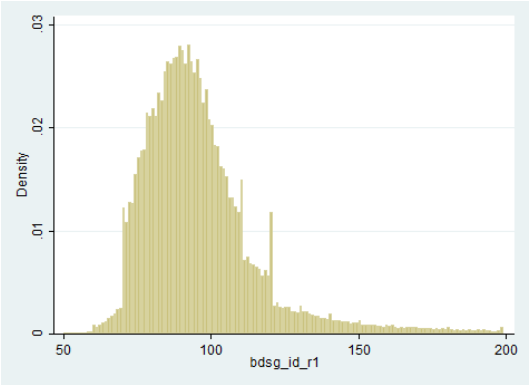


Figure 3.2: Histogram of Baseline Blood Sugar by Counties

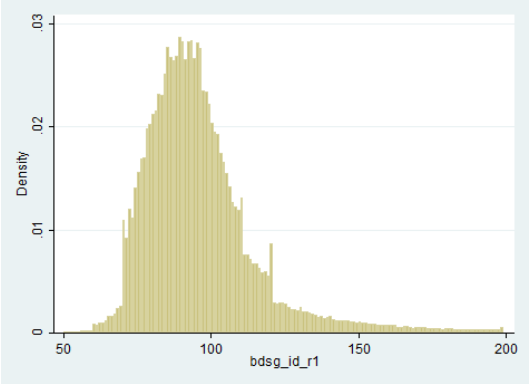
(a) Top Tertiles



(b) Middle Tertiles



(c) Bottom Tertiles



(d) Bottom 10%

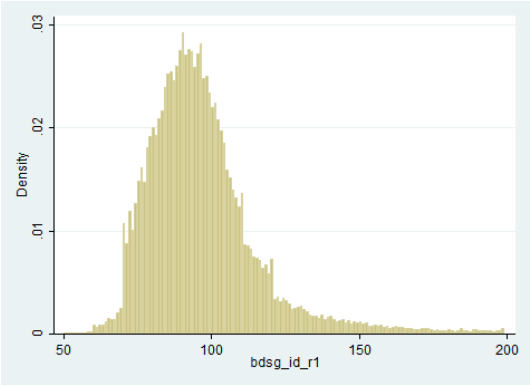
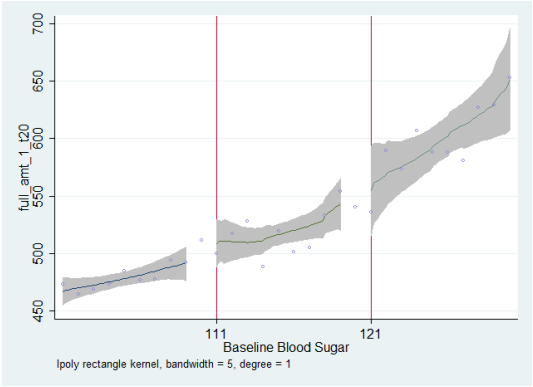
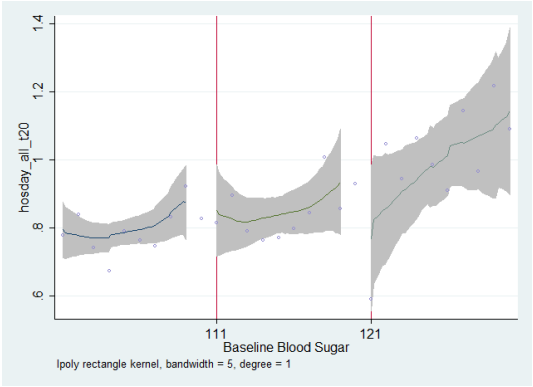


Figure 3.3: Baseline Characteristics

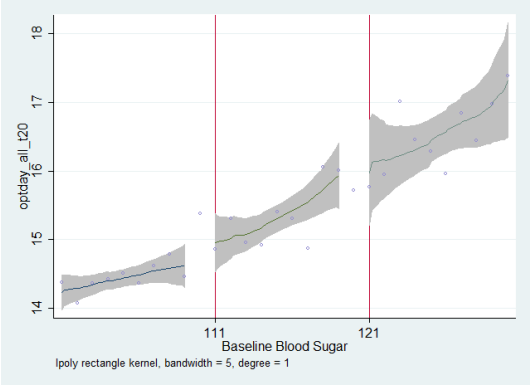
(a) Annual Medical Expenses



(b) Annual Hospital Days

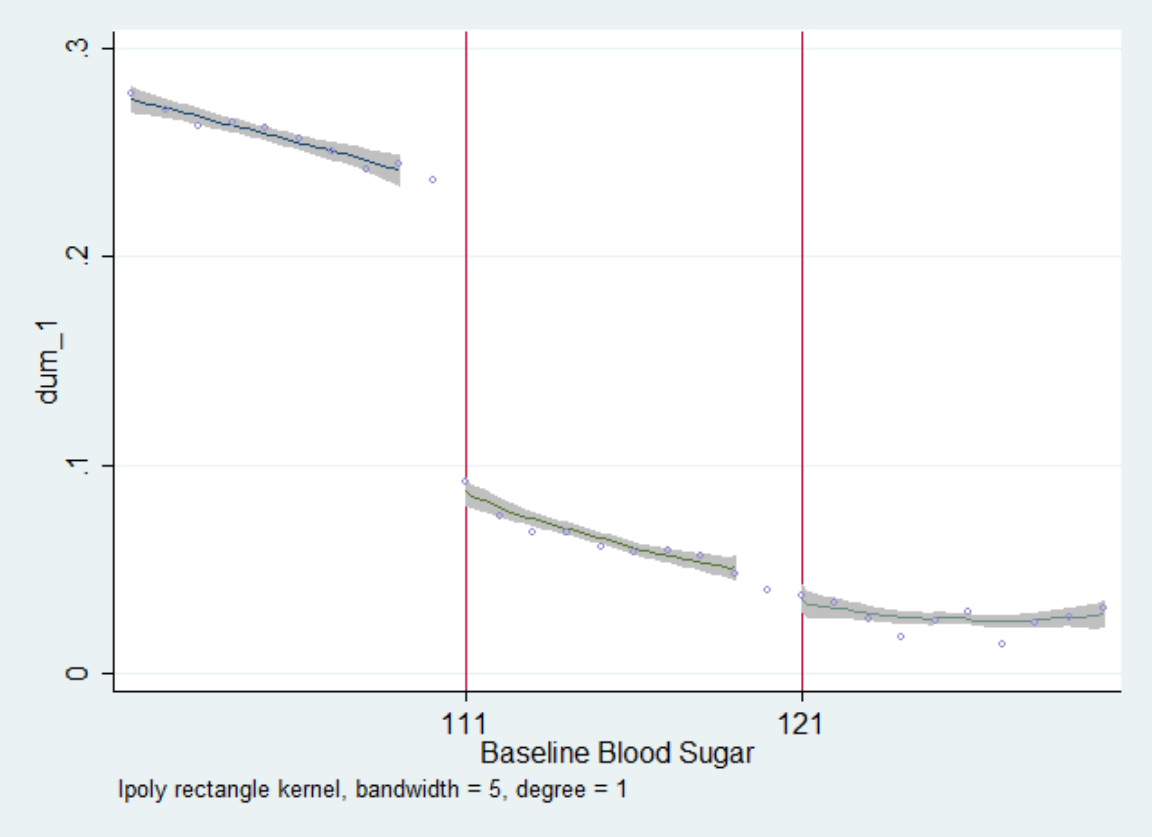


(c) Annual Outpatient Days



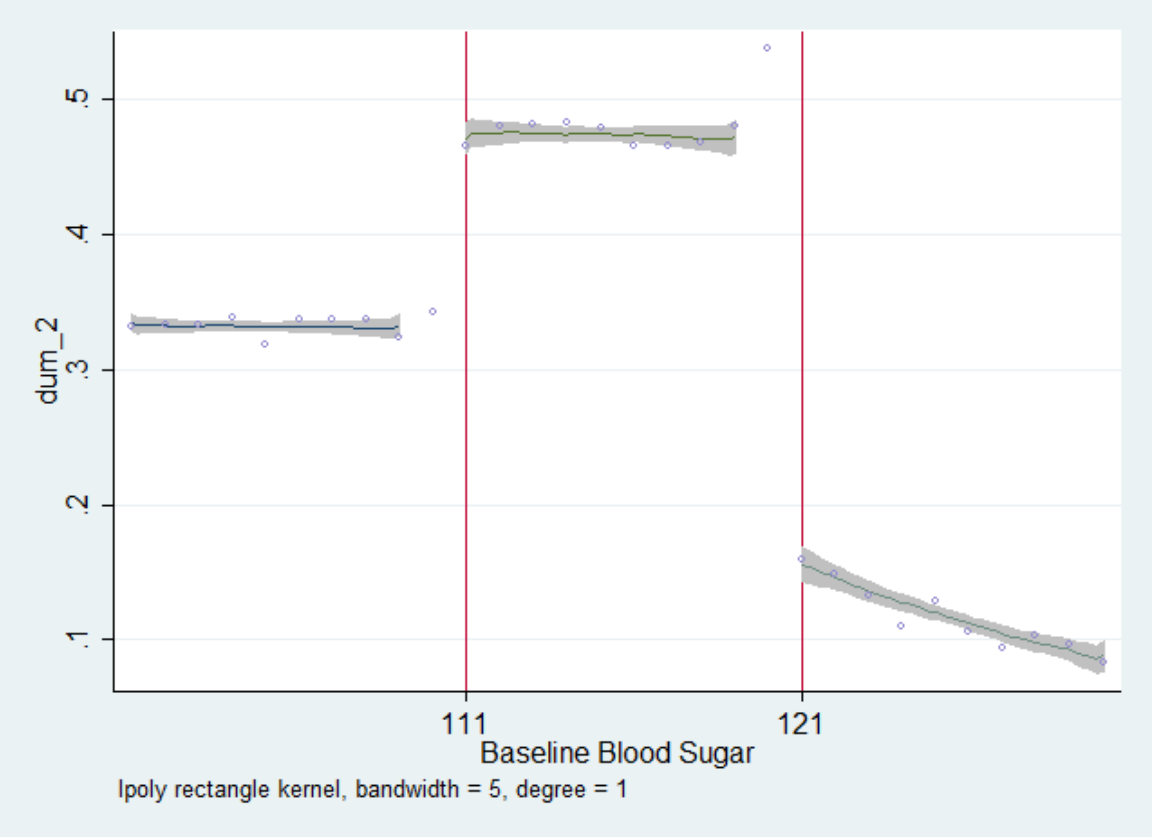
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.4: "Normal" Status



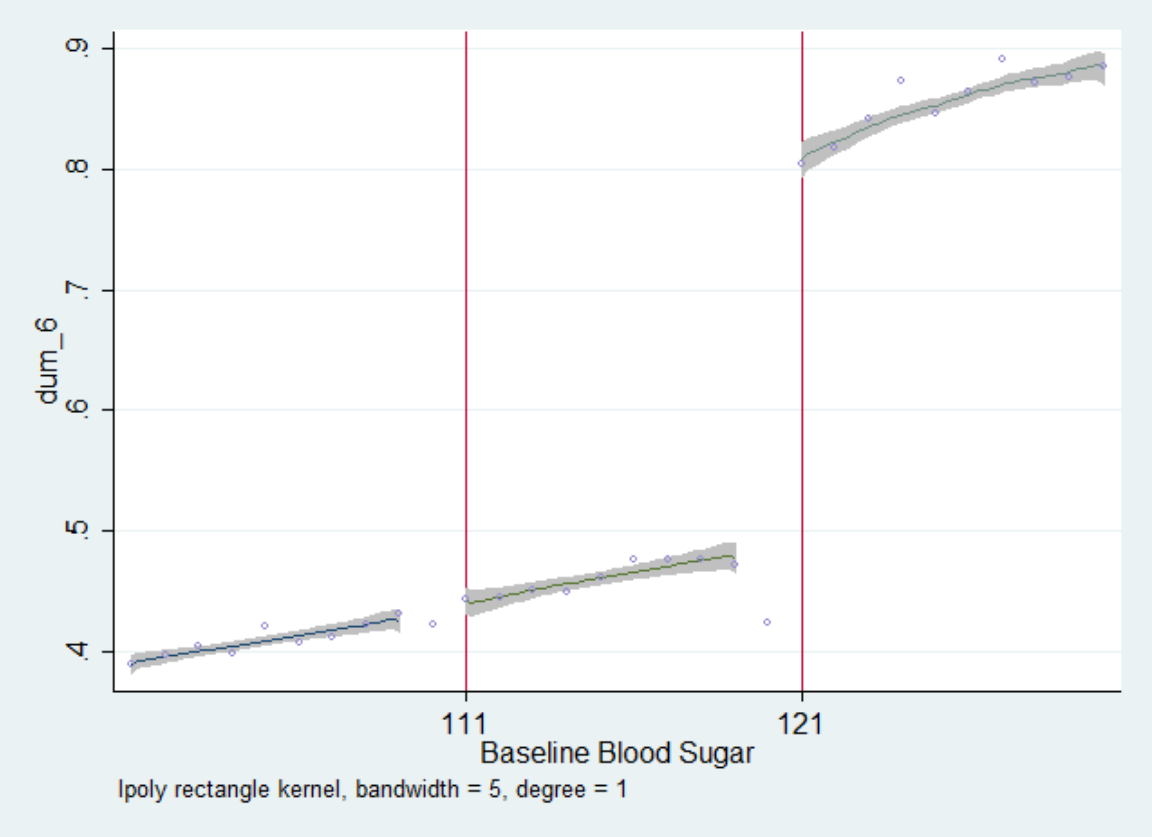
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.5: "Risk Group" Status



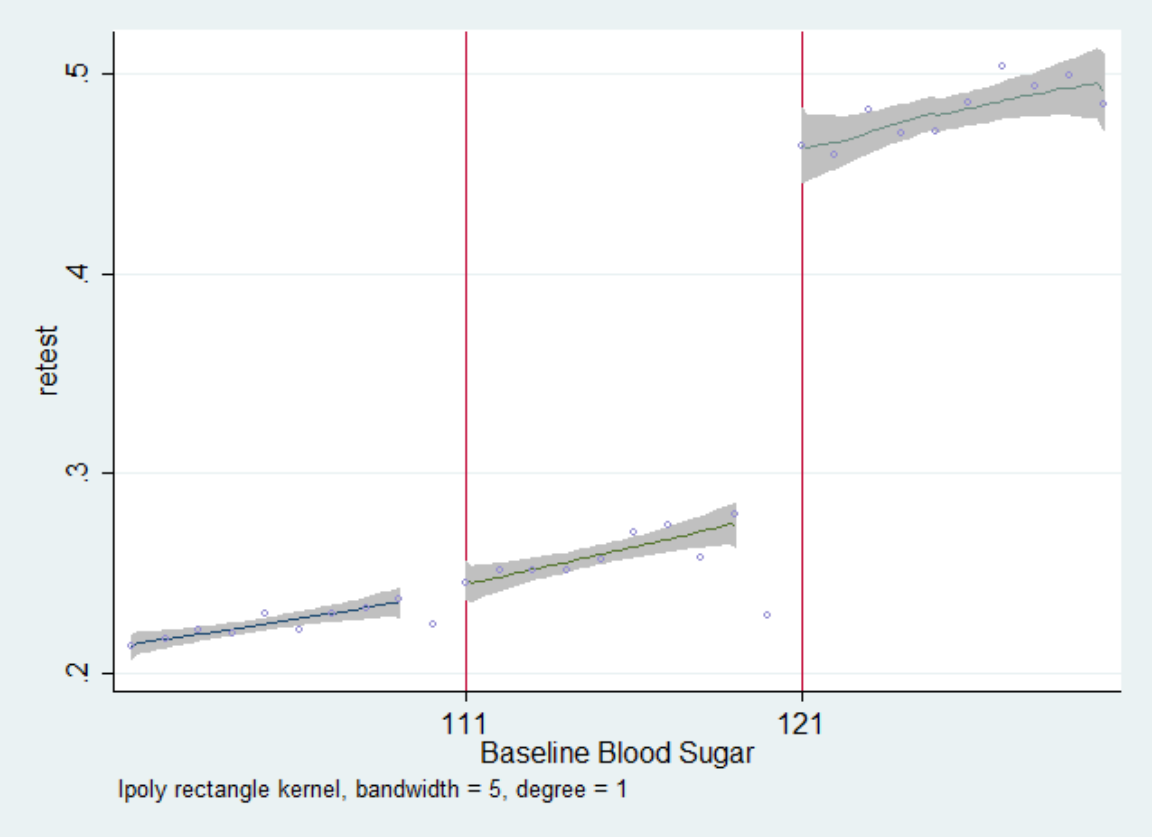
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.6: "Disease Suspected" Status



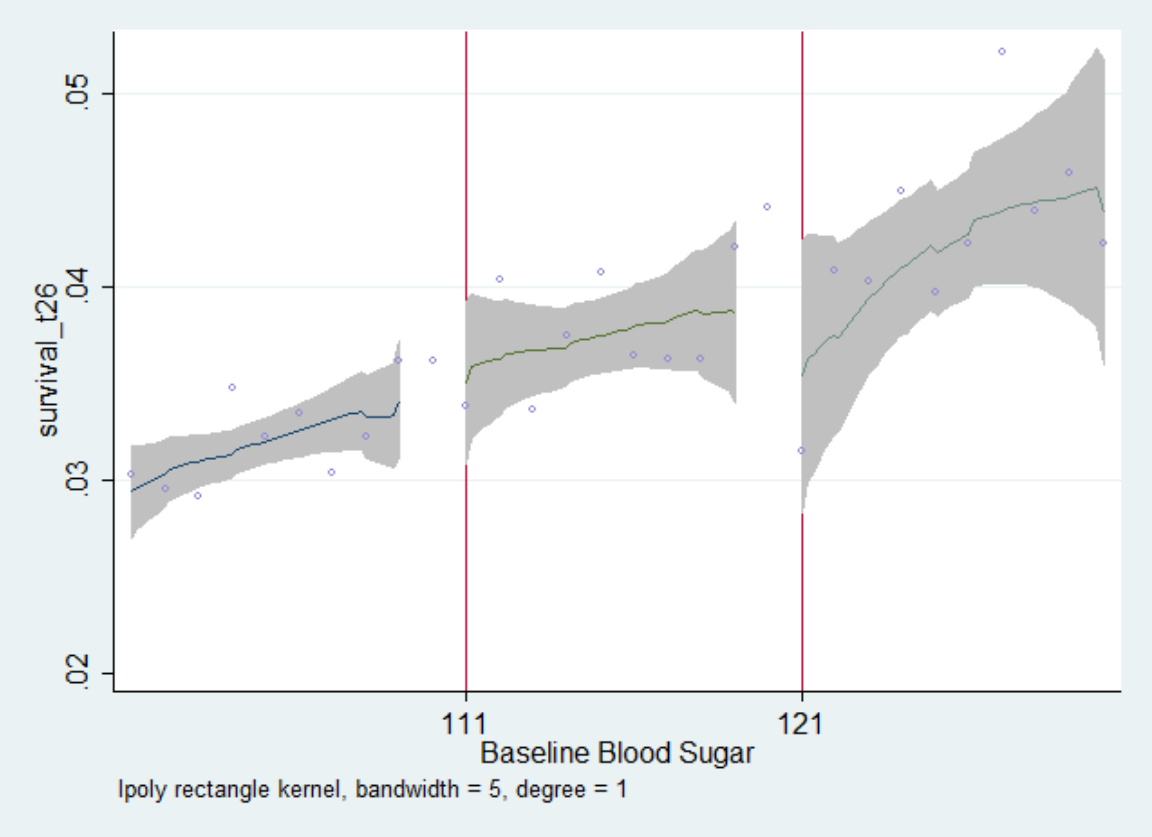
Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.7: Clinic Revisit



Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Figure 3.8: Cumulative Mortality Through 5 Years After Screening



Notes: The running variable is baseline blood sugar level. The open circles plot the mean of the dependent variable at each unit. The solid lines are fitted values from local linear regression of the dependent variable using a rectangular kernel with a bandwidth of 5 mg/dL. The shaded regions are 95 percent confidence intervals.

Bibliography

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

Almond, D. and J.J. Doyle. 2011. "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays." *American Economic Journal: Economic Policy* 3 (3):1–34.

Andriole, G.L., E.D. Crawford, R.L. Grubb III, S.S. Buys, D. Chia, T.R. Church, M.N. Fouad, E.P. Gelmann, P.A. Kvale, D.J. Reding et al. 2009. "Mortality results from a randomized prostate-cancer screening trial." *New England Journal of Medicine* 360 (13):1310–1319.

Ashraf, N., J. Berry, and J.M. Shapiro. 2010. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." *American Economic Review* 100 (5):2383–2413.

Atkin, W.S., R. Edwards, I. Kralj-Hans, K. Wooldrage, A.R. Hart, J. Northover, D.M. Parkin, J. Wardle, S.W. Duffy, J. Cuzick et al. 2010. "Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial." *The Lancet* 375 (9726):1624–1633.

Baicker, K. and D. Goldman. 2011. "Patient Cost-Sharing and Health Care Spending Growth." *Journal of Economic Perspectives* 25 (2):47–68.

Bass, DM and LS Noelker. 1987. "The Influence of Family Caregivers on Elder Use of In-Home Services : An Expanded Conceptual Framework." *Journal of Health and Social Behavior* 28 (2):184–196.

Black, W.C., D.A. Haggstrom, and H.G. Welch. 2002. "All-cause mortality in randomized trials of cancer screening." *Journal of the National Cancer Institute* 94 (3):167–173.

Boyle, P. and B. Levin. 2008. *World cancer report 2008*. World Health Organization.

CBO. 2007. *The Long-Term Outlook for Health Care Spending*. Congressional Budget Office.

CDC. 2011. *National Diabetes Fact Sheet*. Centers for Disease Control.

Chandra, A., J. Gruber, and R. McKnight. 2010. "Patient cost-sharing and hospitalization offsets in the elderly." *The American economic review* 100 (1):193.

Charles, KK and P Sevak. 2005. "Can Family Caregiving Substitute for Nursing Home Care?" *Journal of Health Economics* 24 (6):1174–1190.

Christianson, JB. 1988. "The Evaluation of the National Long-Term Care Demonstration .6. The Effect of Channeling on Informal Caregiving." *Health Services Research* 23 (1):99–117.

Cohen, J. and P. Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment*." *Quarterly Journal of Economics* 125 (1):1.

Colombo, F., A. Llena-Nozal, J. Mercier, and F. Tjadens. 2011. *Help Wanted? Providing and Paying for Long-Term Care*. OECD Health Policy Studies. OECD Publishing.

Cutler, David M. and Adriana Lleras-Muney. 2006. "Education and Health: Evaluating Theories and Evidence." Working Paper 12352, National Bureau of Economic Research.

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

Engelhardt, Gary V and Nadia Greenhalgh-Stanley. 2010. "Home Health Care and the Housing and Living Arrangements of the Elderly." *Journal of Urban Economics* 67 (2):226–238.

GAO. 2005. *Long-Term Care Financing: Growing Demand and Cost of Services are Straining Federal and State Budgets*. U.S. Government Accountability Office.

Golberstein, E, DC Grabowski, KM Langa, and ME Chernew. 2009. "Effect of Medicare Home Health Care Payment on Informal Care." *Inquiry* 46 (1):58–71.

Goldman, D.P., G.F. Joyce, and Y. Zheng. 2007. "Prescription drug cost sharing." *Journal of the American Medical Association* 298 (1):61–69.

Gross, C.P. 2011. "Screening with low-dose computed tomography reduced lung cancer mortality in high-risk patients." *N Engl J Med* 365:395–409.

Grossman, Michael. 2006. "Education and Nonmarket Outcomes." Elsevier, 577 – 633.

Hahn, J., P. Todd, and W. Van der Klaauw. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69 (1):201–209.

Hsu, J., M. Price, R. Brand, G.T. Ray, B. Fireman, J.P. Newhouse, and J.V. Selby. 2006. "Cost-Sharing for Emergency Care and Unfavorable Clinical Events: Findings from the Safety and Financial Ramifications of ED Copayments Study." *Health Services Research* 41 (5):1801–1820.

Imbens, G. and K. Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies* 79 (3):933–959.

Imbens, Guido and Karthik Kalyanaraman. 2009. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." Working Paper 14726, National Bureau of Economic Research.

Kadiyala, S. and E. Strumpf. 2011. "How Effective is Population-Based Cancer Screening? Regression Discontinuity Estimates from the US Guideline Screening Initiation Ages." *Regression Discontinuity Estimates from the US Guideline Screening Initiation Ages (July 1, 2011)* .

Kim, J., S. Ko, and B. Yang. 2005. "The effects of patient cost sharing on ambulatory utilization in South Korea." *Health Policy* 72 (3):293–300.

Kim, Y., J.K. Jun, K.S. Choi, H.Y. Lee, E.C. Park et al. 2011. "Overview of the National Cancer screening programme and the cancer screening status in Korea." *Asian Pac J Cancer Prev* 12 (3):725–730.

Lange, Fabian. 2011. "The Role of Education in Complex Health Decisions: Evidence from Cancer Screening." *Journal of Health Economics* 30 (1):43–54.

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in US House Elections." *Journal of Econometrics* 142 (2):675–697.

Lee, David S and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Econometric Literature* 48 (2):281–355.

Lee, D.S. and D. Card. 2008. "Regression discontinuity inference with specification error." *Journal of Econometrics* 142 (2):655–674.

Lleras-Muney, A. 2005. "The Relationship Between Education and Adult Mortality in the United States." *Review of Economic Studies* 72 (1):189–221.

Lo Sasso, AT and RW Johnson. 2002. "Does Informal Care from Adult Children Reduce Nursing Home Admissions for the Elderly?" *Inquiry* 39 (3):279–297.

Manning, W.G., J.P. Newhouse, N. Duan, E.B. Keeler, and A. Leibowitz. 1987. "Health insurance and the demand for medical care: evidence from a randomized experiment." *The American Economic Review* :251–277.

McKnight, R. 2006. "Home Care Reimbursement, Long-Term Care Utilization, and Health Outcomes." *Journal of Public Economics* 90 (1-2):293–323.

NCC. 2009. *National Cancer Center, National Cancer Screening Survey Report (In Korean), 2009.* National Cancer Center of Korea.

NCI. 2007. *Cancer Trends Progress Report 2007 Update, NIH, DHHS, Bethesda, MD, http://progressreport.cancer.gov.* National Cancer Institute.

Newhouse, Joseph P. and Rand Corporation. Insurance Experiment Group. 1993. *Free for all?: lessons from the RAND health insurance experiment.* Harvard University Press.

Newhouse, J.P., W.G. Manning, C.N. Morris, L.L. Orr, N. Duan, E.B. Keeler, A. Leibowitz, K.H. Marquis, M.S. Marquis, C.E. Phelps et al. 1981. "Some interim results from a controlled trial of cost sharing in health insurance." *New England Journal of Medicine* 305 (25):1501–1507.

NHS. 2008. *NHS Breast Screening Programme Annual Review.* NHS Breast Screening Programme.

Oken, M.M., W.G. Hocking, P.A. Kvale, G.L. Andriole, S.S. Buys, T.R. Church, E.D. Crawford, M.N. Fouad, C. Isaacs, D.J. Reding et al. 2011. "Screening by chest radiograph and lung cancer mortality." *JAMA: the journal of the American Medical Association* 306 (17):1865–1873.

Orsini, Chiara. 2010. "Changing the Way the Elderly Live: Evidence from the Home Health care Market in the United States." *Journal of Public Economics* 94 (1-2):142–152.

Pezzin, LE, P Kemper, and J Reschovsky. 1996. "Does Publicly Provided Home Care Substitute for Family Care? Experimental Evidence with Endogenous Living Arrangements." *Journal of Human Resources* 31 (3):650–676.

Schopper, D. and C. de Wolf. 2009. "How effective are breast cancer screening programmes by mammography? Review of the current evidence." *European journal of cancer* 45 (11):1916–1923.

Schröder, F.H., J. Hugosson, M.J. Roobol, T.L.J. Tammela, S. Ciatto, V. Nelen, M. Kwiatkowski, M. Lujan, H. Lilja, M. Zappa et al. 2009. "Screening and prostate-cancer mortality in a randomized European study." *New England Journal of Medicine* 360 (13):1320–1328.

Shampanier, K., N. Mazar, and D. Ariely. 2007. "Zero as a special price: The true value of free products." *Marketing Science* 26 (6):742–757.

Soldo, BJ. 1985. "In-Home Services for the Dependent Elderly: Determinants of Current Use and Implications for Future Demand." *Research on Aging* 7 (2):281–304.

Stabile, Mark, Audrey Laporte, and Peter C. Coyte. 2006. "Household Responses to Public Home Care Programs." *Journal of Health Economics* 25 (4):674–701.

Swartz, K. 2010. "Cost-sharing: effects on spending and outcomes." *The Synthesis project. Research synthesis report* (20).

Urquiola, Miguel and Eric Verhoogen. 2009. "Class-Size Caps, Sorting, and the Regression-Discontinuity Design." *American Economic Review* 99 (1):179–215.

USPSTF. 2008. "Screening for Prostate Cancer: US Preventive Services Task Force Recommendation Statement." *Annals of Internal Medicine* 149 (3):185–191.

Van Houtven, CH and EC Norton. 2004. "Informal Care and Health Care Use of Older Adults." *Journal of Health Economics* 23 (6):1159–1180.

Ward, Derek, Amy Drahota, Diane Gal, Martin Severs, and Taraneh P. Dean. 2008. "Care Home Versus Hospital and Own Home Environments for Rehabilitation of Older People." *Cochrane Database of Systematic Reviews* (4).

Wennberg, J.E. and M.M. Cooper. 1996. *The Dartmouth atlas of health care*. American Hospital Publishing.

Wolinsky, FD, RR Mosely, and RM Coe. 1986. "A Cohort Analysis of the Use of Health Services by Elderly Americans." *Journal of Health and Social Behavior* 27 (3):209–219.

Wooldridge, J and J Schore. 1988. "The Evaluation of the National Long-Term Care Demonstration .7. The Effect of Channeling on the Use of Nursing Homes, Hospitals, and Other Medical Services." *Health Services Research* 23 (1):119–127.

Zauber, A.G., S.J. Winawer, M.J. O'Brien, I. Lansdorp-Vogelaar, M. van Ballegooijen, B.F. Hankey, W. Shi, J.H. Bond, M. Schapiro, J.F. Panish et al. 2012. "Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths." *New England Journal of Medicine* 366 (8):687–696.

Appendix A

Chapter 1 Appendix

Table A.1: The Price of Cancer Screening

Unit($)

	Year	2002	2003	2004
	Administrarion cost	3.70	3.81	4.16
Stomach cancer screening	UGI	33.34	34.13	34.88
	EGD	33.30	34.28	35.20
	Biopsy	20.73	24.16	24.81
Breast cancer screening	Mammography	12.50	18.31	18.76
	Biopsy	24.02	29.70	30.50

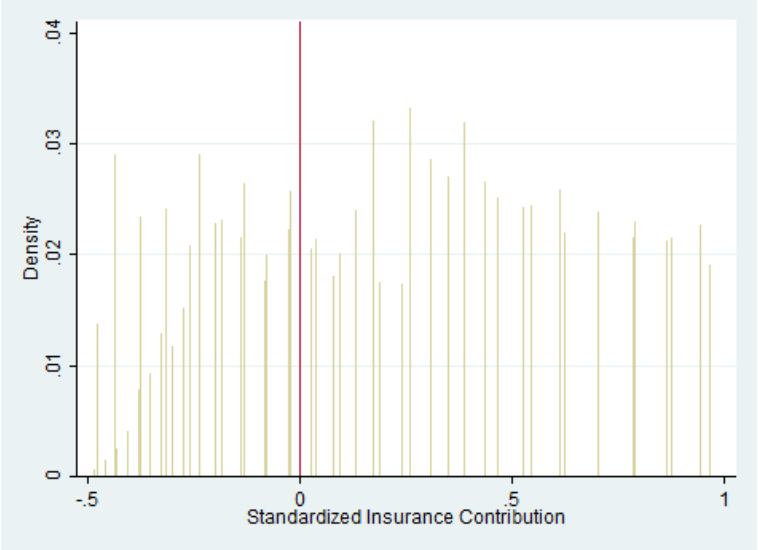
Table A.2: Smoothness around Cutoff

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A. Male	Age		Employment status		Grenal health screening (R1)		Height (Cm)		Obesity (Round 1)		BMI (Round 1)	
Bandwidth	0.3	IK(0.25)	0.3	IK(0.25)	0.3	IK (0.23)	0.3	IK(0.11)	0.3	IK(0.13)	0.3	IK(0.11)
	-0.4545*	-0.3793	0.0191	0.0157	0.0007	0.0093	-0.0232	0.0486	0.0012	0.0143**	-0.0344	0.0635*
	(0.211)	(0.238)	(0.019)	(0.021)	(0.012)	(0.015)	(0.084)	(0.126)	(0.007)	(0.002)	(0.048)	(0.020)
N	1,260,729	935,784	1,260,729	1,066,081	1,260,729	935,785	527,997	195,261	527,725	195,154	527,725	195,152
Panel B. Female	Age		Employment status		Grenal health screening (R1)		Height (Cm)		Obesity (Round 1)		BMI (Round 1)	
Bandwidth	0.3	IK(0.25)	0.3	IK(0.25)	0.3	IK (0.09)	0.3	IK(0.15)	0.3	IK(0.12)	0.3	IK(0.12)
	-0.2994	-0.0799	0.0042	-0.0188	0.0534**	0.0329	-0.0903+	-0.0280	0.0096*	0.0056	0.0605*	0.0354+
	(0.180)	(0.170)	(0.025)	(0.022)	(0.015)	(0.024)	(0.044)	(0.064)	(0.004)	(0.004)	(0.027)	(0.016)
N	1,396,081	1,178,589	1,396,081	1,178,589	1,396,081	445,477	408,967	206,229	408,623	137,755	408,623	137,755

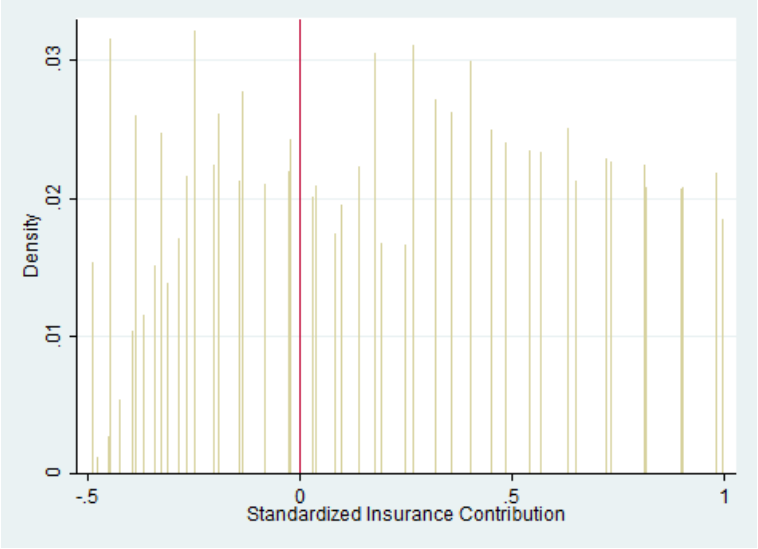
Note: Each cell represents a coefficient β from different local linear regression of equation (1.1). The running variable is the standardized insurance contribution. Rectangular kernel is used. Robust standard errors clustered at standardized insurance contribution in parentheses. **, * and + indicates statistical significance at the 1, 5 and 10 percent level respectively.

Figure A.1: Density of Insurance Contribution

Panel A. Histogram, Male



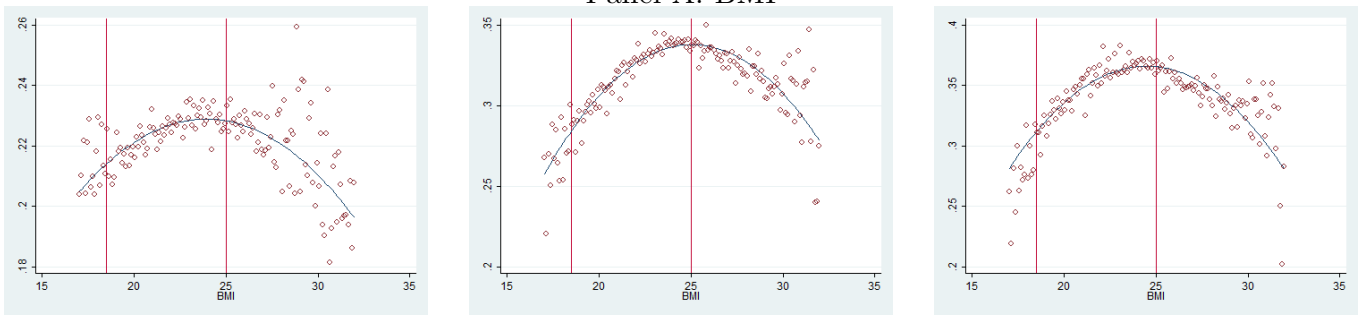
Panel B. Histogram, Female



Note: In each panel histogram with smallest bin size and 0.05 are presented

Figure A.2: Probability of Cancer Screening Take up by Health Status

Panel A. BMI

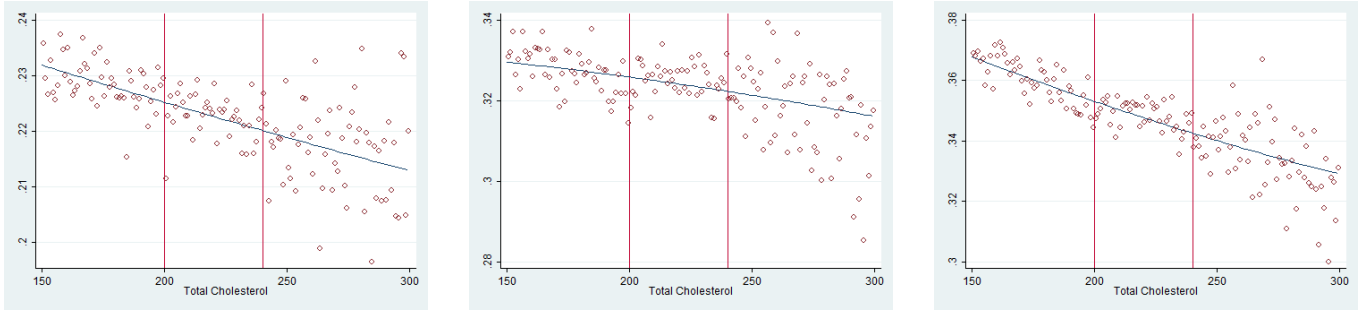


Stomach cancer, Male | Stomach cancer, Female | Breast cancer, Female

Panel B. Blood sugar level

Stomach cancer, Male | Stomach cancer, Female | Breast cancer, Female

Panel C. Total Cholesterol

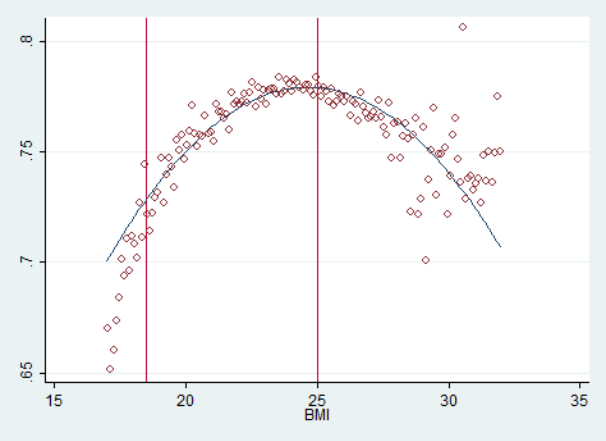


Stomach cancer, Male | Stomach cancer, Female | Breast cancer, Female

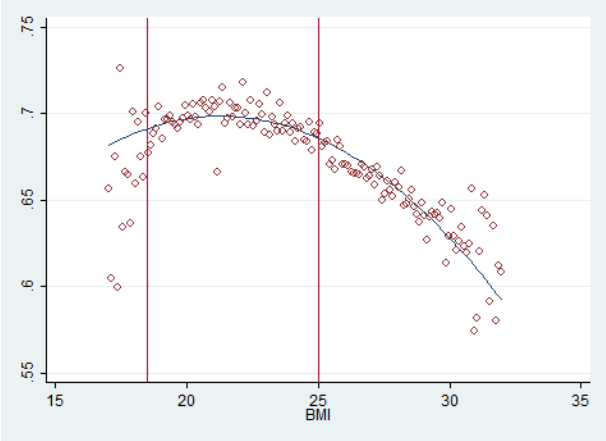
Note: Each figure shows probability of general screening take up in the second round by BMI, blood sugar and cholesterol level in the first round. Normal range of BMI is between 18.5 and 25. Normal level of blood sugar is under 110, and DM is diagnosed if it is greater than 120. Normal level of total cholesterol is under 200, and hyperlipidemia is diagnosed if it is over 240.

Figure A.3: Probability of General Health Screening Take up by Health Status

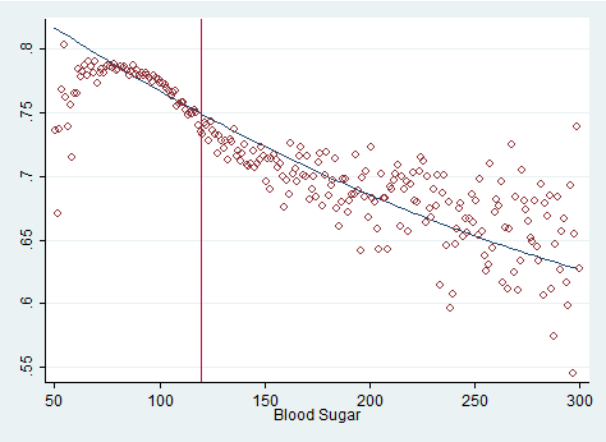
Panel A. BMI, Male



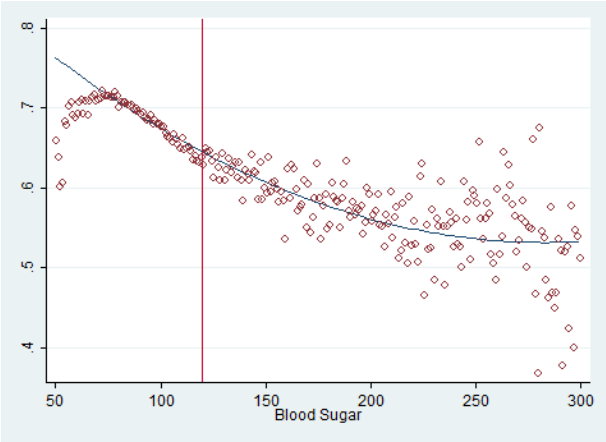
Panel B. BMI, Female



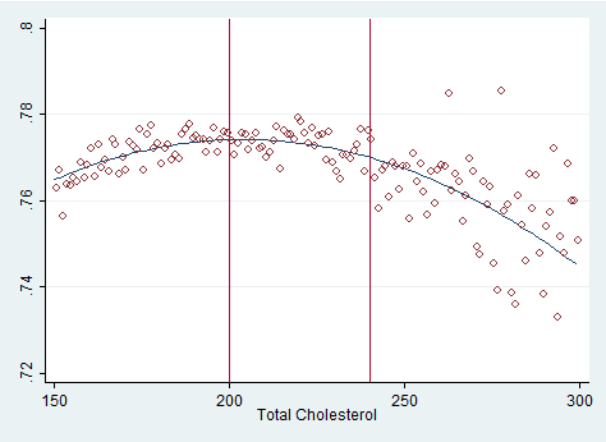
Panel C. Blood sugar, Male



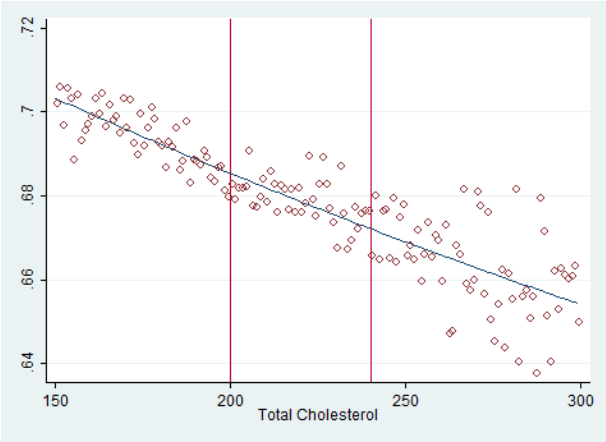
Panel D. Blood sugar, Female



Panel E. Total Cholesterol, Male



Panel F. Total Cholesterol, Female



Note: Each figure shows probability of general health screening take up in the second round by BMI, blood sugar and cholesterol level in the first round. Normal range of BMI is between 18.5 and 25. Normal level of blood sugar is under 110, and DM is diagnosed if it is greater than 120. Normal level of total cholesterol is under 200, and hyperlipidemia is diagnosed if it is over 240.

Appendix B

Chapter 2 Appendix

The preliminary score is calculated from the responses to 52 questions in the eligibility evaluation. A list of these items and possible responses are listed in Table B.1.

The procedure for determining the preliminary score is as follows:

1. Convert responses to point values, according to Table B.1.
2. Sum the points in each category.
3. Based on the category scores and the responses to the 52 items, determine sub-scores for eight service categories: individual hygiene, excretion support, eating, moving, behavior, indirect support, nursing care, and rehabilitation. See Figure B.1 for an illustration of how the eating sub-score is determined.
4. Sum the service sub-scores to arrive at the preliminary score.

We now provide a partial example for calculating the preliminary score. Table B.2 contains a sample set of answers to the eligibility assessment. The category scores for ADL and REH are 16 and 13, respectively. Now follow the eating tree. The first fork depends on the response to "eating" in the ADL category. The response of independent (1) sends us along the left branch. The response to "brushing teeth" is independent (1), which takes us down the first branch. Since the ADL score is 16, we end up along the right branch for a score of 9.4. We repeat this procedure for the remaining service sub-scores:

- Individual hygiene 5.3

- Excretion support 2.6
- Eating 9.4
- Moving 3.6
- Behavior 0.8
- Indirect 21.7
- Nursing Care 9.7
- Rehabilitation 2.7

These sum to the final score of 55.8.

Table B.1: Assessment Questions

Category	**Items**	**Possible Responses (response score)**
Physical Function	Dressing and Undressing, Washing His/Her Face, Brushing His/Her Teeth, Bathing, Eating, Changing Positions, Sitting Up, Moving to Sit Elsewhere, Moving to Another Room, Using the Restroom, Control of Defecation, Control of Urination	Independent (1), Needs Partial Help (2), Needs Complete Help (3)
Cognitive Function	Short Term Memory Disorder, Failing to Perceive Date, Failing to Perceive Location, Failing to Recall Age and Birth Date, Failing to Understand Instructions, Deteriorating Circumstantial Judgement, Communication Disorder	No (0), Yes (1)
Behavior	Hallucination, Delusion, Depression, Sleep Disorder, Resistant to Help, Restless, Gets Lost, Abusive/Aggressive Behavior, Trying to Go Out, Breaks Things, Inappropriate Behavior, Hiding Money/Goods, Dressing Improperly, Lack of Restroom Hygiene	No (0), Yes (1)
Medical Treatment	Tracheotomy, Aspiration, Oxygen Therapy, Bed Sore, Tube Feeding, Cancer Pain Management, Urine Catheter, Fistula Care, Dialysis	No (0), Yes (1)
Rehabilitation	Motor Disturbance (Right/Left Arm/Leg); Joint Disorder (Shoulder, Elbow, Wrist, Hip , Knee, Ankle)	No Disorder/Limitation (1), Partial Disorder/Limitation (2), Complete Disorder/Limitation (3)

Table B.2: Sample Assessment

Category	Item	Response	Score
Activities of Daily Living	Dressing and Undressing	Need Partial Support (NPS)	2
	Washing Face	Fully Self Support (FSS)	1
	Brushing Teeth	Fully Self Support (FSS)	1
	Bathing	Need Full Support (NFS)	3
	Feeding	Fully Self Support (FSS)	1
	Changing Position	Fully Self Support (FSS)	1
	Sitting Up	Fully Self Support (FSS)	1
	Changing Sitting Location	Fully Self Support (FSS)	1
	Ambulation	Need Partial Support (NPS)	2
	Using the Restroom	Fully Self Support (FSS)	1
	Voluntarily Control of Fecal Discharge	Fully Self Support (FSS)	1
	Voluntarily Controlling of Urinary Discharge	Fully Self Support (FSS)	1
			16
Cognitive Function	Short Term Memory Loss	Yes	1
	Disorientation of Date	Yes	1
	Disorientation of Place	Yes	1
	Disorientation of Age and Birth Date	No	0
	Disorientation of Order	Yes	1
	Disorientation of Judgement	Yes	1
	Despair of Communication	No	0
			5
Misbehavior	Illusion	No	0
	Delusion	Yes	1
	Depression	No	0
	Sleep Disorder	Yes	1
	Resistant to Support	No	0
	Being Anxious / Lingering Around	Yes	1
	Being Lost	No	0
	Abusive Language / Aggressive Behavior	No	0
	Trying to Go Out	Yes	1
	Destroys Things	No	0
	Meaningless or Inappropriate Behavior	Yes	1
	Hiding Money / Things	Yes	1
	Inappropriate Clothing	No	0
	Unclean Urination / Defecation	No	0
			6
Nursing	Tracheotomy	No	0
	Aspiration	No	0
	Oxygen Therapy	No	0
	Bed sore	No	0
	Tube Feeding	No	0
	Pain Control of Cancer	No	0
	Care of Urine Catheter	No	0
	Fistula Care	No	0
	Care for Dialysis	No	0
			0
Rehabilitation	Right Arm	No Disorder (ND)	1
	Left Arm	No Disorder (ND)	1
	Right Leg	No Disorder (ND)	1
	Left Leg	No Disorder (ND)	1
	Shoulder	No Limitation	1
	Elbow	No Limitation	1
	Wrist	No Limitation	1
	Hip Joint	Symmetry	3
	Knee Joint	Asymmetry	2
	Ankle	No Limitation	1
			13

Figure B.1: Eating Tree

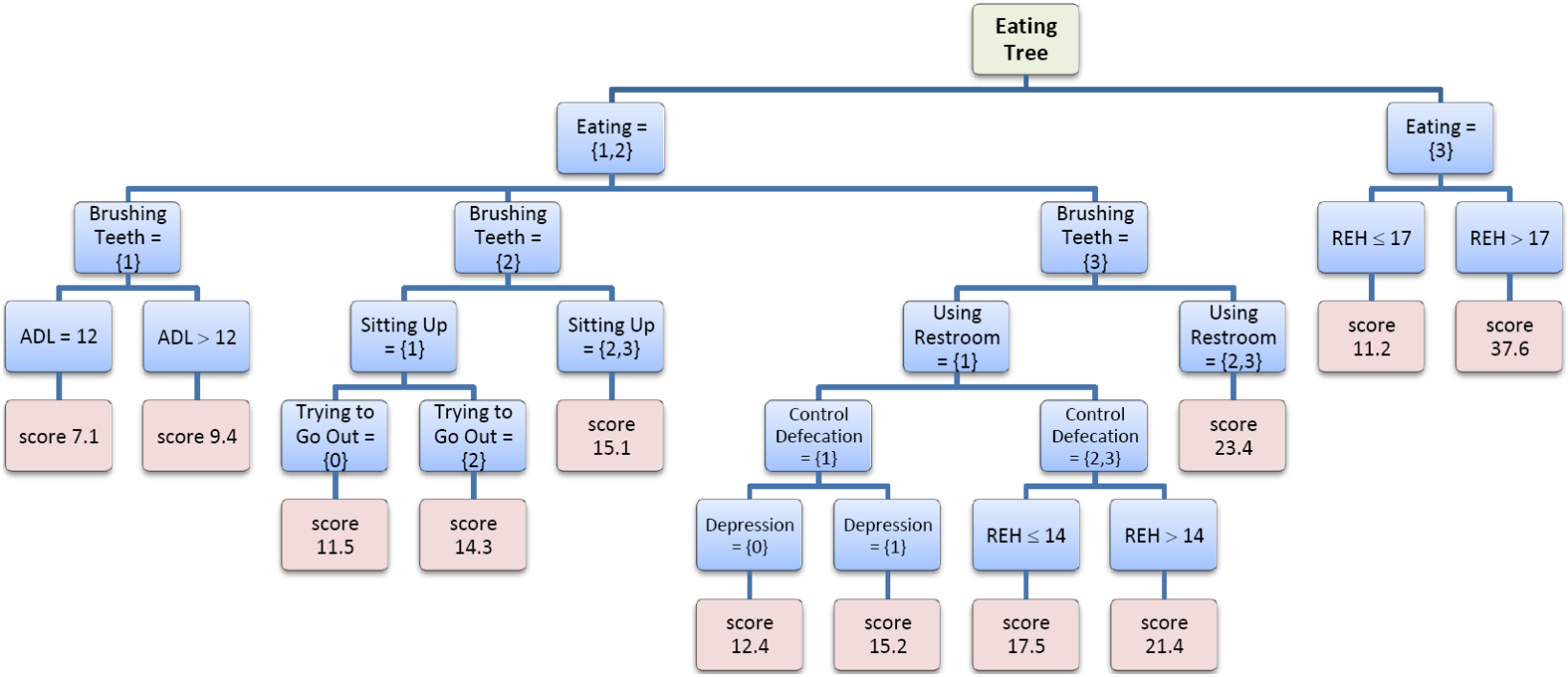
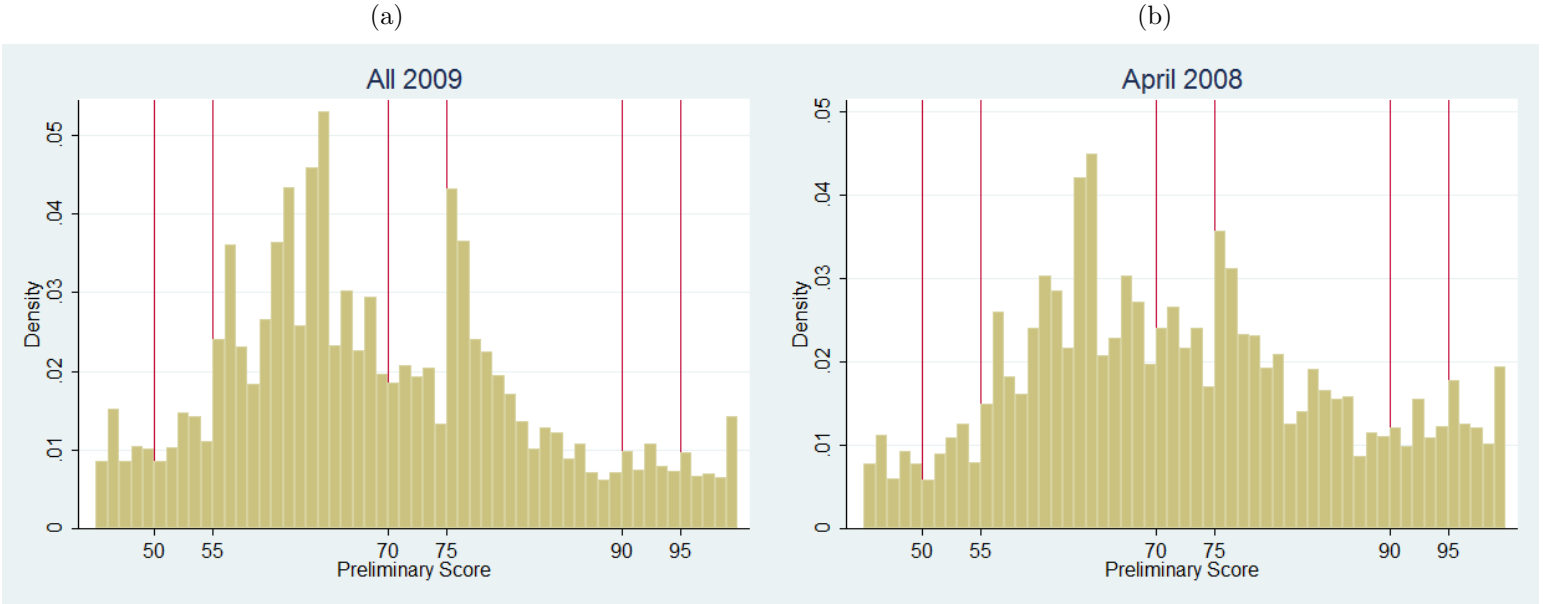


Table B.3: Description of Reimbursed LTC Services

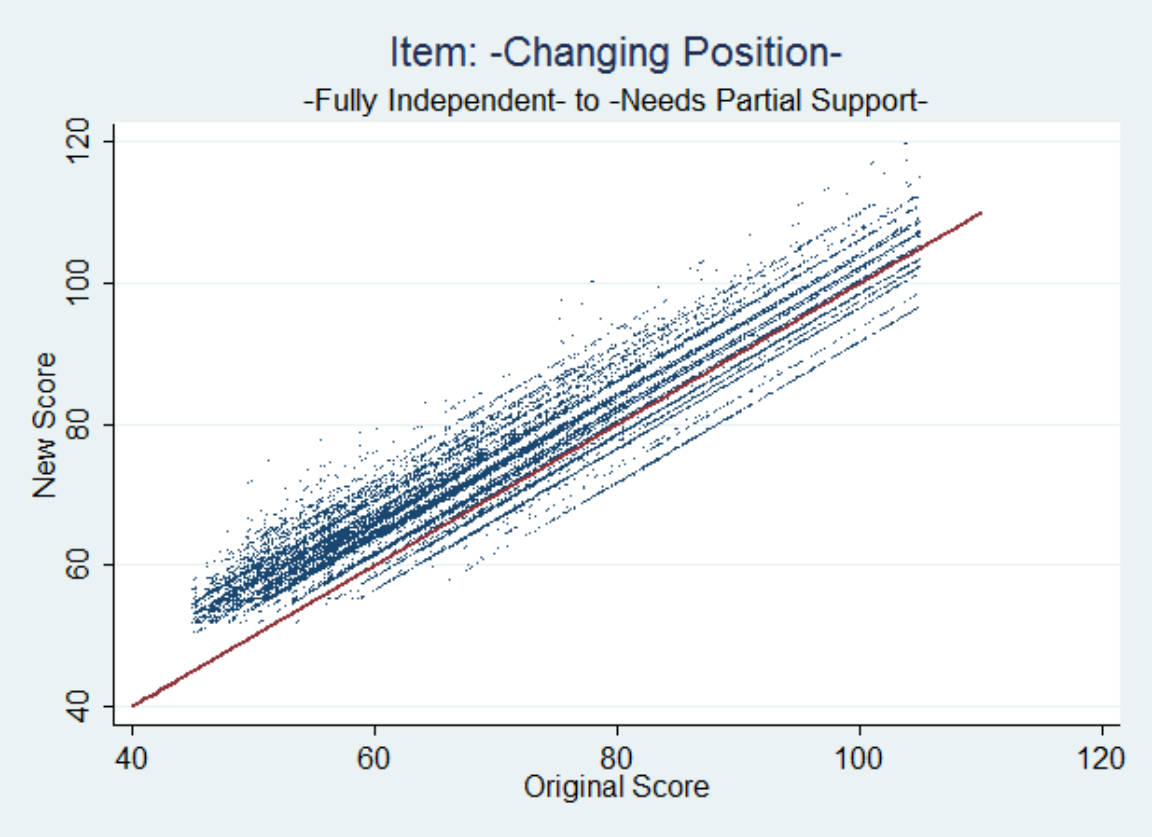
Category	Benefit	Description
Home	Home Help Visit	A care worker visits the beneficiary's residence and helps in the following: bathing, using the restroom, changing clothes, washing hair, cooking, buying daily necessities, cleaning, and clearing up surroundings.
Home	Home Bathing Visit	A care worker visits the beneficiary's residence to provide bathing services using a bathing device.
Home	Home Nursing Visit	A nurse or dental hygienist visits the beneficiary's residence to provide nursing, treatment assistance, care consultation, or dental hygiene services.
Home	Day and Night Care	Facility stay for less than a day where education or training is provided to the beneficiary for maintenance or improvement of physical and mental function.
Home	Short-Term Respite Care	Short term facility care in order to provide temporary relief for the regular caregiver.
Home	Medical Supplies	Equipment is provided for the support of the beneficiary's daily tasks and physical activities (e.g. bath seat or walker).
Institutional	Elder Care Facility or Group Home	Residence, meals, care, and other conveniences requried for daily function.

Figure B.2: Density of Scores, 2009 vs. April 2008



Notes: Preliminary scores in 1 point bins. "All 2009" consists of the preliminary scores from assessments in 2009. "April 2008" consists of the preliminary scores from assessments in April 2008.

Figure B.3: Score Sensitivity Example



Notes: Sample of individuals whose preliminary scores fall between 45 and 105, with the response "Fully Independent" to the item "Changing Position". The original preliminary score is on the x-axis. The new preliminary score after changing the response from "Fully Independent" to "Needs Partial Support" is on the y-axis. Also graphed in red is the 45 degree line.