Three Essays on Health Care

Hitoshi Shigeoka

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2012

ABSTRACT

Three Essays on Health Care

Hitoshi Shigeoka

This dissertation has been motivated by the question of how countries should optimally structure health care. Especially, there are two important economic and policy questions asked that extend beyond the area of health economics. The first is how the expansion of health insurance coverage affects the utilization and health of its beneficiaries (extensive margin); the second is how generous should health insurance be (intensive margin) to balance the provision of care and financial protection against risk while containing medical expenditures. The three chapters in this dissertation aim to make empirical contributions to these ongoing research questions.

First Chapter, "The Effect of Patient Cost-Sharing on Utilization, Health and Risk Protection: Evidence from Japan" addresses the second question. It investigates how cost-sharing, requiring patients to pay a share of the cost of care, affects the demand for care, health itself, and risk protection among the elderly, the largest consumers of health service. Previous studies of cost-sharing have had difficulty separating the effect of cost-sharing on patients from the influence of medical providers and insurers. This paper overcomes that limitation by examining

a sharp reduction in cost-sharing at age 70 in Japan in a regression discontinuity design. I find that price elasticities of demand for both inpatient admissions and outpatient visits among the elderly are comparable to prior estimates for the non-elderly. I also find that the welfare gain from risk protection is relatively small compared to the deadweight loss of program financing, suggesting that the social cost of lower cost-sharing may outweigh social benefit. Taken together, this study shows that an increase in cost-sharing may be achieved without decreasing total welfare.

Third Chapter, "Effects of Universal Health Insurance on Health Care Utilization, Supply-Side Responses and Mortality Rates: Evidence from Japan" (with Ayako Kondo) address the first question. Even though most developed countries have implemented some form of universal public health insurance, most studies on the impact of the health insurance coverage have been limited to specific subpopulations, such as infants and children, the elderly or the poor. We investigate the effects of a massive expansion in health insurance coverage on utilization and health by examining the introduction of universal health insurance in Japan in 1961. We find that health care utilization increases more than would be expected from previous estimates of the elasticities of individual-level changes in health insurance status such as RAND Health Insurance Experiment in the US.

The two chapters addressed above focus on consumers' incentives. Second chapter, "Supply-Induced Demand in Newborn Treatment: Evidence from Japan" (with Kiyohide Fushimi) examines the incentives faced by medical providers. Since

medical providers exert a strong influence over the quantity and types of medical care demanded, measuring the size of supply-induced demand (SID) has been a long-standing controversy in health economics. However, past studies may underestimate the size of SID since it is empirically difficult to isolate SID from other confounding hospital behaviors, such as changes in the selection of patients. We overcome these empirical challenges by focusing on a specific population: at-risk newborns, and we measure the degree of SID by exploiting changes in reimbursement caused by the introduction of the partial prospective payment system (PPS) in Japan, which makes some procedures relatively more profitable than other procedures. We find that hospitals respond to PPS adoption by increasing utilization and increasing their manipulation of infant's reported birth weight, which determines infants reimbursement and maximum length of stay. We also find that this induced demand substantially increases hospital reimbursements without improving infant health, implying that the additional money spent has no commensurate health gains.

Contents

List of Figures

List of Tables

Acknowledgements

During writing this dissertation, I benefited from a number of people. First of all, I would like to thank my main advisor, Douglas Almond, for his guidance at every stage of the my research. Without his continuous encouragement through my entire dissertation, I could not complete the dissertation. Wojciech Kopczuk and Tal Gross gave me invaluable advices and supports during job market.

I am also grateful to helpful comments and suggestions from Prashant Bharadwaj, Kasey Buckles, Janet Currie, Joseph Doyle, Mark Duggan, Amy Finkelstein, Kiyohide Fushimi, Michael Grossman, Hideki Hashimoto, Masako Ii, Amanda Kowalski, Ilyana Kuziemko, Frank Lichtenberg, Jason Lindo, Bentley MacLeod, Shinya Matsuda, Robin McKnight, Matt Neidell, Cristian Pop-Eleches, Heather Royer, Bernard Salanié, Miguel Urquiola, Eric Verhoogen, Till von Wachter, Reed Walker, and the seminar participants at Bank of Japan, Columbia University, McGill University, National University of Singapore, Osaka University, Simon Frazer University, University of Michigan, Uppsala University, and NBER Japan project meeting.

Special thanks go to Hideo Yasunaga and Hiromasa Horiguchi for their invaluable help in obtaining the data and for helpful discussions.

I would also like to thank my friends and colleagues at Columbia for their helpful discussions and for making my time in graduate school memorable. At the risk of forgetting some names, I thank Bruno Giovannetti, Mariesa Herrmann, Takakazu Honryo, Ayako Kondo, Marcos Yamada Nakaguma, Yoichi Sugita, Kensuke Teshima, and Zhanna Zhanabekova.

Finally, my family have supported the entire period of my graduate study in New York. I dedicate this dissertation to my family members in token of affection and gratitude.

CHAPTER 1

The Effect of Patient Cost-sharing on Utilization, Health and Risk Protection: Evidence from Japan

1.1. Introduction

Governments increasingly face an acute fiscal challenge of rising medical expenditures especially due to aging population and expansion of coverage. Spending growth for Medicare, the public health insurance program for the elderly in the United States, has continued unchecked in spite of a variety of government attempts to control costs.[1] As more than one third of current health spending is on the elderly, future cost control efforts can be expected to focus on seniors.[2]

One main strategy for the government to contain cost is cost-sharing, requiring patients to pay a share of the cost of care. However, cost-sharing has clear tradeoffs. While cost-sharing may reduce direct costs by decreasing moral hazard of health

[1]Examples of supply-side attempts by the government to control cost are the introduction of prospective payment for hospitals and reductions in provider reimbursement rates (Cutler, 1998).

[2]The elderly are the most intensive consumers of health care. Patient over age 65 consume 36 percent of health care in the US despite representing only 13 percent of the population (Centers for Medicaid and Medicare Services 2005). Furthermore, Medicare costs are expected to comprise over a quarter of the primary federal budget by 2035, or between five and six percent of GDP (CBO, 2011). Likewise, in Japan, the elderly consume five times as many health services as non-elderly (Okamura et al, 2005). Also Japan has the most rapidly aging population in the world (Anderson and Hussey, 2000).

care services, it may also reduce access to beneficial and necessary health care that could mitigate future severe and costly health events. Moreover, very high levels of cost-sharing may undermine one of the primary reasons of having health insurance, which is financial protection from catastrophic health events. Thus, there is a desperate need for knowledge on how cost-sharing affects utilization, health itself and risk protection, especially among the elderly, to determine the appropriate level of cost-sharing.

Credible evidence on the price sensitivity of health care consumption among the elderly is limited. For instance, individuals above age 62 were excluded from the well-known RAND Health Insurance Experiment (hereafter, RAND HIE), which randomly assigned individuals to insurance plans with different generosities. It is not clear *a priori* whether the elderly are expected to have a larger or smaller price elasticity of demand for health care services than the non-elderly. On one hand, the price elasticity for the elderly may be larger if they tend to be poorer or more credit-constrained than the non-elderly. On the other hand, it can be smaller if their health problems are more severe than those of non-elderly. An exception that studied the elderly is Chandra et al. (2010) who examined the effect of a small increase in the copayments for physician office visits and prescription drugs in a supplemental Medicare insurance policy.

Most U.S. studies, however, have difficulty separating the demand elasticities of patients from the responsive behavior by insurers and medical providers. This limitation arises because insurers prevent patients from freely choosing medical providers through managed-care, and medical providers determine which treatments to provide based on the patients' health insurance plans. Indeed, there is substantial evidence that the medical providers are reluctant to treat patients with government-funded health insurance beneficiaries due to low reimbursement rates as well as frequent delays in reimbursement.[3] If insurers and medical providers limit the patients' demand for health care services, the elasticities of demand that are estimated in these studies could be underestimated.

By contrast, the unique setting in Japan permits isolation of the demand elasticity for health care services since medical providers and insurers typically play a small, if any, role in patients' demand for health care services. Under universal health insurance coverage in Japan, there are no restrictions on patients' choices of medical providers. Also physicians' payments are based on a national fee schedule that does *not* depend on patients' insurance type. This institutional setting limits physicians' incentives to influence patient demand and prevents cost-shifting, a well-known phenomenon in the U.S. where medical providers charge private insurers higher prices to offset losses from the beneficiaries of government-funded health insurance (Cutler, 1998).

My research design exploits a sharp reduction in patient cost-sharing at age 70 in Japan in a regression discontinuity design to compare the outcomes of those just below versus those just over age 70. Due to national policy, cost-sharing for

[3]For example, see Cunningham and O'Malley (2009) and Garthwaite (2011).

outpatient visits and inpatient admissions is as much as 60-80 percent lower at age 70 than at age 69 in Japan. This reduction is substantial, especially for inpatient admissions: out-of-pocket medical expenditures for inpatient admissions can reach as much as 25 percent of the average annual income of a 69-year-old patient among those admitted. Since turning 70 in Japan does not coincide with changes in any other confounding factors such as employment or pension receipt, I can plausibly isolate the effect of the cost-sharing on demand for health care services.

This setting also offers additional advantages over previous empirical settings. While the change in co-payment in Chandra et al. (2010) is limited to office visits and prescription drugs, in Japan cost-sharing for inpatient admissions also changes abruptly at age 70. Thus I can estimate the elasticity of inpatient admissions of the elderly as well. Also, since I have detailed information on outpatient visits, I can investigate the price sensitivity of preventive care in the outpatient setting.[4] In contrast, most existing datasets capture either outpatient visits or inpatient admissions.[5] Finally, I examine the effect of cost-sharing on exposure to out-of-pocket medical expenditure risk. While there is a large literature on the impact of cost-sharing on health care utilization and health, there is remarkably little

[4]Outpatient visits are visits to a clinic or hospital without being admitted. It is common for individuals to visit hospitals for outpatient care rather than clinics (similar to physicians' office visits in the U.S.) in Japan.

[5]In fact, the Agency for Healthcare Research and Quality (AHRQ) has recognized the need to develop a methodology for studying preventive care in an outpatient setting by using inpatient data to identify admissions that should not occur in the presence of sufficient preventive care (AHRQ, 2011). This issue is more discussed in section 4.

analysis of the impact of cost-sharing on expenditure risk, which is arguably the primary purpose of health insurance (e.g., Zeckhauser, 1970).[6]

I reach three conclusions. First, I find that reduced cost-sharing at age 70 discontinuously increases health care consumption. The corresponding elasticity is modest, around -0.2 for both outpatient visits and inpatient admissions. As it turns out, the elasticity I estimate is similar to the estimates found in the HIE for the non-elderly, and slightly larger than that estimates for the elderly by Chandra et al. (2010). The finding indicates that the price elasticity of the elderly is similar in magnitude to that of the non-elderly.

Second, looking in more detail at patterns of utilization, I find that lower cost-sharing is associated with increase in the number of patients presenting with both serious and non-serious diagnoses. Thus, I find that demand for both more and less beneficial care is price sensitive. For example, I find large increases in outpatient visits for diagnoses that are defined as Ambulatory Care Sensitive Conditions (ACSCs), for which proper and early treatment reduce subsequent avoidable admissions.

Finally, on the benefit side, I do not find statistically significant improvements in health at age 70. Both mortality, and self reported physical and mental health are unchanged despite utilization changes, implying that patient cost-sharing can reduce health care utilization without adversely affecting health. But I find that

[6]See Chandra et al. (2008) and Swartz (2010) for an excellent summary of the past literature on cost-sharing and utilization.

lower cost-sharing at age 70 yield reductions in out-of-pocket expenditures since lower cost-sharing overwhelms the increase in utilization. I then compute the gain in risk premiums through increased generosity in health insurance at age 70 by combining the expected utility framework with the quantile RD estimates. Although somewhat speculative, my estimates suggest that the welfare gain of risk protection from lower cost-sharing is small for most, suggesting that the social cost of lower cost-sharing may outweigh the social benefit. Taken together, this study shows that increased cost-sharing may be achieved without decreasing the total welfare.

This paper is related to an influential literature that examines Medicare eligibility at age 65 in a similar RD framework as this paper. Card et al. (2009) and Chay et al. (2010) show that Medicare eligibility has a modest positive effect on the health of those above age 65. However, these studies cannot definitely address whether these health improvements are the result of health insurance provision *per se* (extensive margin) or changes in health insurance generosity (intensive margin). This issue arises because turning age 65 in the US entails a number of coincident changes: transitions from private to public health insurance, increases in multiple coverage due to supplementary coverage (e.g., Medigap), and fewer gatekeeper restrictions due to the change from managed care to fee-for-services. Indeed, Card et al. (2009) conclude that it is not clear whether reductions in mortality are due to health insurance provision or generosity.[7] In contrast, the change at age 70 only reflects increases in benefit generosity in my case.

The rest of the paper is organized as follows. Section 1.2 briefly describes the institutional background. Section 1.3 describes the data, and presents the identification strategy. Section 1.4 shows the main results on utilization. Section 1.5 turns to the analysis on benefit, and examines the health outcomes as well as risk reduction. Section 1.6 carries out simple cost-benefit analysis and section 1.7 concludes.

1.2. Background

This section describes the universal health insurance system in Japan, focusing on the differences in cost-sharing between the elderly and non-elderly.[8]

1.2.1. Institutional Setting

Japan's universal health insurance system consists of two parallel subsystems: employment-based health insurance and National Health Insurance (hereafter,

[7] In a companion paper, Card et al. (2008) also find that both supply-side incentives and shifts in insurance characteristics play an important role for the utilization of health care services.

[8] Japan achieved universal health insurance coverage in 1961. See Kondo and Shigeoka (2011) for more details about the effect of the introduction of universal health insurance on utilization and health.

NHI). Employment-based health insurance covers the employees of firms that satisfy certain requirements and employees' dependents.[9] NHI is a residential-based system that provides coverage to everyone else, including the employees of small firms, self-employed workers, the unemployed, and the retired.

For this study, there are two important features of Japanese medical system that arguably permits isolation of the patient demand for health care services from responsive behavior by insurers and medical providers: universal coverage and the uniform national fee schedule. First, under universal coverage, patients in Japan have unrestricted choices of medical providers unlike in the U.S where managed-care often restricts the set of the providers at which beneficiaries can receive treatment. For example, it is common for individuals to visit hospitals for outpatient care rather than clinics (similar to physicians' office visits in the U.S.) in Japan. Patients have direct access to specialist care without going through a gatekeeper or referral system. There is also no limit on the number of visits a patient can have. Patients may go either hospitals or clinics for outpatient visits and go to hospitals for admissions, unlike in the U.S., where those who lack insurance use hospitals as primary care.

[9] Employment-based health insurance is further divided into two forms; employees of large firms and government employees are covered by union-based health insurance, whereas employees of small firms are covered by government-administered health insurance. Enrollment in the government-administered health insurance program is legally required for all employers with five or more employees unless the employer has its own union-based health insurance program.

Second and perhaps more importantly, all medical providers are reimbursed by the national fee schedule, which is uniformly applied to all patients regardless of patients' insurance type and age. Since patients' insurance type and age do not affect reimbursements, physicians have few incentives to influence patients' demand.[10] For example, from physicians' perspective, there are few reasons to delay surgeries until age 70 because reimbursements do not differ by age of patients. The uniform fee schedule also implies that there is little room for cost-shifting, a well-known behavior of medical providers in the U.S. where they charge private insurers higher prices to compensate for losses from beneficiaries of public health insurance (Cutler, 1998).[11]

As a result, while people in Japan enjoy the relatively easy access to health care services, Japan has the highest per-capita number of physician visits among all OECD countries; physician consultations (number per capita per year) is 13.2 in Japan, which is more than three times larger than 3.9 in the U.S. (OECD, 2011). While some blame universal coverage for high frequency of unnecessary physician visits, others claim that these medical services contribute to the longevity of the Japanese (Hashimoto et al., 2011).

[10]The national schedule is usually revised biennially by the Ministry of Health, Labor and Welfare through negotiation with the Central Social Insurance Medical Council, which includes representatives of the public, payers, and providers. See Ikegami (1991) and Ikegami and Campbell (1995) on details.

[11]Japan introduced prospective payment for hospitals since 2003 for only acute diseases, but the reimbursement does not differ by the insurance type or age of the patients. See Shigeoka and Fushimi (2011).

1.2.2. Changes in Cost-sharing at Age 70

Unlike a normal health insurance plan that has three basic components (a deductible, a coinsurance rate, and a stop-loss), there is no deductible in Japan.[12] A patient pays coinsurance which is the percentage of medical costs for which beneficiary is responsible.[13] Since inpatient admissions are more expensive than outpatient visits, coinsurance rate of inpatient admissions tends to be set lower than that of outpatient visits in Japan. The insurer pays the remaining fraction of expenses until the beneficiary meets the stop-loss (also known as the maximum out-of-pocket), and the insurer pays all expenses above the stop-loss.

The Japanese government passed the Act on Assurance of Medical Care for Elderly People, which imposed cost-sharing on those over 70 starting in February 1983 after the 10 years of generous policy that provided free care for the elderly over age 70.[14] Even after its introduction, there has been still a large discrepancy in cost-sharing between those just above and below age 70 as described in detail below.

[12]A deductible is lump-sum amount of spending that beneficiary must pay before the insurers cover any expenses.

[13]Typically coinsurance is applied for medical costs above the deductible in the US.

[14]Japan introduced free care for the elderly in January 1973. However, this policy substantially increased the utilization of health care services and medical expenditures. In fact, the medical expenditures rose by 55 percent in just one year, from 429 billion Yen in 1973 to 665 billion Yen in 1974. Due to data availability, this study focuses on the period after the implementation of the cost-sharing for the elderly.

The elderly become eligible for lower cost-sharing on the first day of the next month after they turn 70. They receive a notice from the government that indicates that they are eligible for Elderly Health Insurance and a new insurance card, which they can present at medical institutions to receive the discount. Elderly Health Insurance is also provided to bedridden people between the ages of 65 and 70. Figure 1.1 shows the age profile of health insurance coverage from the pooled Patient Surveys described later in the data section. Age is aggregated into months. The percent of patients with Elderly Health Insurance abruptly rises from 20 percent to nearly 100 percent once they turn 70. I also see a small jump in Elderly Health Insurance coverage at age 65.

Table 1.2 displays the cost-sharing formulas for those below and above age 70 for outpatient visits and inpatient admissions separately for each survey year of the Patient Survey. For those below age 70, the coinsurance rate is determined by the type of health insurance (employment-based health insurance or NHI), employment status (retired or not), and whether the person is a (former) employee or is a dependent. Employment-based health insurance had a lower coinsurance rate than NHI until 2003, when both were equalized to a common coinsurance rate of 30 percent for both outpatient visits and inpatient admissions. At the age of 70, people switch to Elderly Health Insurance and in principle face the same cost-sharing.[15] Note that on the other hand, physicians' reimbursements are based on a national fee schedule that does not depend on patients' insurance type or age.

Figure 1.2 illustrates the amount of out-of-pocket expenditures with respect to total monthly medical expenditures for year 2008 as an example based on the formula in Table 1.2. Unlike in the US, in Japan, the stop-loss is set monthly rather than annually.[16] The horizontal axis is total monthly medical expenditures, and the vertical axis shows the corresponding monthly out-of-pocket medical expenditures. Since the stop-loss differs for outpatient visits and inpatient admissions for those over age 70, I show separate lines for outpatient visits and inpatient admissions. For those below 70, there is no distinction between these two services in 2008. Figure 1.2 shows that the price schedule of out-of-pocket medical expenditures for those above 70 always lies below that of those below age 70.

Unfortunately, the actual out-of-pocket expenditure information among the general population is only available for year 2007, and this data does not distinguish outpatient visits and inpatient admissions. However, I have individual level insurance claim data for outpatient visits and inpatient admissions respectively,

[15]In fact, high income earners above age 70 are charged higher coinsurance rate (20 percent instead of 10 percent) since October 2002. The bar for high income level is set quite high, so that a limited number of patients is in this category (7 percent according to Ikegami et al. 2011). Since income is not collected in the Survey of Medical Care Activities in Public Health Insurance, which I use to derive the monthly out-of-pocket expenditures, I compute the monthly out-of-pocket expenditures for a normal family. See Appendix A.1 for detail.

[16]This is purely administrative reason; reimbursements to the medical institutions are conventionally paid monthly in Japan and thus stop-loss is set monthly.

which is the monthly summary of medical expenditures claimed for insurance reimbursement to medical institutions (called the Survey of Medical Care Activities in Public Health Insurance). Since a portion of this monthly total medical expenditure is paid as patient cost-sharing according to the formula in Table 1.2, I can compute the average out-of-pocket medical expenditures at each age for each survey year of the Patient Survey.

Table 1.3 summarizes the actual monthly out-of-pocket expenditures of the average 69-year-old, and the counterfactual monthly out-of-pocket medical expenditures for a 70-year-old. For those age 70-year-old, since out-of-pocket medical expenditures are endogenous (i.e., observed out-of-pocket medical expenditure already reflects the change in cost-sharing), I compute their counterfactual out-of-pocket expenditures by applying the cost-sharing rules of Elderly Health Insurance to the utilization of the average 69-year-old. See Appendix A1 for details on these derivations. Note here that I do not exploit the year-to-year variation in cost-sharing in this paper, and rather pool all the survey rounds to increase the statistical power and to smooth out cohort-size effect.[17] The overall out-of-pocket medical expenditure conditional on using medical institutions in Table 1.3 is the weighted average of the out-of-pocket medical expenditure across all survey years, using the population of 69-year-old in each survey year as weights.

[17]Due to the smaller sample size, the estimates from separate years are noisier and do not have any consistent pattern. Also I need to view these results with caution since cohort-size may affect the estimates in this RD framework since I use counts rather than rate in most of the specifications. These results are available from the author.

Table 1.3 reveals a couple of interesting facts. First, out-of-pocket medical expenditures, especially from inpatient admissions, can pose a substantial financial burden on the near elderly (those just below age 70). Since the average annual income for 69-year-old is 1,822 thousand Yen (or roughly 18,220 US dollars), out-of-pocket medical expenditures for inpatient admissions can reach as much as 25 percent of an average person's total annual income for those admitted.[18] On the other hand, once the patient turns 70, the counterfactual ratio of medical expenditures to the average income is reduced to as small as 8.2 percent.[19]

It is also important to note that stop-loss plays a role in reducing the out-of-pocket medical expenditures for those *below* 70, especially for inpatient admissions. In the absence of stop-loss, the gap between above and below 70 would be even larger. Since coinsurance rate is much higher for those below age 70 than those over 70 (30 percent vs. 10 percent), the stop loss kicks in at a much lower total amount, which is jointly paid by the patient and the insurers, for those below 70 (267 thousands Yen) than those above 70 (444 thousand Yen = 44.4/0.1). Indeed, column (4) in Table 1.3 shows that while only 0.1 percent of outpatient visit claims for 69-year-old reach the stop-loss, 14.6 percent of inpatient admissions reach the

[18] One thousands Yen is roughly $10 US dollars. Author's calculation from the Comprehensive Survey of Living Conditions (38.0*12)/1,822 = .25

[19] Author's calculation from the Comprehensive Survey of Living Conditions (12.4*12)/1,822 = .082

stop-loss conditional on the use of the medical institutions. Interestingly, no 70-year-old patients reach the stop-loss for inpatient admissions in my data, since their coinsurance rate is set particularly low, as seen in column (5) in Table 1.3. I explore the effect of cost-sharing on out-of-pocket medical expenditures in detail in Section 1.5.

1.3. Data and Identification

I use one of the most comprehensive sources of health-related datasets ever assembled on Japan. Here I summarize the most important datasets in the study; further details can be found in the Appendix A.3. My main outcomes are health care utilization on the cost-side, and health outcomes, and out-of-pocket expenditures on the benefit-side.

1.3.1. Data

The dataset for health care utilization is the Patient Survey, a nationally representative repeated cross-section that collects administrative data from both hospitals and clinics.[20] Since the survey is conducted every three years, I have individual patient level data for nine rounds of surveys between 1984 and 2008. One of the biggest advantages of this survey relative to usual hospital discharge data is that the Patient Survey includes information for outpatient visits as well. In contrast, most existing datasets capture either outpatient visits or inpatient admissions. In

[20]See Bhattacharya et al. (1996) for an example of a study that uses the Patient Survey.

fact, the Agency for Healthcare Research and Quality (AHRQ) has recognized the need to develop a methodology for studying preventive care in an outpatient setting by using inpatient data to identify admissions that should not occur in the presence of sufficient preventive care (AHRQ, 2011).[21] In my case, I can look at changes in the number of patients for beneficial and preventive care in the outpatient setting.[22] The disadvantage of this data is that, as in the case for most of the discharge data, it only includes limited individual demographics such as gender, and place of living (no education or income).

The Patient Survey consists of two types of data: outpatient data and discharge data. I use the former to examine outpatient visits and the latter for inpatient admissions. The outpatient data is collected during one day in the middle of October of the survey year and provides information on all patients who had outpatient

[21]The interaction between outpatient visits and inpatient admissions may be crucial since Chandra et al. (2010) find evidence of offset effects; copayment increases reduce outpatient visits but increase subsequent hospitalizations. Offset effects are not observed in the RAND HIE. I cannot really answer whether I see the offset effects because coinsurance rate for both outpatient visits and inpatient admissions change at age 70, making it harder to examine the interaction of two services.

[22]Another advantage of the Patient Survey, which is unique to Japan's medical system, is that it has information on patients in both hospitals and clinics. In Japan, hospitals are defined as medical institutions with 20 or more beds, and clinics are defined as medical institutions with no more than 19 beds. Unlike in the U.S., direct outpatient visits to hospitals are common practice in Japan since there are no restrictions on the patients' choice of medical providers. Therefore, the government aims at having clinics provide primary care and hospitals serve more serious cases to increase the total efficiency of the health care system. However, the reduction in cost-sharing at age 70 may increase the flow of outpatient visits to hospitals for non-serious reasons. This possibility is investigated briefly in section 1.4.1.

visits to the surveyed hospitals and clinics during the survey day.[23] This data includes patients' exact date of birth and the survey date, which is equivalent to the exact date of the visits. The discharge data contain the records of all patients who were discharged from surveyed hospitals and clinics in September of the survey year. The discharge data report the exact dates of birth, admission, surgery, and discharge, which enable me to compute age at admission.[24] Hospital and clinic information are obtained from the Survey of Medical Institutions and merged with Patient Survey.

As health outcomes, I examine both mortality and morbidity. I examine mortality since it is one of the few objective, well-measured health outcomes and is also often easily available, and comparable across different countries. I use the universe of death records between 1987-1991, which report the exact dates of birth, death, place of death, and cause of death using International Classification of Diseases (ICD) Ninth. The main advantage of the death records is that they cover all deaths that occur in Japan, unlike hospital discharge records, which only report deaths that occur in the hospital.[25] I complement the mortality results by examining

[23]Since outpatient visits are collected on only one day, the survey is susceptible to external factors such as weather. Therefore it is important to include the survey year fixed effects in the specification to account for this common shock within years. This short survey period is another reason why I do not exploit the year-to-year variation in cost-sharing in this paper.

[24]I describe these dates in chorological order for simplicity, but each unit of data is per discharge.

[25]A rare exception is hospital discharge records in California used in Card et al. (2008, 2009) that tracks mortality within one year of discharge. To my knowledge, data that tracks post-discharge mortality does not exist in Japan.

other morbidity related measures in the Comprehensive Survey of Living Conditions (CSLC), which is survey of a stratified random sample of Japanese population conducted every three years between 1986 and 2007. The survey asks questions about insurance coverage, self-reported physical and mental health, stress levels, and so forth. Age is reported in month in this dataset. Descriptive statistics for Patient Survey (outpatient data and discharge data respectively) and CSLC are reported in the Table 1.1.

1.3.2. Identification Strategy

My identification strategy is very similar to studies from the U.S. that use a regression discontinuity design to examine the effect of turning 65 (Card et al. 2004, 2008, 2009; Chay et al. 2011). However, in Japan, the change at age 70 only reflects increases in benefit generosity rather than combined effect of receiving health insurance coverage and change in benefit generosity, and turning age 70 in Japan does not coincide with changes in any other confounding factors such as employment or pension receipt as shown later.

Even though the idea behind the identification strategy is the same, for clarity, I write two regression equations, one for the CSLC and the other for the Patient Survey and mortality data. The difference comes from the nature of the datasets; while I see all the individuals in former dataset, I only observe those who are present in medical institutions or deceased in the latter two datasets.

My basic estimation equation for CSLC is a standard RD model as follows:

$$Y_{iat} = f(a) + Post70_{iat}\beta + \gamma X_{iat} + \varepsilon_{iat} \tag{1.1}$$

where Y_{iat} is a measure of morbidity or out-of-pocket medical expenditure for individual i at age a in survey year t, $f(a)$ is a smooth function of age, X_{iat} is a set of individual covariates, and ε_{iat} is an unobserved error component. $Post70_{iat}$ is a dummy that takes on the value of one if individual i is over age 70. My parameter of interest is the coefficient β. All coefficients on Post70 and their standard errors have been multiplied by 100 unless otherwise specified, so they can be interpreted as percentage changes. Other controls include a set of dummies for gender, marital status, region, birth month, and survey year. I use a quadratic in age fully interacted with the post dummies as a baseline specification, and run several robustness checks by limiting the sample to narrower age window (ages 67-73), and adding cubic terms in age. To account for common characteristics within the same age cells, the standard errors are clustered at the age in month, following Lee and Card (2008).

Unlike the CSLC, a unique feature of the Patient Survey and mortality data is that I only observe those who are present in the medical institutions or deceased. My approach to deal with this issue is to assume that the underlying population at risk for outpatient visits, inpatient admissions and deaths trends smoothly with age. Card et al. (2004) formally show that under the assumption that the underlying population counts varies smoothly, the estimated discontinuities in log admission counts can be attributed to a corresponding discontinuity in the log of the probability of admission.[26] Since I pool several years of surveys, this assumption seems plausible.[27] Therefore, I use the log of counts as the dependent variable for these datasets and modify the regression equation as follows:

$$\log(Y_{at}) = f(a) + Post70_{at}\ \beta + \mu_{at} \tag{1.2}$$

where Y_{at} is counts of patients or deaths at age a in year t.[28]

Equation (1.2) implies that this RD framework is conceptually different from the typical RD design which relies on assumptions of imprecise control over the running variable (i.e., age in this case), and hence the smoothness of the density of the running variable to identify treatment effects (Lee, 2008). Here, it is precisely

[26] I follow the notation in Card et al. (2004) here. Let p_{ia}, the probability that an individual i of age a is admitted to the hospital in a given time interval, to be written as $log(p_{ia}) = g(a)+Post70_a\pi+v_{ia}$ where $g(a)$ is a smooth age function, $Post70_a$ is a dummy for age 70 or older, and v_{ia} is an error component. Let N_a represent the population of age a and let A_a represent the number who are admitted to hospital, so the ratio A_a/N_a is an estimate of p_a. Finally, assume that the log of the population around age 70 follows a smooth trend: $log(N_a) = h(a)$. Two equations combined implies that the log of the number of hospital admissions at age a is given by $log(A_a) = [g(a) + h(a)] + Post70_a\pi + v_{ia} + \varepsilon_a$ where $\varepsilon_a = log(A_a/N_a) - log(p_a)$.

[27] Note that I am using 9 rounds of Patient Survey. Thus, the people in a given age group in my samples are actually drawn from 9 different age cohorts, smoothing any differences in cohort size.

[28] See Carpenter and Dobkins (2010), and Card et al. (2009) which use the log counts of deaths as outcomes in a similar RD design.

the discontinuity in the density of age at age 70 that I am attributing to an effect of lower cost-sharing on utilization (see e.g., Lee and McCrary 2009; Card et al. 2004, 2008).

There is one remaining empirical issue in estimating equation (1.2) using the Patient Survey. As seen in Figure 1.3, there is substantial seasonality and heaping in the reported birthdays of patients observed in the Patient Survey. First, heaping on the first day of the month is observed, which is likely due to reporting.[29] Second, there are many more births in the first quarter than in the other three quarters throughout the sample period. Some argue that this observation is due to farmers timing births for winter, when there is less work, but the evidence on this observation is little (Kawaguchi 2011).

Whatever the reason, heaping and seasonality in birthdays pose a challenge for estimating equation (1.2) since the Patient Survey is only conducted in one day in October for outpatient visits, and one month in September for inpatients admissions.[30] To account for heaping within the month, I collapse the data into age in months. Since people become eligible for Elderly Health Insurance at the beginning of the next month after their 70th birthday, this approach allows me to code age in months and the post age-70 dummy using the dates of birth and

[29]For example, individuals (or their designated respondents) who do not know their exact birthday may report the first day of their birth month. Other heaps occur at multiples of five and ten days and at the end of the month.

[30]If the data covers the entire year, seasonality is more likely to be smoothed out.

dates of visits without error.[31] To account for the seasonality in birth distribution, I include the birth month fixed effects in addition to survey year fixed effects in all specifications (see e.g., Barreca et al. 2010; Carneiro et al. 2010). Thus the cell is the birth month for each age for each survey year. There are 120 observations (12 month of birth months for each year times 10 years of age 65-75 windows) per survey round, and there are 9 rounds of surveys, and thus there are 1,080 cells in the estimation for outpatient visits.

I also tried two different approaches to account for heaping and seasonality. One approach is to collapse the data into age in quarters, and convert the counts into rates, since I have population data by quarter of birth from the population censuses, which are conducted every five years. The disadvantage of this approach is that the interpolation of population may introduce additional noise in the estimates. In fact, the estimates from this approach tend to be smaller than the main approach due probably to measurement error in the population estimates. Another approach is to include 365 day-of-birth fixed effects as well as year-of-birth fixed effects into the equation in (1.2) to account for the seasonality and cohort-size effects, and use age in days at the time of outpatients visit or inpatient admissions as the running variables (Gans and Leigh 2009; Barreca et al. 2010). The disadvantage of this approach is that when I divide the sample into finer subsamples (e.g., by diagnoses), there are many birthdays without any observations, which may cause

[31] I assign a person who reaches his 70th birthday in October of the survey year to the age 69 and 11 months.

noise in the running variable. The approach of using age in months does not suffer from this problem much since I usually observe at least one observation in each month cell. The results using this alternative approach yield similar results as the main approach as long as there are not many "zero" cells in the data. Since both alternative approaches face different disadvantages, I prefer to take an approach I first described. Some of the results using age in days as running variable are shown in an Appendix Table A.3.[32]

The discharge data pose a slightly more complicated problem. Unlike the outpatient data, the admission day can be any day of the year, as long as they are discharged in September. To avoid including patients with unusually long hospitals stays, I limit the sample to those admitted within three months from discharge in September (July, August, and September) in the survey year. This approach is reasonable since 90 percent of admissions in my data are concentrated within these three months. Since there are 1,080 cells for each admission month, there are a total of 3,240 observations in the estimation of inpatient admissions.[33] In the result sections, I show that the estimates are robust to using different windows from the discharge date.

[32]Other results from different approaches to handle the heaping and seasonality in the dataset (not shown in this paper) are available from author.

[33]The cell for discharge is the month of birth, month of admission, and survey year. The estimation include the birth month fixed effects, admission month fixed effects as well as survey year fixed effects.

For the mortality data, I estimate the same equation as (1.2), replacing Y_{at} as death counts, and using age in days as the running variable.[34] I suffer less from the seasonality of birth issues when using annual mortality census data, since deaths occur throughout the year, and pooling many years of data smoothes out the cohort size.[35] The main drawback of using death records is that I only observe exact date of the death, not exact date of admission as in the hospital discharge data. Note that this may attenuate the estimates since people who died immediately after their 70th birthday may not be eligible for Elderly Health Insurance at the time of admission even though I consider them as treated.

Importantly, the age RD design is distinct from the standard RD design because the assignment to treatment is essentially inevitable, i.e. all individuals will eventually age into the program.[36] As Lee and Lemieux (2010) point out, there are two issues specific to the age RD design. One is that, because treatment is inevitable, individuals may fully anticipate the change in the regime and, therefore,

[34]In fact, since people become eligible for Elderly Health Insurance at the beginning of the next month in which their 70th birthday falls, I use the distance in days from exact day of death to the day of being eligible for Elderly Health Insurance as running variable in mortality analysis.

[35]Interestingly, I observe the same pattern in the mortality data as in the Patient Survey that births are more concentrated in the first quarter of birth, and also on the first day of the month.

[36]Age RD settings are prevalent everywhere. Examples of age RD settings in the United States are eligibility for the Medicare program at age 65 (Card et al., 2008, 2009), young adults aging out of their parents' insurance plans at age 19 (Anderson et al., 2010), legal drinking age at age 21 (Carpenter and Dobkin, 2009), being subject to more punitive juvenile justice system at the age of majority (Lee and McCrary, 2009). There are also many age RD settings in Japan as well; mandatory retirement used to be age 60, the pension receipt usually begins either at age 60 or 65, legal drinking and smoking age is 20, and government sponsored cancer screening start at age 40.

may behave in certain ways before treatment is turned on. This issue is particularly relevant for the analysis of utilization measures since there is a possibility that people may delay some expensive medical procedures until they reach 70, which may accentuate the size of the discontinuity.[37] However, in the RD setting I can visually examine whether the discontinuity is accentuated or not since if the increase is transitory rather than permanent, I should observe tendency after age 70 to revert to the previous level as well as drop-off just below the age 70.

Second, even if there is an effect on the outcome, if the effect is not immediate, it will not generally generate a discontinuity. This issue is particularly relevant for the analysis of health outcomes. For example, lower cost-sharing at age 70 induces individuals to receive preventive care that has long-run, but not short-run, effects on mortality. In this case, I will not find any discontinuity at age 70 even though there is a long-run effect. It is infeasible to estimate long-run effects because individuals age into treatment.[38]

The underlying assumption of typical RD model still applies to age RD design; in this case, the assumption is that the expected outcomes below and above age

[37]It is not always the case that anticipation accentuates the magnitude of the discontinuity; it can also mute the discontinuity. For example, simple life-cycle theories without liquidity constraints suggest that the age profile of consumption will exhibit no discontinuity at age 67, when Social Security benefits start payment in the US.

[38]One potential way to detect the mortality change in age RD setting is to look at the change in the slope of the age profile of mortality below and above age 70 rather than change in mean at the threshold in a similar spirit as regression kink design (RKD) proposed by Card, Lee, and Peri (2009).

70 are continuous at age 70 (Hahn et al. 1999). Continuity requires that all other factors that might affect the outcome of interest trend smoothly at age 70. My empirical setting is potentially better than those using Medicare edibility of age 65 in the US, since age 70 in Japan does not coincide with changes in any other confounding factors such as employment or pension receipt.[39] A simple test for the potential impact of discontinuities in confounding variables is fitting the same models like (1.1) for confounding variables and testing for discontinuities at age 70 (Lee and Lemieux, 2010).

Table 1.4 presents estimation results that test for discontinuities in the age profiles of employment, and other outcomes from the 1986-2007 pooled CSLC (age measured in months). The estimated jumps in employment-related outcomes are small in magnitude and statistically insignificant. Figure 1.3 displays the actual and fitted age profiles of employment for the pooled CSLC sample. These profiles all trend relatively smoothly through age 70 for both genders.[40] Row (1) in Table 1.4 confirms that there is no jump in employment at age 70. In the remaining rows

[39]Even though Card et al. (2008, 2009) shows no discontinuity in employment at age 65, as Dong (2010) points out, there is an obvious difference in *slopes* above and below age 65 in the age profiles of employment. In this case, treatment effects based on standard RD estimators may be weakly identified.

[40]The mandatory retirement age in Japan used to be 60 and has gradually shifted to 65 since 2003. Pension receipt starts either 60 or 65 years of age depending on the type of job. In fact, I find that there is a sharp drop in employment at 60, and a large increase in fraction of people receiving pensions at both age 60 and 65 (not shown). Also long-term care (LTC) health insurance was introduced in Japan in 2000, but age at 70 is not used to determine the edibility for LTC. Indeed, I do not see any change at age 70 in probability of receiving LTC as shown in Table 1.4.

in Table 1.4, I also investigate the age profiles of marriage, and income related variable in the CSLC, but none of these outcomes show any discontinuities at age 70. These results lead me to conclude that employment, family structure, and family income vary relatively smoothly at age 70, and are unlikely to confound the impact of cost-sharing at age 70.

1.3.3. Elasticity under Non-Linearity and Catch-up Effects

Before showing the results on utilization, I discuss the potential bias in the estimation of the elasticity. There are two issues that may potentially bias my estimates on elasticity: non-linearity in the budget set and the catch-up effect. To illustrate the direction of potential bias, it is convenient to write the elasticity ϵ simply:

$$\begin{aligned} \epsilon &= \frac{\log(Q\_above70) - \log(Q\_below70)}{\log(P\_above70) - \log(P\_below70)} \\ &= \frac{RD\ estimates\ at\ Age\ 70}{\log(P\_above70) - \log(P\_below70)}. \end{aligned} \tag{1.3}$$

First, the non-linearity imposed by the cap on out-of-pocket medical expenditures and deductibles is classic but important challenge in estimating elasticities that dates back to the RAND HIE (Keeler et al., 1977; Ellis, 1986; Keeler and Rolph, 1988).[41] The problem is that although many medical expenditures are

[41] See also Kowalski (2011) that discusses challenges of estimating the demand elasticity under non-linear budget set. My case is simpler than her case since there is no deductible.

caused by unpredictable illnesses, economically rational individuals can anticipate some spending and can take advantage of varying prices by spending more during periods when the price is low. In the extreme case, for those whose monthly medical expenditures are already above or are expected to exceed the stop-loss, the effective price, the shadow price of consuming additional medical services, is near zero. In general, the "true" out-of-pocket price is smaller than the nominal out-of-pocket price. The size of the difference depends on the probability that the individual will subsequently exceed the stop-loss. Indeed, under fairly restrictive assumptions, it can be shown that the effective price before the stop-loss has satisfied is the simple form $(1-x)P$, where P is nominal price, and x is the probability of exceeding the stop-loss (Keeler and Rolph 1988). Since those below age 70 are more likely to reach the stop-loss, the true $P\_below70$ may be smaller than that of the nominal price, thus the bias incurred from using the observed price is downward.

Second is the catch-up effect. As I mentioned earlier, individuals may anticipate the lower cost-sharing once turning 70 and, therefore they may delay some expensive medical procedures until they reach 70, which may accentuate the size of the discontinuity. This may cause $Q\_above70$ to be larger and $Q\_below70$ to be smaller, and therefore may bias the estimates of the elasticity upward. Fortunately, I can to some extent visually examine whether the discontinuity is magnified by looking at the dip just below 70 and surge just above 70.

These two issues are less relevant for outpatient visits, since I will show later that there does not appear to be a catch-up effect, and reaching the stop-loss is very unlikely since outpatient visits are not costly. The more relevant case is inpatient admissions. I will show later that overall age trend does not seem to display any catch-up effects, but close inspection of inpatient admissions with elective surgery shows some drop-off just below age 70, and a sudden surge just over age 70. Though not far from perfect, to partially account for the catch-up effect, I run a "donut-hole" RD by excluding a few observations around the threshold. This approach was initially proposed by Barreca et al. (2011) to account for pronounced heaping in the observations around the threshold in RD framework.[42] The caveat of this methodology is that there is no clear economic or statistical consensus on the optimal size of the donut and excluding observations near the threshold undermines the virtue of the RD design, that is, comparing outcomes just below and above the threshold. Nonetheless, this donut-hole RD may show whether my RD estimates are sensitive to the catch-up effects.

Accounting for non-linearity associated with stop-loss is much harder, since to fully understand the size of the difference between true and nominal price, I may need data on episodes of illness rather than monthly aggregated data (Keeler

[42]See Bharadwaj and Neilson (2011) for an example of the donut-hole RD.

and Rolph, 1988).[43] I argue that the effect of the stop-loss on over-utilization is probably much smaller in my case rather than RAND HIE because the stop-loss is set by monthly in Japan rather than annually like the RAND HIE and most health insurances in the U.S. To the extent that illnesses are unpredictable, this shorter interval may make it harder for people to time and overuse the medical services. Keeler et al. (1977) and Ellis (1986) formally show that the more time left in the accounting period, the more the effective price falls. Furthermore, even under an annual stop-loss, Keeler and Rolph (1988) empirically shows that people in the RAND HIE respond myopically to stop-loss, i.e., people do not appear to change the timing of medical purchases to reduce costs. Nonetheless, to partially account for this effect, I simply apply formula of $(1 - x_t)P_t$ for those whose out-of-pocket medical expenditures are more than median in each survey year t since this problem is most relevant for consumers who are close to reaching the stop-loss. Since the probability of reaching the stop-loss is not high even for the inpatient admissions (14 percent for those admitted, and 2 percent for non-conditional population), the nominal price (38.0 thousand Yen) for those just below age 70 is not so different from the "true" price (35.3 thousand Yen). Therefore, the bias coming from the non-linearity associated with stop-loss may be negligible in this case.

[43] If I had disaggregated data with individual characteristics, I might have been able to partially separate the income effect from the substitution effect by identifying those who almost certainly would be beyond the stop-loss, since those on the stop-loss is only affected by income effects.

1.4. Utilization Results

In this section, I examine the effect of changes in cost-sharing on utilization. I use the pooled 1984-2008 Patient Survey for people between ages 65 and 75. I examine outpatient visits and inpatient admissions, respectively.

1.4.1. Outpatients Visits

I use the pooled outpatient data to examine changes in the number and characteristics of outpatient visits at 70. As I mentioned earlier, I collapse counts of patients by age in months, and include birth month fixed effects as well as survey year fixed effects to account for heaping and seasonality in birthdays. Therefore for most of the graphs shown in this section, the plotted average is residual from a regression of the log outcome on birth month fixed effects and survey year fixed effects.

Panel A in Figure 1.5 shows the actual and fitted age profiles of outpatient visits based on the pooled outpatient data. The markers in the figure represent actual averages of the log number of outpatient visits (by age in months). The lines represent fitted regressions from models with a quadratic age profile fully interacted with a dummy for age 70 or older. Overall outpatient visits steadily increase prior to age 70, and then jump sharply at age 70. Also, the increase appears to be permanent rather than transitory, with no tendency after age 70 to revert to the previous level, which might occur if the jump in outpatient visits only represents catching up on deferred visits.

Table 1.5 presents the summary of the estimated discontinuity for outpatient visits. All the estimates in the Table 1.5 come from the preferred model, which uses a quadratic in age, fully interacted with dummy for age 70 or older. The first entry in first column shows that the jump in Panel A in Figure 1.5 corresponds to a 10.3 percent increase.

The implied elasticity of the outpatient visits is -0.17 =(10.3/((log(1.0)-log(4.0))/100), where the denominator is the log difference in price between age 69 and age 70 from the first row in Table 1.3.[44] This estimated elasticity is similar to the estimates found in the HIE for the non-elderly (roughly -0.2), and slightly larger than that estimates for the elderly (-0.07 to -0.10) by Chandra et al. (2010). The finding indicates that the price elasticity of outpatient visits for the elderly is similar in magnitude to that of the non-elderly. Since I do not visually observe catch-up effects, and the stop-loss is rarely reached, the bias on the estimating elasticity of outpatient visits seems minimal.

Another way to look at more frequent access to outpatient care is to examine the change in the interval since the last outpatient visits. A shorter interval indicates a higher frequency of outpatient visits.[45] As much as 94 percent of patients are

[44]Note that the price in the denominator I used is the average price rather than the marginal price. Thus the elasticity estimated is with respect to the average price. However, the marginal price and the average price may not differ much. For example, as for 2008, the log marginal price difference would be log(0.1)-log(0.3) without stop-loss, while what I used here as log average price difference is log(1.0)-log(4.0) for outpatient visits and log(12.4)-log(38.0) for inpatient admissions.

[45]For this question, the Patient Survey first asks whether the outpatient visit is new or repeated. For repeated patients, then it reports the exact day of the last visit.

repeated visit patients (i.e., visits for the same underlying health conditions *and* the same hospitals or clinics as last time) rather than first-time visit patients as shown in the summary statistics in Table 1.1. The Patient Survey asks the exact day of the last outpatient visits for these repeated patients. Panel B in Figure 1.5 plots the age profile of days from the last outpatient visit for repeated patients. Consistent with the increase in outpatient visits, the duration from the last visit steadily decreases prior to age 70, and then drops sharply at age 70 by roughly one day.[46]

So far, I find compelling evidence that people use more outpatient care once they turn 70. Next, I investigate whether the increase in outpatient visits solely reflects moral hazard or increases in beneficial care. If most of the increase reflects discretionary and "ineffective" care, it suggests that increase in patient cost-sharing can reduce unnecessary health care utilization. On the other hand, if some useful preventive treatments are also price-sensitive, it may caution against raising the patient cost-sharing.

To investigate this question, I divide the sample into various dimensions in the remaining rows in Table 1.5. In Panel B, I divide outpatient visits by first visit or a repeated visit. Interestingly, the results indicate not only repeated visits but also first visits increase by more than 10 percent. Since repeated visits accounts for 94

[46]Additionally, I can use the age at the time of the last visit as a running variable to investigate whether the last outpatient visit also jumps at age 70. I find that last outpatient visits also increase discontinuously at age 70 (not shown).

percent of all outpatient visits, the increase in first visits is small in magnitude relative to total outpatient visits. But the increase in new visits raises the possibility that those newly receiving the outpatient care may avoid outpatient care due to cost reasons before turning age 70.[47]

For repeated visits, Panel C in Table 1.5 shows that most of the increases in the repeated outpatient visits are concentrated within a short interval from the last visits. In fact, most of the increase is concentrated among those who receive their last outpatient care within 7 days.[48] In Panel D, I divide outpatient visits by institutions. The increase in outpatient visits is concentrated at clinics rather than at hospitals. Since people have much easier access to small clinics than large hospitals, this result indicates that these outpatient visits are more discretionary and less serious. In Panel E, I stratify the sample by the presence of a referral. Since most referrals to hospitals are provided at clinics, an increase in non-referral outpatient visits is consistent with the increase in outpatient visits at clinics.

Most of the findings so far suggest that those who visit medical institutions for outpatient reasons once they turn age 70 are less seriously ill than those who visit at age 69. Finally, I investigate the size of discontinuity at age 70 by type of diagnoses.

[47]Appendix Figure A.1 shows the age profiles for first time and repeated outpatient visits, respectively. The age profiles of first time visits show a very interesting trend; the number of first time visits steadily decreases prior to age 70, reflecting the trend of deteriorating health as people get older, and then jumps sharply at age 70. The age profiles of repeated visits are very similar to that of total outpatient visits, since most of total outpatient visits are repeated visits.

[48]Average days from last outpatient visits among ages 65-75 are 13.6 days.

A key advantage of the Patient Survey is that I can break down outpatient visits by diagnoses. Appendix Table A.1 lists the top 10 diagnoses by three digit ICD 9 codes, which account for roughly half (45 percent) of all outpatient visits. By far the most frequent diagnosis is hypertension, which accounts for nearly 16 percent of all outpatient visits. Untreated high blood pressure can be an important risk factor for the elderly, and thus proper treatment may prevent subsequent hospitalization or even death from conditions such as heart failure, cerebrovascular disease or stroke, and heart attacks (Pierdomenico et al., 2009). Panel F in Table 1.5 first presents the results for the top 5 outpatient diagnoses: essential hypertension, spondylosis, diabetes, osteoarthrosis, and cataracts. Even though most of the large increases come from relatively elective diagnoses such as two degenerative joint diseases (spondylosis and osteoarthrosis), I also find an 8 percent statistically significant increase for essential hypertension visits.[49]

The results on hypertension raise the possibility that increases in outpatient visits may include useful preventive treatments. Figure 1.6 displays the age profile of outpatient visits for commonly examined diagnoses: heart disease, cerebrovascular disease, and respiratory disease (see e.g., Chay et al., 2010). While I do not find a statistically significant jump in visits for heart disease in Panel A, Panel B and C show that there is sharp increase in the number of outpatient visits for

[49]Indeed, a recent paper in the Lancet "What Has Made the Population of Japan Healthy?" (Ikeda et al., 2011) points out that the interventions to control blood pressure (e.g., salt reduction campaigns, and antihypertensive drugs) have contributed to the sustained extension of Japanese longevity after the mid-1960s.

cerebrovascular disease and respiratory disease, which may cause serious problems without proper preventive treatments.

I also look at the diagnoses defined as the Prevention Quality Indicators (PQI), which are measures of potentially avoidable hospitalizations for Ambulatory Care Sensitive Conditions (ACSCs) developed by Agency for Healthcare Research and Quality (Appendix Table A.3 for the list of PQI).[50] This measure is intended to study preventive care in an outpatient setting using inpatient data to identify admissions that should not occur in the presence of sufficient preventive care. Since I *do* have outpatient datasets, I can directly look at changes in the number of patients for these beneficial and preventive care. Panel D in Figure 1.6 shows that there is a large jump at age 70 for ACSCs diagnoses.

The remaining rows in Panel F in Table 1.5 confirm these patterns in the figures. In sum, I find that demand for both more and less beneficial care is price sensitive. While most of the largest increase can be found for diagnoses that may not be life-threatening but treating probably enhance the quality of life, such as diseases of genitourinary system, skin, and musculoskeletal system, I also find an increase in potentially more serious diagnoses; I find increases in outpatient visits for cerebrovascular disease, respiratory disease, and ACSCs of 15.2, 14.3, and 8.2

[50] See also Weissman et al. (1992) for another list of avoidable admissions. Both list have substantial overlaps.

percents respectively. All the estimates mentioned here are statistically significant at 1 percent level.[51]

Appendix Table A.2 summarizes the results of alternative specifications that use age in days as the running variable with birthday fixed effects, and yield quantitatively similar results for most of the outcomes.[52] As a falsification test, I also run the same estimation at other ages (each single age of 66-74) that should not have any discontinuity, and did not find any statistically significant change in other ages (not shown). This result is not surprising since I do not see any visible discontinuity in other ages in either Figure 1.5 or Figure 1.6.

1.4.2. Inpatient Admissions

Before starting the analysis of the inpatient admissions, I need to mention one potential threat to interpreting the results for impatient admissions. Since a sharp change in cost-sharing in inpatient admissions coincides with that of outpatient visits, it may be difficult to separate whether the change in inpatient admissions

[51] I also investigate each PQI measure separately, but due to smaller sample sizes, I could not obtain precise estimates for most PQIs. The two exceptions are Chronic Obstructive Pulmonary Disease (COPD; PQI5), a progressive disease that makes it hard to breathe, and hypertension (PQI7). The increase for patients with COPD is 17.2 percent (t-stat=2.10) and for all hypertension is 8.5 percent (t-stat=3.54).

[52] I choose outcomes that do not have "zero" cells for any age in days in Appendix Table A.1. It is a convention to add one or small positive value before taking log for those "zero" cells, but the "zero" cells introduces the noises and hence attenuate the estimates. In fact the estimates obtained by using age in days as running variables start to deviates from those of age in months as the number of "zero" cells increases.

for a certain condition is the result of lower inpatient cost-sharing *per se* or complementarity or substitution with increased outpatient visits. For example, effective outpatient treatments may replace avoidable impatient admissions. However, since I do not see a discontinuity with time lag, it is more likely that the jump I observe is the reflection of the lower cost-sharing rather than any complementarity.

Figure 1.7 shows the actual and fitted age profiles of inpatient admissions based on my 1984-2008 pooled discharge data. The plotted average is the residual from a regression of the log outcome on birth month, admission month and survey year fixed effects. Overall inpatient admission steadily increases prior to age 70, and then jumps sharply at age 70. The increase appears to be permanent in this case as well as outpatient visits, with no tendency after age 70 to return to the pre age 70 level.

Table 1.6 presents the summary of the estimated discontinuity for inpatient admissions. All the estimates in this Table 1.6 come from the preferred model, which includes a quadratic in age, fully interacted with a dummy for being age 70 or older. The first entry in Table 1.6 shows that the jump in overall inpatient admissions in Figure 1.7 corresponds to an 8.2 percent increase. Panel 1 in Appendix Figure A.2 shows that the result is not an artifact of how I limit the sample by admission dates; the results are pretty robust to the length of windows from the discharge date. Note that more than 90 percent of inpatient admissions occurred within three months from discharges.

The implied elasticity of the inpatient admissions is -0.17 (= 8.2/((log(12.4)-log(38.0))/100), where the denominator is the log difference in price between age 69 and age 70 from the second row in Table 1.3. As I discussed earlier, there is a potential bias in estimating elasticity especially due to the catch-up effect. To account for the catch-up effect, I run a "donut-hole" RD by excluding a few months of observations around the threshold. Since there is no guide as to the size of the donut-hole statistically or economically, I experiment with zero month to six months.[53] However, removing six months from both side of age 70 may be too drastic since it means that I am essentially comparing those aged 69.5 and 70.5, so there is one year age gap between those above and below threshold. Panel 2 in Appendix Figure A.2 shows that the estimates get smaller and the standard errors get larger as the "hole" is expanded. But as long as the removal of the data is within three months of 70, the estimates are statistically significant at 5 percent level. Taking the conservative RD estimate from the three-month donut-hole RD, the lower bound of the implied elasticity is -0.15 (= 7.2/((log(12.4)-log(38.0))/100), not so different from the "naive" elasticity.

Next, I examine the characteristics of inpatient admissions in the remaining rows in Table 1.6. First, I divide the sample by whether patients received surgery in Panel B. Interestingly, I find that the increase in admissions for people who receive surgery is larger than the overall growth in admissions (10.8 percent versus

[53]It is not clear what magnitude of delay is fathomable/medically low cost for patients. It may vary substantially by the severity of the conditions and type of diagnosis.

an overall increase of 8.2 percent) while estimates from non-surgery admissions are smaller in magnitude (5.4 percent) and marginally statistically significant. Indeed, close inspection of the age profile of patients with surgery in Panel A in Figure 1.8 reveals a drop-off just prior to 70, coupled with a temporary surge shortly after 70. This pattern suggests that some people who are close to 70 delay surgery until they become eligible for Elderly Health Insurance to reduce the out-of-pocket expenditures.

This finding raises two possibilities for physicians' and patients' role in the demand for health care services. First, it may imply that physicians may consider the financial effects of treatments on patient since there are no financial incentives for physicians to delay surgeries until age 70 because reimbursements do not differ by patient age. Or alternatively, it may raise the possibility that patients play a more active role in determining their treatments. Hai and Rizzo (2009) indeed point out that recent organizational changes (e.g., alternative sources of medical information such as the internet, health care report cards, and direct-to-consumer advertising of pharmaceuticals) may have fostered patient-initiated requests for specific treatments.

In Panel C, I further investigate the discontinuities across types of surgeries. Unfortunately, this information is only collected in the most recent four survey years (1999, 2002, 2005, and 2008), and the categorization is quite coarse. Therefore, it is difficult to obtain the precise estimates. Nonetheless, the estimates indicate that the open-stomach surgery and intraocular lens implantation, which has substantial overlap with admissions for cataracts, show statistically significant jumps at age 70.[54] Appendix Figure A.3 displays the age profile of inpatient admissions for these two procedures. Similar to the overall age profiles for inpatient admissions with surgery (Panel A of Figure 1.8), I find a drop-off just prior to 70, coupled with a temporary surge shortly after 70 for both procedures. These results are plausible since one hand these procedures are easily deferred, and on the other, they are relatively expensive but routine interventions that are thought to have a beneficial effect on quality of life (Card et al. 2008).

Appendix Table A.1 lists the top 10 diagnoses in three digit ICD 9 codes, which account for roughly half (29 percent) of all inpatient admissions. Panel D in Table 1.6 first presents the results for top 5 inpatient admission diagnoses: cataracts, angina pectoris, occlusion of cerebral arteries, diabetes, and stomach cancer. The leading diagnosis is cataracts, clouding of the lens of the eye, and I find as much as 22 percent increase in the number of inpatient admissions for cataracts. This result is consistent with the increase in surgeries for intraocular lens implantation. As expected, I do not find an increase of inpatient admissions for chronic diseases such as diabetes or stomach cancer. Surprisingly though, I find a 14 percent statistically significant increase in occlusion of cerebral arteries,

[54] Unlike Card et al. (2008), I do not find a statistically significant increase in musculoskeletal surgery, which includes joint replacements for hips and knees.

which without proper treatment may lead to one of the three most common causes of death in Japan: cerebrovascular disease (or stroke).[55]

Figure 1.9 displays the age profile of inpatient admissions for the same set of broad diagnoses as outpatient visits. The graphs in Panel A and B show that there is a sharp increase in the number of inpatient admissions for heart disease and cerebrovascular disease, which may potentially be fatal if they are acute ones.[56] The remaining rows in Panel D in Table 1.7 confirm the patterns in the figures. While I do not find any increases for chronic diseases such as cancer, I find large increases for heart disease and cerebrovascular disease. The jump in inpatient admissions for heart disease and cerebrovascular disease in Figure 1.9 corresponds to 11.5 percent and 10.5 percent increases, respectively.

I further divide heart disease and cerebrovascular disease into finer diagnoses to see whether these are acute ones recognizing the disadvantage of small sample size. The results reveal that most of the increase in admissions for heart disease come from ischemic heart disease - but chronic and not acute ones since I do not find any increase in heart attacks (clinically referred to as an acute myocardial Infarction or AMI) - and most of the increase in cerebrovascular disease, comes from the cerebral infarction, which is consistent with the increase in admissions

[55]The three leading causes of death in Japan are cancer, heart disease, and cerebrovascular disease.

[56]Unfortunately, the discharge data in the Patient Survey do not collect data on route into the hospital or whether the admission was for elective, urgent, or emergency care.

for the occlusion of cerebral arteries. On the other hand, I do not find statistically significant increase for Ambulatory Care Sensitive Conditions (ACSCs).[57]

Interestingly, the observed patterns by admission diagnoses I find here are similar to the findings in Card et al. (2008), which examines the Medicare eligibility at age 65; they find smaller increases for conditions that are typically treated with medication or bed rest (heart failure, bronchitis, and pneumonia), and large increases for those are treated with specific procedures (chronic ischemic heart disease, and osteoarthrosis). While I do not find an increase in admissions for respiratory diseases, and ACSCs that are typically treated with medication, I also find increases for cataracts, cerebral infarction (including occlusion of cerebral arteries), (chronic) ischemic heart disease, which may require procedures, such as intraocular lens implantation, open-head or open-heart surgery.[58] These results imply that diagnoses that are treated with expensive but elective procedures are quite price sensitive, probably due to its large cost, and hence patients delay to reduce the out-of-pocket expenditures.

Finally, I also examine the interaction between the outpatient visits and inpatient admissions by looking at the route before admission to hospitals. Panel

[57]The RD estimates for COPD is 1.6 percent (t-stat=0.34) and for hypertension is 3.2 percent (t-stat=0.58).

[58]The fact that I did not find any decline in inpatient admissions for ACSCs is potentially interesting. If the outpatient care takes care of these conditions, and hence replace inpatient admissions, I should see a corresponding decline in inpatient admissions for these conditions. On the other hand, if seemingly "effective" care at outpatient visits still includes some moral hazard, I may not see any change in inpatient admissions from these conditions.

E in Table 1.6 shows that there is statistically significant 9.7 percent increase in admissions that come from the outpatient visits within the same hospitals. This increase is slightly larger than the overall increase in admissions (8.2 percent), implying that patients wait and switch from outpatient visits to inpatient admissions within the hospital once cost-sharing for inpatient admissions is reduced drastically at age 70. This pattern is consistent with the possibility that physicians take the financial burden on patients into account when they provide expensive medical services.[59]

Appendix Table A.4 shows the results of alternative specifications for selected outcome variables. The table shows that the results are quite robust to different specifications such as limiting the sample to narrower age window (ages 67–73) and including a cubic polynomial in age, fully interacted with a dummy for age 70 or older. However, specifications with a cubic polynomial in age sometimes give larger estimates due to a drop-off in number of inpatient admissions just prior to 70.

[59] I also divide the inpatient admissions by the characteristics of hospitals in Appendix Table E. Consistent with the notion that patients can freely choose medical institutions, patterns do not differ by hospital ownership. This result is in stark contrast to the U.S.; Card et al. (2008) finds that with the onset of medical eligibility, hospital admissions to both private non-profit and private for-profits hospitals experience relatively large increases in admissions, while hospitals owned by large and long-established HMOs show little change, and county hospitals experience a sharp decline. Another possibility for this difference is that there is not much difference in the quality of hospitals by ownership or size in Japan. Also note that there are no for-profit hospitals in Japan since the hospitals are not allowed to issue shares and distribute the earnings.

1.5. Results on Benefit

To look at the benefit side of cost-sharing, I first explore whether lower cost-sharing benefits the health of those above age 70, and next examine risk reduction.

1.5.1. Health Outcomes

As a measure of health outcomes, I examine both mortality and morbidity. Overall, I do not find statistically significant improvements in health at age 70 despite utilization changes.

A priori, the impact of cost-sharing on mortality is ambiguous. On the one hand, cheaper access to health care services may reduce mortality.[60] On the other hand, lower cost-sharing may increase mortality if those who are just below 70 delay life-saving treatment. Most importantly, if the marginal patient is not severely ill, I may find no effects on mortality.

Figure 1.10 shows the age profiles of the log of overall deaths among those between the ages of 65 and 75 using pooled 1987-1991 mortality data. Even though there is slight decline at age 70 in the log counts of mortality, first entry in Column (1) in Table 1.7 shows that the size of the estimates (-0.7 percent) is not statistically significant at conventional level. I also estimated different specifications, including

[60] Also it is possible that more frequent interactions with physicians could increase peoples' awareness of the health consequences of behavioral risk factors such as smoking. Alternatively, it is also possible that by reducing the adverse financial consequences of poor health, lower cost-sharing may discourage investments in health and health-related behaviors, and thereby worsen health outcomes (*ex-ante* moral hazard).

local-linear regressions, but they yield similar results as shown in the remaining columns.[61]

I also examine cause-specific deaths for three leading causes of death among the elderly in Japan: cancer, heart disease, cerebrovascular disease, plus respiratory disease. Appendix Figure A.4 show the there are no disenable patterns for any causes of death. The remaining rows in Table 1.7 confirm that there is no clear change in the cause-specific mortality at age 70, even though in some specifications the estimates become marginally statistically significant. These results are to some extent as expected, since in general, it is hard to detect the effect on health in a regression discontinuity framework, since health is stock (Grossman, 1972); thus it may take a while for most observable effects to be realized, unless the causes of death are acute, such as heart attacks or stroke (see e.g., Card et al., 2009; Chay et al., 2010). I also examined more acute causes of death such as heart attacks or stroke but did not find any disenable patterns in age profile (not shown).[62]

Next, I examine trends in self-reported health as a morbidity measure before and after age 70. It is also not clear whether self-reported health will improve. On one hand, it is possible that more preventative care leads to improvements in subjective health if certain health problems can be resolved quickly, or if uncertainty

[61]For bandwidth selection, I use rule of thumb bandwidth procedure proposed by the Fan and Gijbels (1996) assuming a triangular kernel. I then estimate the local linear regression using the triangular kernel with the estimated bandwidth, and also report asymptotic standard errors (Porter 2003).

[62]Results are available from author.

about a chronic condition can be resolved. On the other hand, it may worsen subjective health if increasing contact with the physicians causes individuals to learn about previously unrecognized health problems (Card et al., 2004).

The respondents to the CSLC report health on a five-point scale (very poor, poor, fair, good, or very good). Appendix Figure A.5 shows the age profiles of the fraction of the people who report themselves to be in good, or very good health (31 percent of the population), based on pooled 1984-2008 CSLC samples. The graph shows that self-reported health is gradually declining with age but I do not find any observable change at age 70. Appendix Table A.6 confirms this age pattern. Column (2) presents estimates from linear probability models for the probability that people report that their health is good or better. Column (4) reports estimates from a simple linear regression for the mean assessment of health (assigning 1 to poor health and 5 to very good). Consistent with the patterns in Figure A.5, none of the estimates in Table A.6 are associated with statistically significant changes in any of self-reported health. In the remaining columns, I also look at the mental health, but I did not find any changes in mental health outcomes either.

Overall, I do not find any evidence that lower cost-sharing leads to a discrete jump in morbidity or mortality.[63] These results are not surprising, since the findings in the utilization imply that the marginal patient receiving health care because

[63]Card et al. (2004) also did not find any impact of Medicare eligibility on self-reported health, while Finkelstein et al. (2011) find large improvement among the Medicaid beneficiaries in Oregon. The difference may arise from the fact that Medicaid recipients in Oregon are poorer and less healthy, so there is a large scope for improvement of self-reported health.

of lower cost-sharing is not severely ill, and also it is unlikely that people delay life-saving procedures.

1.5.2. Risk Reduction

Other than improved health, another benefit of lower cost-sharing is a lower risk of unexpected out-of-pocket medical spending. As Finklestein and McKnight (2008) point out, this benefit is often overlooked in the literature. For example, neither the RAND HIE nor Chandra et al. (2010) analyze the impact of cost-sharing on exposure to out-of-pocket medical expenditure risk. And yet, some claim that protection against large medical expenditure risk is arguably the primary purpose of health insurance (e.g., Zeckhauser, 1970). Indeed, for risk averse individuals, the largest welfare gains from lower cost-sharing come from reducing catastrophic negative shocks to consumption.

To examine the effect of cost-sharing on risk reduction, I use self-reported out-of-pocket medical expenditure in the CSLC. Unfortunately, CSLC started collecting this information in 2007, thus I only have one survey year of individual out-of-pocket expenditures. The out-of-pocket medical expenditure includes any medical expenses such as over-the-counter drug spending which is not covered by health insurance, and does not distinguish the outpatient visits and inpatient admissions. With these caveats in mind, my primary interest is to examine total individual out-of-pocket medical expenditures, regardless of how they were spent. Therefore in the analysis in this section, I focus on the data in year 2007. My analysis is based on 66,112 individuals between age 65 and 75 with non-missing out-of-pocket medical expenditure. The average annual out-of-pocket spending among those aged 65-69 is 142 thousand Yen ($1,420) while median out-of-pocket medical expenditure is 48 thousand Yen ($480).

I start with presenting an RD estimate at the mean on out-of-pocket medical expenditures by estimating (1.1) where the model assumes quadratic in age fully interacted with post 70 dummy. First row in Table 1.8 shows that lower cost-sharing is associated with decline in out-of-pocket medical expenditure by 52 thousands Yen ($520), but the estimate is close to but not marginally statistically significant at the conventional level (t-stat = -1.47). However, the mean impact may miss the distributional impact of the lower cost-sharing (Bitler et al., 2006). As is well known, the distribution of out-of-pocket spending is highly right-skewed. Among those age 65-69, the top 5 percent of spenders account for almost 40 percent of the out-of-pocket medical spending, while 72 percent of the sample has out-of-pocket spending below 100 thousands Yen ($1,000) in a year.

Panel A in Figure 1.11 shows that lower cost-sharing at age 70 overwhelms the utilization effect. The graph compares the distribution of out-of-pocket medical expenditure in 2007 for 65-69 year olds (not covered by Elderly Health Insurance) and 70-74 year olds (covered by Elderly Health Insurance) in 2007. The graph reveals that 70-74 year-olds at the top of the distribution spend substantially less than 65-69 year-olds despite the large benefits from stop-loss for 65-69 year-olds. This result is consistent with other studies in the US that show a pronounced decline in a right-tail in the distribution of the out-of-pocket medical expenditures through Medicare Parts A and B (Finkelstein and McKnight, 2008), Medicare Part D (Englehardt and Gruber, 2011), and Medicaid (Finkelstein et al., 2011). These studies look at the effect of insurance coverage rather than changes in generosity.

One concern in the above analysis is that I may merely pick up an underlying change in the spending distribution that differs systematically by age group. Panel B in the same figure examines out-of-pocket medical expenditures among an adjacent age group (age 60-64) to the near-elderly (age 65-69), neither of whom benefit from lower cost-sharing. The figure shows that out-of-pocket medical expenditures among 65-69 year-olds is higher than among 60-64 year-olds, showing that medical expenditure tend to increase with age. This finding is reassuring; it suggests that that I am not measuring any systematic change in spending by age groups.

1.5.2.1. RD Estimates at Each Quantile. To put this analysis into more RD framework, Panel A in Figure 1.12 shows the age profiles of the out-of-pocket medical expenditures at 75th, 90th, and 95th percentiles. Out-of-pocket medical expenditures steadily increase prior to age 70, reflecting worse health as people age, and then decline sharply at age 70 at all three percentiles, with the largest decline at the highest percentile.

To gauge the magnitude of the decline, I estimate the following equation for each quantile q

$$M_i^q = \alpha_0^q + \alpha_1^q Post70_i + f^q(a) + X_i^{'}\gamma^q + \varepsilon_i, \tag{1.4}$$

where M_i^q is the out-of-pocket medical expenditure at quantile q, and $f^q(a)$ is a quantile-specific smooth function of age, where age a is normalized to zero at age 70. X_i are demographic controls in the form of dummy variables for marital status, gender, region and birth month.

Panel B in Figure 1.11 plots the RD estimates at age 70 on each quantile (α_1^q), along with their 95 percent confidence interval. The standard error is computed based on the empirical standard deviation of 200 bootstrap repetitions of quantile treatment estimates.[64] Note that the coefficient and standard errors on the post70 dummy are not multiplied by 100 throughout this section. The figure shows that lower cost-sharing at age 70 is associated with declines in out-of-pocket spending at almost all (non-zero) quantiles of the distribution.

Table 1.8 reports the RD estimate (α_1^q) of each tencile above 40 percentile, and 95th and 99th percentile in column (2), with a value just below age 70 (α_0^q) in column (1). While the lower cost-sharing has a very small effect at the low quantiles, it grows consistently with baseline spending. At the median, the impact on out–of-pocket spending is a reduction of 23.5 thousands Yen; at the 95th quantile

[64]See Frandsen, Froelich and Melly (2010), and Froelich and Melly (2010) that propose the nonparametric estimator for quantile treatment effects in a RD design. Recognizing the potential bias due to the misspecification, I choose to use parametric approach since I also want to obtain the coefficients on other controls variables that are used to derive the distribution of out-of-pocket medical expenditure at each quantile conditional on individual characteristics later in the welfare analysis. In fact, I also estimate the proposed non-parametric estimators, and compare it to the parametric ones. The estimates are quite similar throughout the percentile except for slight deviation among the top 3 percentile. The results are available from the author. The stata code for the non-parametric estimator is available at Frandsen's website. http://econ-www.mit.edu/grad/frandsen/software

it grows to 115 thousands Yen, roughly a 30 percent decline from the value just below age 70. Note that the estimates reflect the effect of treatment on the distribution, not the effect of treatment on any particular individual without a rank invariance assumption.

1.6. Cost-Benefit Analysis

In this section, I carry out a simple cost-benefit analysis. Since it requires making a number of assumptions, the results here are more speculative. But the exercise provides a rough estimate on the social costs and benefits of marginal change of the cost-sharing at age 70.

To understand the costs and benefits in this framework, I first describe the items of social costs and benefits associated with the change in the price of the health care services at age 70. The program incurs two types of the costs. First is extra spending for mechanical reasons, i.e., the government has to bear additional payments due to higher reimbursements for the consumers above age 70 (denote this item #1). The other is efficiency costs from moral hazard on increased health spending (#2). The sum of #1 and #2 is the amount of the increase in spending out of government funds. Since there are marginal costs associated with raising public revenue, these numbers have to be multiplied by the marginal cost of funds (MCF) to estimate the total social cost. On the benefit side, there are two benefits. First is the mechanical gain by the lower cost-sharing accrued to the consumers, which is exactly the mirror image of the increase in the government reimbursement (i.e., #1). The other benefit is risk protection against unexpected out-of-pocket medical spending which I explain in length later (#3). Thus net benefit can be written as follows.

$$\begin{aligned} (1.5) \qquad Net\ Benefit &= (Total\ Benefit) - (Total\ Cost) \\ &= (\#3 + \#1) - MCF * (\#1 + \#2) \\ &= \#3 - (MCF - 1) * \#1 - MCF * \#2 \end{aligned}$$

Note that the mechanical cost is multiplied by the (MCF-1), which is the excess burden of the public fund or dead weight loss, while the moral hazard is multiplied by MCF, since there is no benefit accrued by consumers to offset the cost. In the following, I estimate each component, #1, #2, and #3 accordingly.

1.6.1. Social Cost

The first cost is the mechanical cost. Since the out-of-pocket medical expenditures reported in CSLC do not distinguish the outpatient visits and inpatient admissions, I need to make an assumption to estimate the out-of-pocket spending distribution that mechanically adjusts for what the Elderly Health Insurance would have covered if it were applied to those just below age 70. Since the coinsurance rate for both inpatient admissions and outpatient visits is 30 percent for those below 70, and 10 percent for those above age 70 in 2007, I assume that two thirds of the out-of-pocket medical expenditures just below age 70 is the mechanical cost (i.e., I assume that the cost-sharing would have been one third if Elderly Health Insurance was mechanically applied to those just below age 70).[65] Since the average out-of-pocket medical expenditure just below age 70 from the first row of Table 1.8 is 152 thousand Yen, the average mechanical cost is 102 thousand Yen ($1,020).

Second, there are efficiency costs from the moral hazard on increased health spending. As seen from the results on utilization, some of the increased spending may have been socially inefficient. However, it is difficult to know exactly what would be the socially efficient use of the medical services. By treating all of the increase in utilization as a social cost, I provide an upper bound on the efficiency costs of the lower cost-sharing. The difference between the counterfactual and actual out-of-pocket medical expenditure just above age 70 should be moral hazard. From first row in column (1) in Table 1.8, the counterfactual mean value of the out-of-pocket medical expenditure is 51 thousand Yen (=152/3). The actual out-of-pocket medical expenditure just above the cut-off is 100 thousand Yen (152-52) from the first row of Table 1.8, and therefore moral hazard is remaining 49 thousand Yen.

[65] This assumption is reasonable since only 2 percent of those aged 65-69 pay beyond the stop-loss in the sample. Note that Table 1.3 shows that 14.6 percent of those ages 65-69 reach stop-loss conditional on being admitted.

1.6.2. Social Benefit: Welfare Gains from Risk Protection

To estimate the value of the reduction in risk exposure, I combine the expected utility framework with the quantile RD estimates in the previous section, and calculate the change in the risk premium associated with out-of-pocket expenditure as a measure of the welfare gain from the lower cost-sharing at age 70. This approach is akin to Feldstein and Gruber (1995), Finkelstein and McKnight (2008), and Englehardt and Gruber (2011).[66]

Specifically, I assume that each individual has utility $U(C)$ that is the function of net non-health consumption C. I then assume the individual must satisfy a budget constraint each period $C = Y - M$, where Y is per-period income and M is individual's out-of-pocket medical expenditures. M is a random variable with probability density function $f(M)$ with support $[0, \bar{M}]$.

I calculate the change in the risk premium associated with lower cost-sharing by computing the risk premium for both just below (denoted as zero) and above 70 (denoted as one). For those just below age 70, the risk premium (or certainty equivalence) π_0 can be defined by a fixed amount such that

$$U(Y - \pi_0) = \int_0^{\bar{M}} U(Y - M_0) f(M_0) dM_0, \tag{1.6}$$

[66]My welfare estimates may be bound to be lower than those in the US since it is much less likely to have catastrophic health expenses in Japan due to stringent control of national fee schedules by the government (Ikegami and Campbell 1995).

and measures the amount a risk-averse individual would be willing to pay to insure against random variation in out-of-pocket spending.

For those just above age 70, lower cost-sharing at age 70 reduces not only the variance but also the mean of the out-of-pocket spending distribution. However, since the difference between the mean values of M_0 and M_1 is simply a transfer between the insured and insurers (or government), I calculate the certainty equivalence for the out-of-pocket risk distribution just above age 70 with an adjustment to make the mean of the risk distribution just above age 70 equal to that of just below age 70 distribution (i.e., I evaluate the mean preserving spread in risk).

Thus I define the risk premium π_1 for those just above age 70 as

$$U(Y - \pi_1) = \int_0^{\bar{M}} U(Y - M_1 + \mu_1 - \mu_0) f(M_1) dM_1, \tag{1.7}$$

where μ_0, and μ_1 are the mean of M_0, and M_1 respectively.

A decrease in risk exposure just above relative to just below 70 is reflected as decline in the risk premium; the absolute value of this decline Δ provides a measure of the insurance value and hence welfare gain of the lower cost-sharing:

$$\Delta = \pi_1 - \pi_0. \tag{1.8}$$

I measure Δ in the two steps as follows. First, I use the quantile estimates of the parameters in (1.4) to calculate for each individual i in the sample the quantiles of the out-of-pocket spending distribution $\hat{M}_i^q$, conditional on individual's characteristics X_i' just below and above age 70.

Specifically, for each $i = 1, ..., N$ in the sample, $\hat{M}_{i0}^q$ for those below age 70 can be written as

$$\hat{M}_{i0}^q = \hat{\alpha}_0^q + X_i'\hat{\gamma}^q, \tag{1.9}$$

respectively for $q = 1, ..., 99$ where $\hat{\alpha}_0^q$ and $\hat{\gamma}^q$ come from equation (1.4) at each quantile q.

For those above age 70, I compute the counterfactual out-of-pocket spending distribution the individual faces once the quantile treatment estimates of lower cost-sharing estimated in equation (1.4) are applied. Therefore $\hat{M}_{i1}^q$ for those above age 70 can be written as

$$\hat{M}_{i1}^q = \hat{M}_{i0}^q + \hat{\alpha}_1^q, \tag{1.10}$$

where $\hat{\alpha}_1^q$ is the RD estimate from equation (1.4) for each quantile q. Because there are 99 quantile estimates for each individual i, to make sure that the sum of the probabilities is one, I set conditional out-of-pocket spending at the very bottom of the distribution to zero, $q = 0$, i.e., $\hat{M}_{i1}^0 = \hat{M}_{i0}^0 = 0$. Then I now have 100 points of equal probability of occurrence in the out-of-pocket spending distribution for each individual. Following Finkelstein and McKnight (2008), and Englehardt and

Gruber (2011), I truncate predicted out-of-pocket spending from below at zero and from above at 80 percent of individual income as a benchmark.

Finally, I calculate the risk premium π_{0i} for those below age 70 for each individual i by solving

$$U(Y - \pi_{0i}) = \frac{1}{100}\left[\sum_{q=1}^{99} U(Y_i - \hat{M}_{0i}) + U_0\right], \tag{1.11}$$

where $U_0 = U(Y_i)$, and the right hand side is the average utility given its income Y_i for each individual. In a similar vein, I calculate the risk premium π_{1i} for just above age 70 by solving

$$U(Y - \pi_{1i}) = \frac{1}{100}\left[\sum_{q=1}^{99} U(Y_i - \hat{M}_{1i} + \hat{\mu}_1 - \hat{\mu}_0) + U_1\right],$$

where $U_1 = U(Y_i + \hat{\mu}_1 - \hat{\mu}_0)$, and I made an adjustment by subtracting from the individual's income the average difference in out-of-pocket expenditures between one's 100 estimates for the original distribution just below age 70 ($\hat{\mu}_0$) and one's 100 estimates for the counterfactual distribution ($\hat{\mu}_1$).

Following the literature, I specify constant relative risk aversion (CRRA) utility function $U(C) = -\frac{1}{\phi-1}C^{1-\phi}$, which implies Arrow-Pratt measure of relative risk aversion of $\phi = -\frac{CU''}{U'}$. Table 1.9 summarizes the results. For a typical risk aversion of 3 in CRRA utility (see e.g., McClellan and Skinner, 2006), I estimate that this decline in risk premium, or welfare gain, is 20 thousands Yen (\$200) per person. This is just half of the average cost through moral hazard.

However, it is important to note that the previous estimate on the decline in risk exposure is understated since the out-of-pocket expenditures include the behavioral response of increased utilization of the health care services. Here I once again assume that the cost-sharing would have been one third if Elderly health Insurance was mechanically applied to those just below age 70. Column (2) in Table 1.9 shows the decline in risk premium associated with lower cost-sharing using this mechanically adjusted out-of-pocket spending. For a typical risk aversion of 3 in CRRA utility, I estimate that this decline in risk premium is doubled from 20 to 46 thousands Yen per person.

These estimates are somewhat sensitive to two particular assumptions: risk aversion and maximum level of out-of-pocket medical expenditures as a share of income. The remaining row in column (2) shows the sensitivity of the welfare gain to these two parameters. First, I examine the sensitivity to the choice of risk aversion coefficient (assuming the cap on out-of-pocket spending is 80 percent of income). Compared to an estimated welfare gain of 46 thousand Yen per person with a relative risk aversion of 3, the welfare gain falls to about 7 thousand Yen with relative risk aversion of 1, and rises to 110 thousand Yen with the relative risk aversion of 5.

Next, the welfare estimates are also sensitive to the assumption I make about the maximum level of out-of-pocket medical expenditures as a share of income (assuming relative risk aversion of 3). If I replace my baseline 80 percent cap with a cap of 60 percent, the estimated welfare gain falls from 46 thousand Yen to 22 thousand Yen, and if I impose a cap of 90 percent the welfare estimate rises to 74 thousand Yen.

Finally, the row B in Table 1.9 shows the risk premium at other percentiles. Recall that my central estimate of risk premium on average is 46 thousand Yen. I assume a relative risk aversion of 3 and out-of-pocket expenditure cap at 80 percent of income here. The median is 25, suggesting that benefits accrue more to those on the right tail. The 95th percentile is 126 thousand Yen. The results suggest that the risk-reduction gain was modest for most elderly, but sizeable for those at the highest risk of spending.

1.6.3. Discussion

My central estimate of risk reduction is 46 thousand Yen per person (\$460). One way to gauge the size of the estimate is to simply plug estimated benefits and costs into equation (1.5) and calculate the MCF that would have for the two to be equal each other. Since I have the estimated values for all components (#1, #2, and #3), it is straightforward to derive that such MCF is equal to 0.98, or in other words, the MCF should be less than 0.98 to have positive net benefits. This value is smaller than the most of the estimates of MCF in 1990s like 1.3 (see e.g., Poterba, 1996; Jorgenson and Yun, 2001).[67] Put differently, assuming the MCF

[67] There is no consensus estimate of MCF since MCF depends on behavioral responses to taxation and may differ by every country at every point in time. Nonetheless, to have a rough estimate, I here focus on income tax since it is a major source of taxes. The simplest formula is $\frac{1}{(1-\rho*(\frac{t}{1-t}))}$ where ρ is the elasticity of taxable income and t is the income tax rate (Kopczuk, 2005). Assuming

is 1.3, the sum of the program financing costs and moral hazard suggests that the total annual social cost was 94.3 thousands Yen (102*0.3+ 49*1.3) per elderly beneficiary; the deadweight loss associated with program financing is responsible for one third of the total cost, and moral hazard accounts for two-thirds. Therefore, with a relatively high risk aversion of five where risk reduction is 110 thousand Yen is the only case I examined here that average social benefit outweighs average social cost.

1.7. Conclusion

Rising medical expenditures present a serious challenge for many developed countries as countries age since the elderly consume many more medical services than the non-elderly. Expansion of health care coverage is another concern for rising medical expenditure. Even the United States, which has been a rare exception in developed countries without universal coverage, is moving towards near-universal coverage through health care reform passed in March 2010 (Patient Protection and Affordable Care Act). Once the universal coverage is achieved, the only way to control cost on the demand-side is the cost-sharing in a form of coinsurance, deductable, and stop-loss.

In this paper, I exploit a sharp change in cost-sharing at age 70 in a regression discontinuity framework to examine whether cost-sharing can affect utilization,

that both the elasticity of taxable income and the tax rate are 0.4, MCF would be 1.36, which is close to 1.3 used here.

health and risk reduction of the elderly in Japan. I find that a reduction in cost-sharing at age 70 substantially increases health care utilization. The corresponding elasticity I find is modest, around -0.2 for both outpatient visits as well as inpatient admissions, which is comparable to estimates found in the RAND HIE for the non-elderly. I also find that lower cost-sharing at age 70 overwhelms the utilization effect yielding reductions in out-of-pocket expenditures. However the welfare gain of risk protection from the lower cost-sharing is relatively small compared to the deadweight loss of program financing, suggesting that the social costs may outweigh the social benefits. This study shows that increased cost-sharing may be achieved without decreasing total welfare.

There are a number of caveats to my welfare calculation. On the one hand, the stylized welfare calculations may overstate the welfare gains since the use of a one-period model ignores the possibility that individuals can use savings or other mechanisms to smooth expenditure risk over several periods, which may lead me to over-state the welfare gains from lower cost-sharing. This may be the case since the elderly seem to have some savings.[68] On the other hand, the welfare gains may be understated because the calculations were based on an annual, rather than lifetime, measure of medical expenditure risk. In fact, there is some evidence that out-of-pocket medical expenditures are positively serially correlated (Feenberg and

[68] Average net savings at age 69 is 5,418 thousands Yen, which is roughly two and half times of average annual income (1,860 thousand Yen). Since saving and debt is only reported at the household level, I divide the net saving (i.e., saving minus debt) by the number of household members.

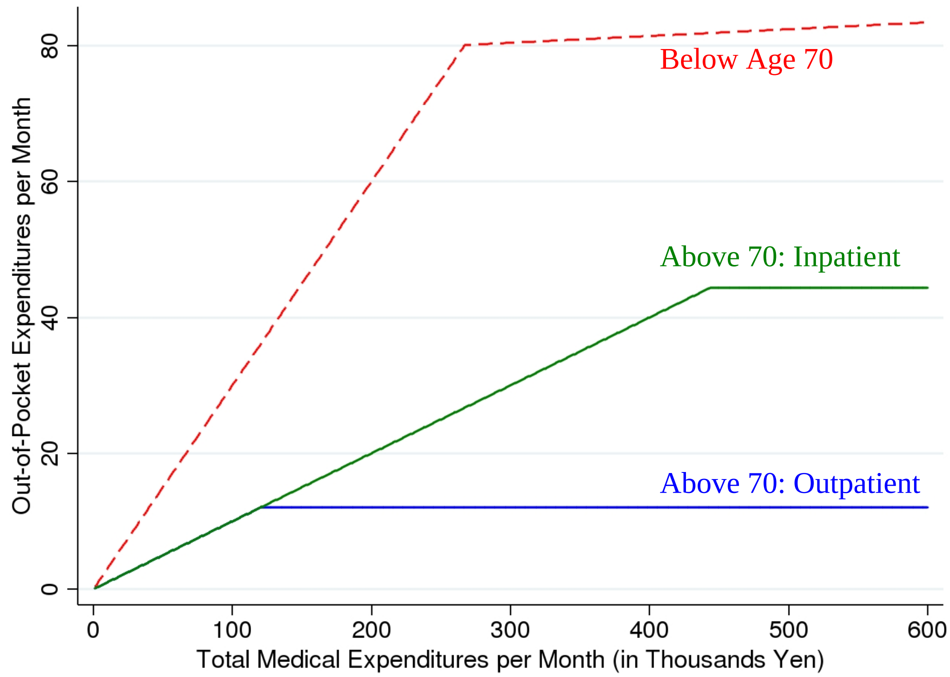
Skinner, 1994; French and Jones, 2004). These studies suggest that the lifetime distribution of out of pocket spending may be even more right-skewed than the annual distribution; therefore, the reduction in risk exposure in the lifetime scale may be even greater.[69] Furthermore, my welfare calculation does not incorporate the welfare gains from the health improvements. While I do not find any *short-term* reduction in mortality or improvement in any self-reported health measures, it is possible that preventive care induced by the lower cost-sharing may prevent future severe health events, and thus improve health in the long run. Estimating the long-term effect of cost-sharing on health is beyond the scope of the current paper, but it clearly remains an important topic for future research.

[69] Also the stylized model treats medical expenditures as affecting the budget constraint only and does not allow for any utility change from increased medical expenditures.

Figure 1.1: Age Profile of Health Insurance Type

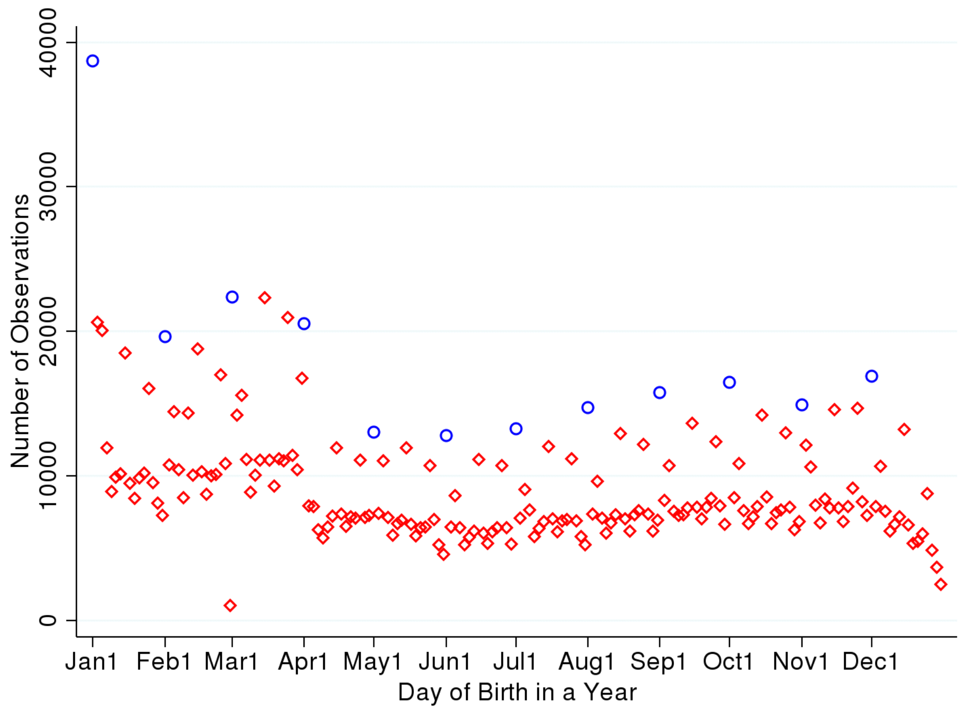
Note: The data come from the pooled outpatient visit data in the Patient Survey. Age is aggregated by month. People over 70 and bedridden people over age 65 are eligible for Elderly Health Insurance. NHI stands for National Health Insurance, by which most of the retired are covered. Employment-based Health Insurance covers both employees and dependents of employees.

Figure 1.2: Cost-Sharing Below 70 and Above 70: Year 2008 as an Example



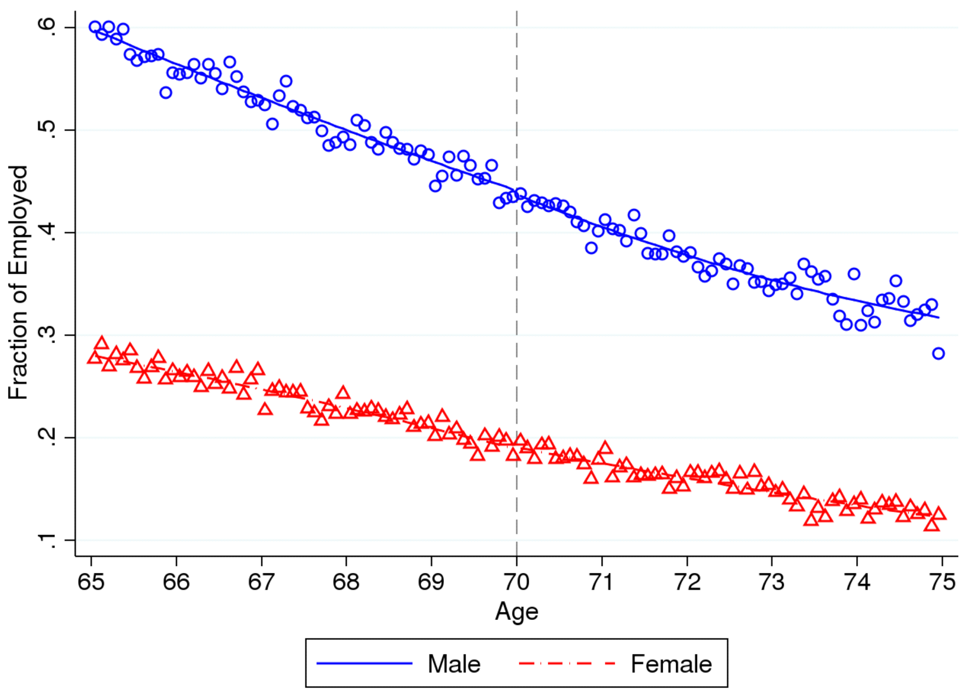
Note: See Table 2 for the formula for cost-sharing below and above 70. For those above 70, since the coinsurance rate and stop loss differs by outpatient visits and inpatient admissions, there are two separate lines for each outpatient visits and inpatient admissions. For those below 70, there is no distinction between outpatient visits and inpatient admissions in year 2008. One thousands Yen is roughly $10 US dollars.

Figure 1.3: Seasonality in Day of Birth in the Patient Survey Data



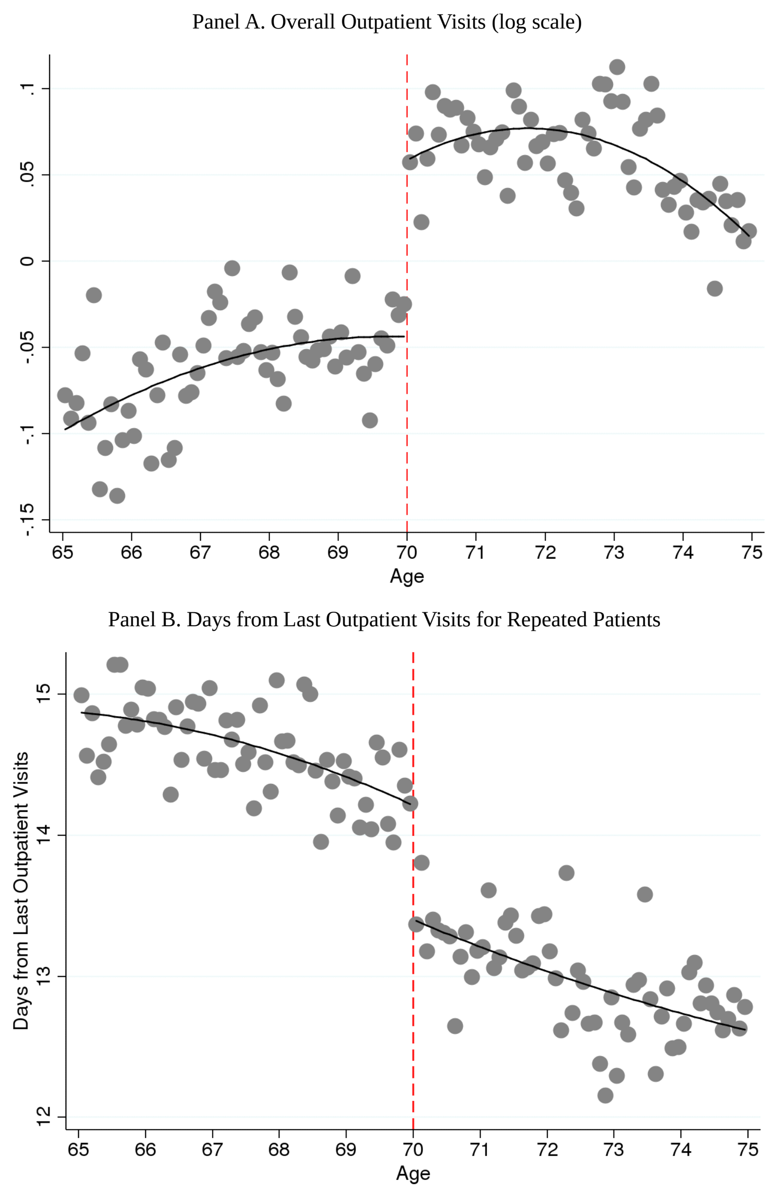
Note: The data comes from pooled 1984-2008 outpatient visit data in the Patient Survey. The circles indicate the first day of the month. Very similar patterns of birth distribution are observed in discharge data in the Patient Survey and mortality data as well.

Figure 1.4: Age Profile of Employment by Gender (1987–2007 CSLC)



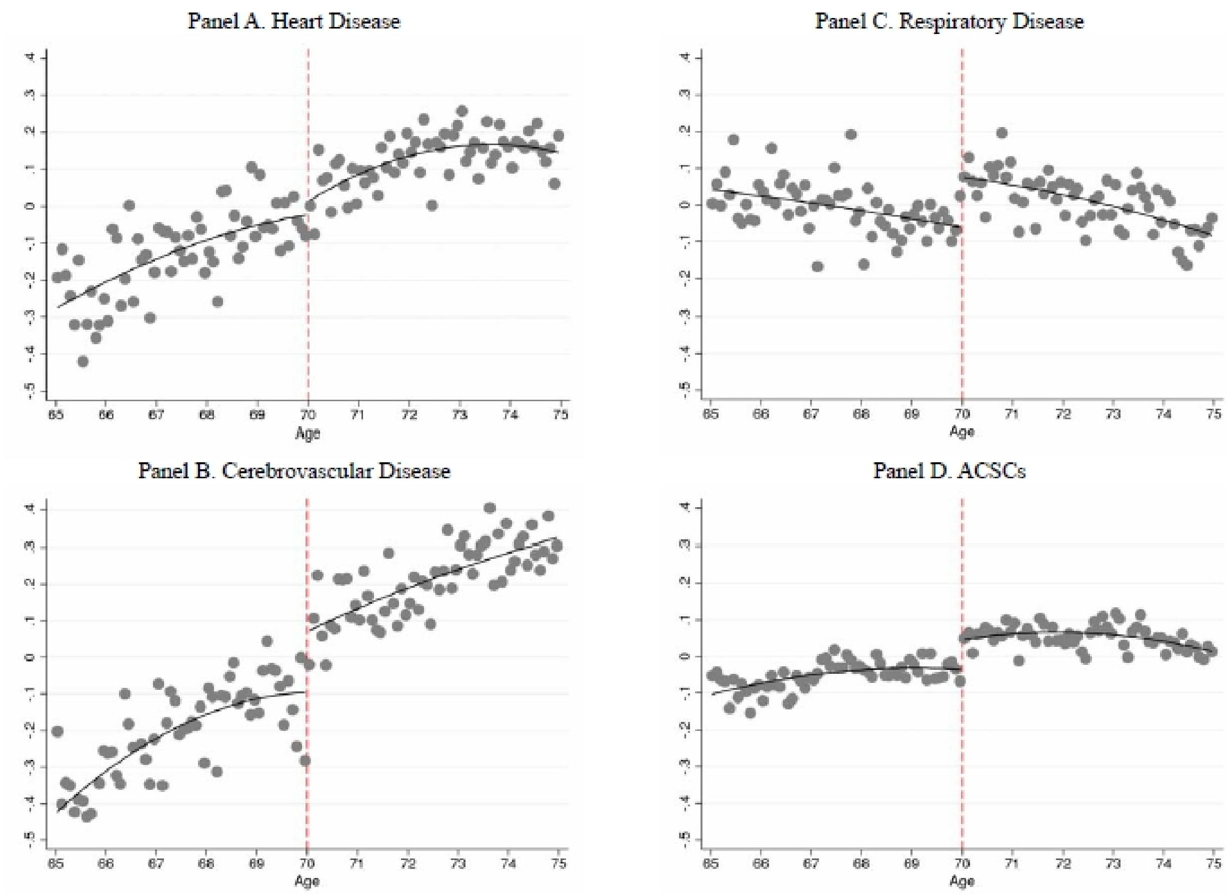
Note: The data come from the pooled 1986-2007 Comprehensive Survey of Living Conditions. The markers represent actual averages (age in month), and the lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older for male and female separately.

Figure 1.5: Age Profile of Outpatient Visits



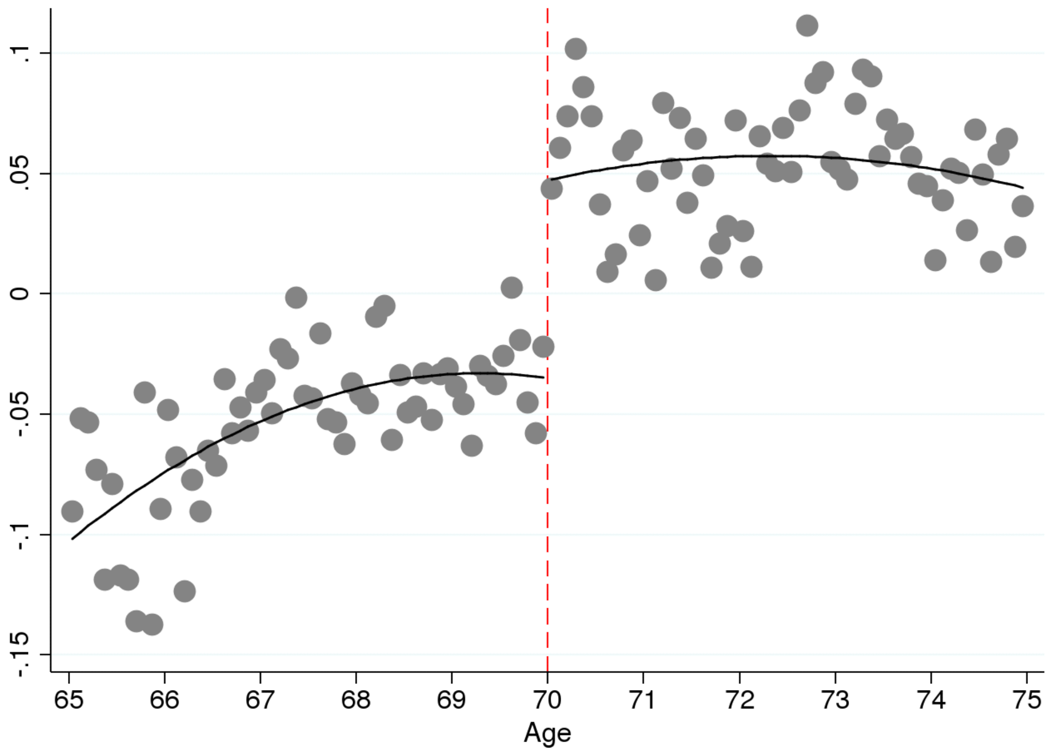
Note: The data come from pooled 1984-2008 outpatient visits data in the Patient Survey. The markers in Panel A represent the averages of residuals from a regression of the log outcome on birth month fixed effects and survey year fixed effects (aggregated by age in month), and the simple average in Panel B. The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure 1.6: Age Profile of Outpatient Visits for Selected Diagnosis (log scale)



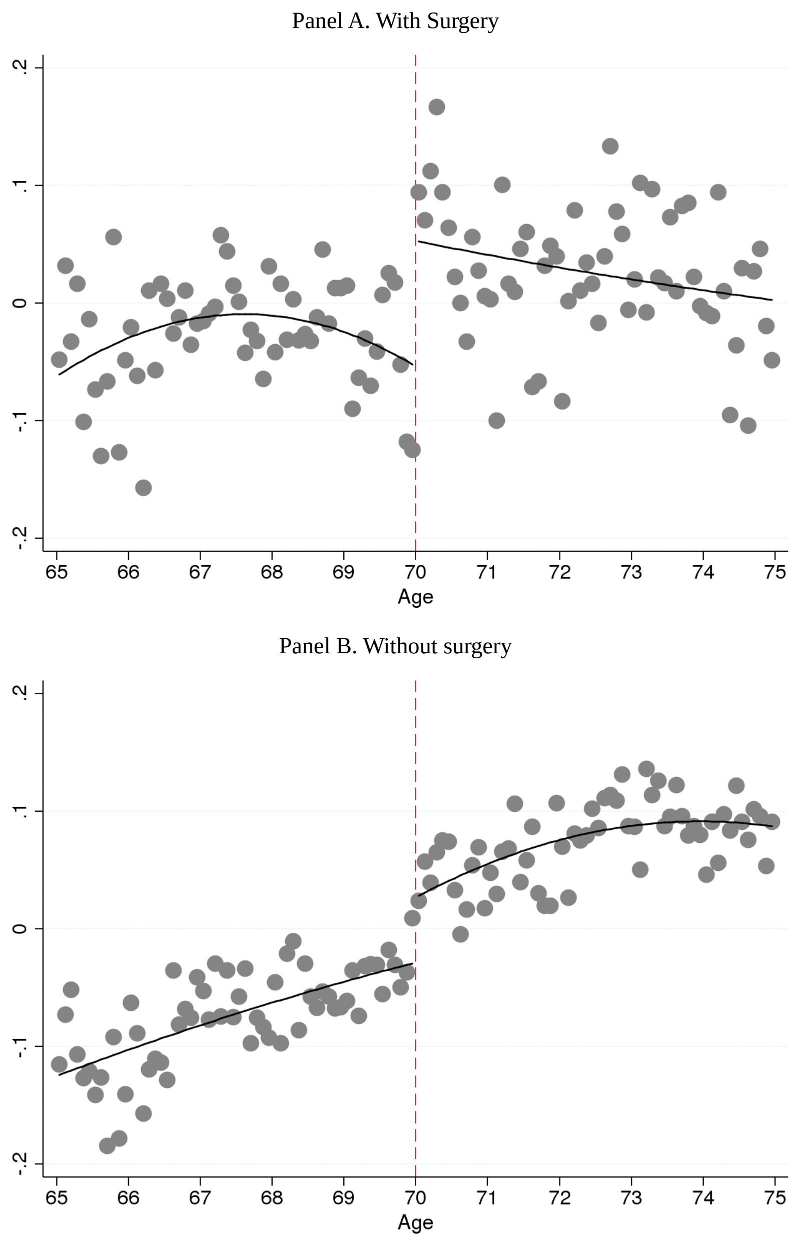
Note: The data come from pooled 1984-2008 outpatient data in the Patient Survey. The corresponding RD estimates at age 70 are statistically significant at 5 % level except for Panel A. The markers represent the averages of residual from a regression of the log outcome on birth month fixed effects and survey year fixed effects (aggregated by age in month). The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older. ACSCs stand for Ambulatory Care Sensitive Conditions developed by AHRQ. See Appendix Table C for the list of ACSCs.

Figure 1.7: Age Profile of Inpatient Admissions (log scale)



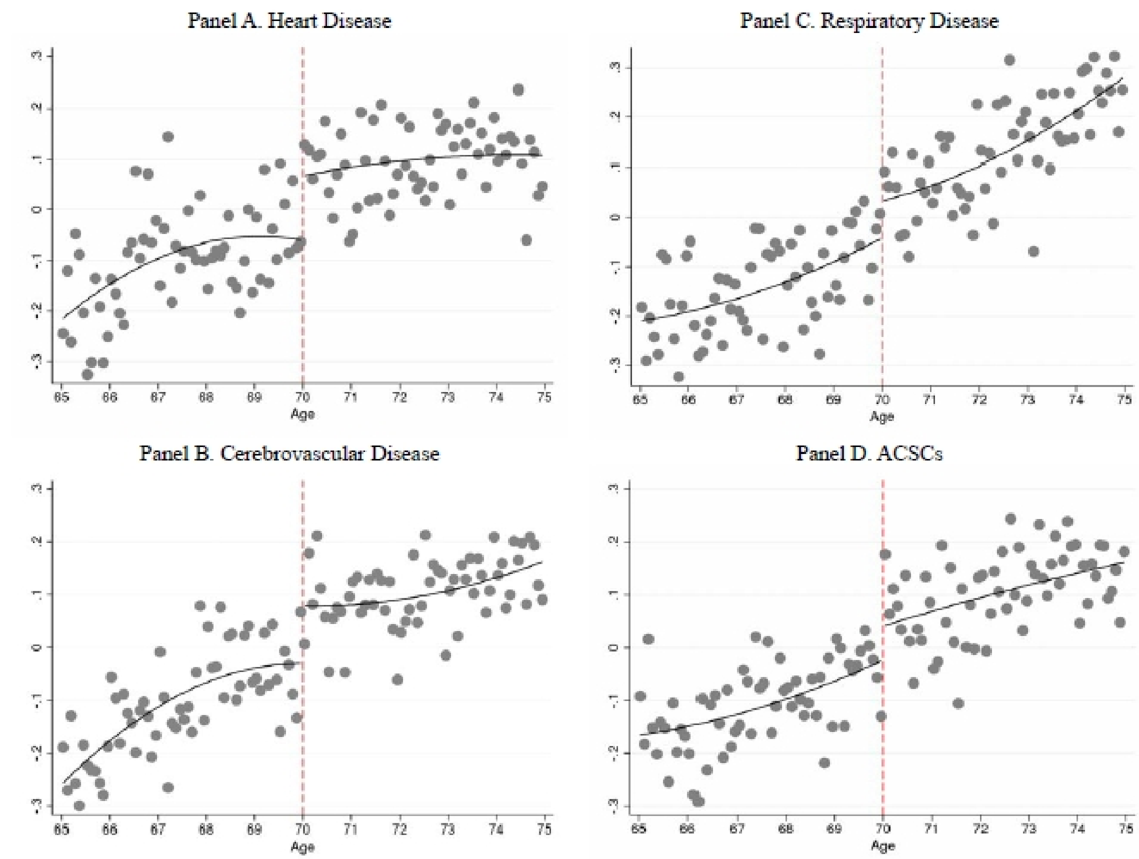
Note: The data come from pooled 1984-2008 discharge data in the Patient Survey. The markers represent the averages of residual from a regression of the log outcome on birth month fixed effects, admission month fixed effects and survey year fixed effects (aggregated by age in month). The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure 1.8: Age Profile of Inpatient Admissions with and without Surgery (log scale)



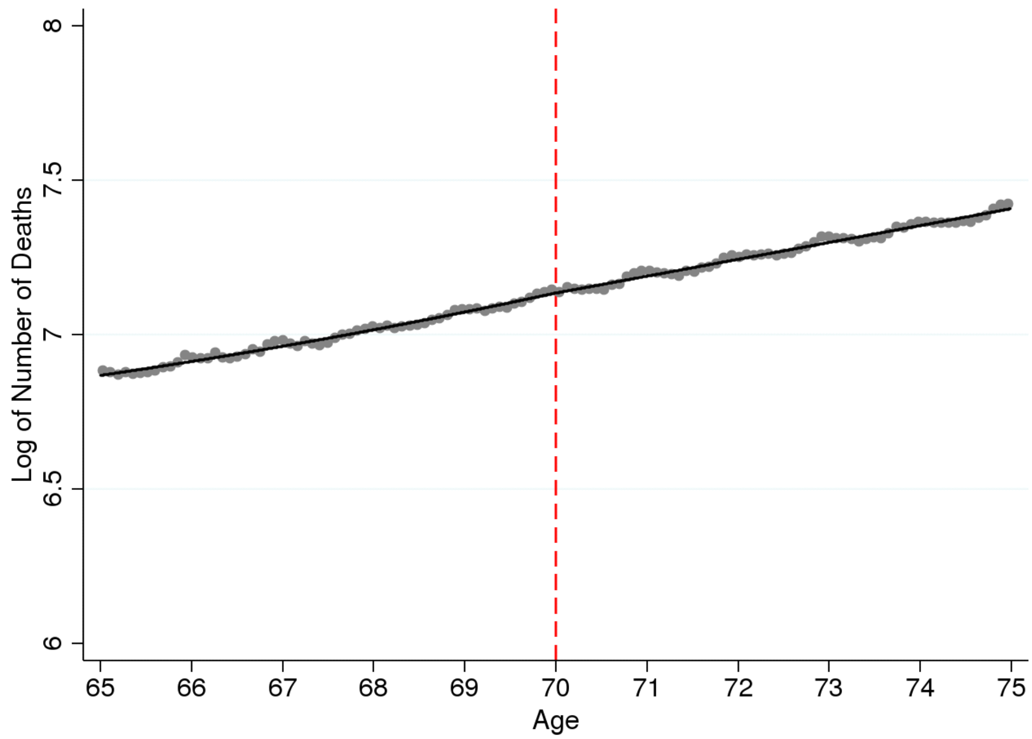
Note: The data come from pooled 1984-2008 discharge data in the Patient Survey. The markers represent the averages of residual from a regression of the log outcome on birth month fixed effects, admission month fixed effects and survey year fixed effects (aggregated by age in month). The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure 1.9: Age Profile of Inpatient Admissions for Selected Diagnosis (log scale)



Note: The data come from pooled 1984-2008 discharge data in the Patient Survey. The corresponding RD estimates at age 70 are statistically significant at 5 % level for Panel A and B only. The markers represent the averages of residual from a regression of the log outcome on birth month fixed effects, admission month fixed effects, admission of month fixed effects, and survey year fixed effects (aggregated by age in month). The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older. ACSCs stand for Ambulatory Care Sensitive Conditions developed by AHRQ.

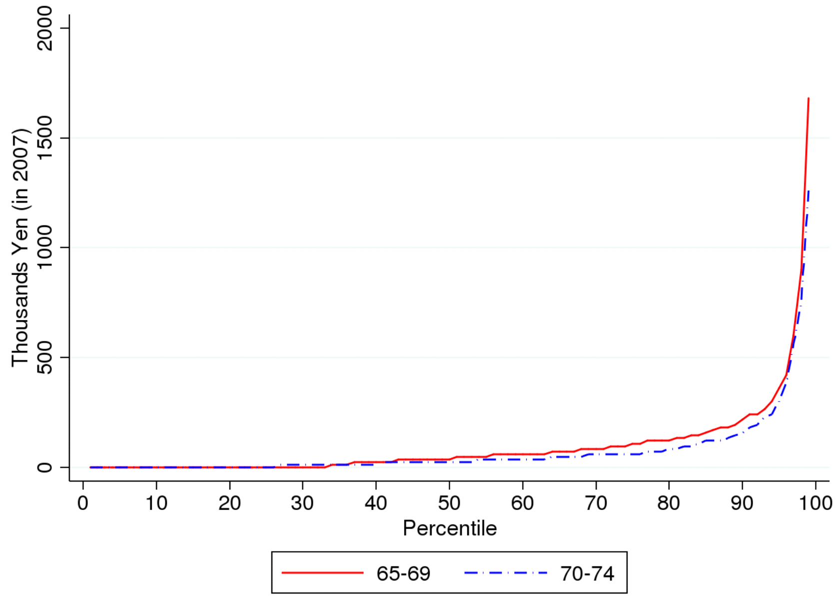
Figure 1.10: Age Profile of Overall Mortality



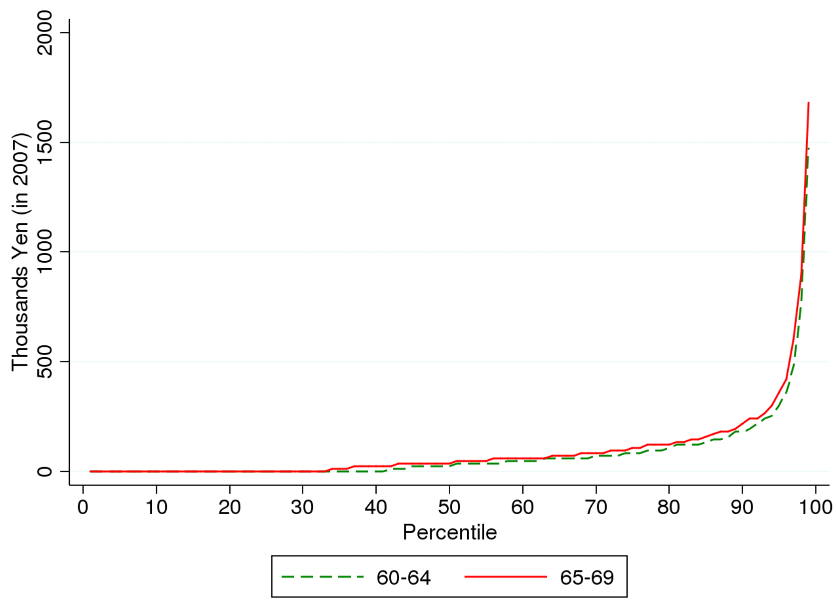
Note: The data come from pooled 1984-2008 mortality data. I use days to eligibility for the Elderly Health Insurance as a running variable. The cell is each 30 days interval from the day of eligibility at age 70. The markers represent the averages, and the lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure 1.11: Distribution of Out-of-Pocket Health Expenditure in 2007

Panel A. Ages 65-69 (Near Elderly) and Ages 70-74 (Elderly)

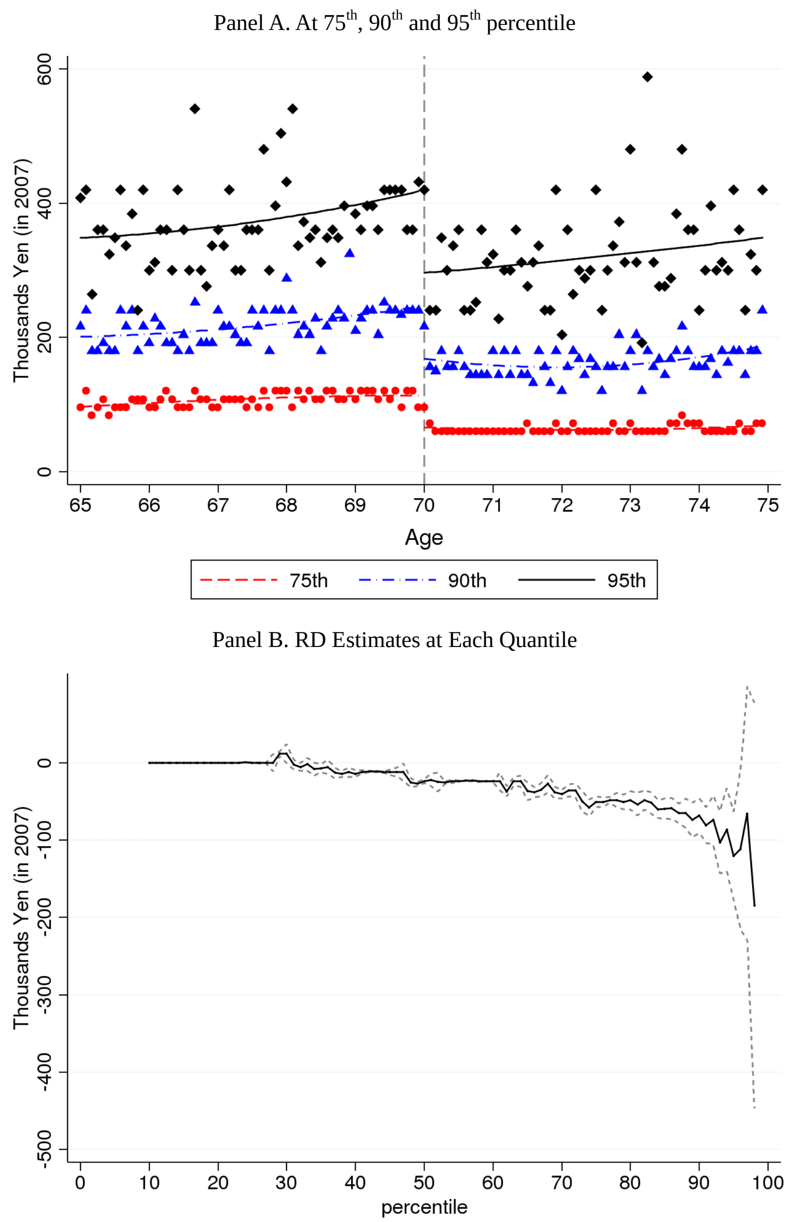


Panel B. Ages 60-64, and Ages 65-69 (Near Elderly)



Note: The data come from 2007 Comprehensive Survey of Living Conditions. I have multiplied the monthly out-of-pocket expenditures by twelve to convert to annual basis. One thousands Yen is roughly $10 US dollars.

Figure 1.12: Age Profile of Out-of-Pocket Medical Expenditures in 2007



Note: The data come from 2007 Comprehensive Survey of Living Conditions. I have multiplied the monthly out-of-pocket expenditures by twelve to convert to annual basis. One thousands Yen is roughly $10 US dollars. Panel A: The markers represent actual averages (age measured in month), and the lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older. Panel B: This figure plots the RD estimates at each quantile along with their 95 percent confidence interval. I do not show 99th percentile in the graph.

Table 1.1: Summary Statistics (Ages 65-75)

	Variables	Mean (SD)
A	Outpatient Data	
	Repeated Visits	0.94
	Hospital	0.44
	Clinic	0.56
	Male	0.42
	With Referral	0.05
	Days from Last Outpatient Visits (Days)	13.6
		(20.2)
B	Discharge Data	
	With Surgery	0.35
	Hospital	0.99
	Clinic	0.01
	Open-head surgery	0.00
	Open-heart surgery	0.01
	Open-stomach surgery	0.04
	Musculoskeletal surgery	0.03
	Endoscopic surgery: stomach	0.01
	Intraocular lens implantation	0.02
	Length of stay (Days)	18.1
		(17.7)
C	CSLC	
	Self Reported Health: Good or Better	0.31
	Being Stressed	0.41
	Male	0.45
	Currently Married	0.74
	Employed	0.31
	Hours of Work per Week	6.82
	Income (Thousands Yen)	1,860
		(1,920)
	Receiving Pension	0.95
	With Long Term Health Insurance	0.03

Note: One thousands Yen is roughly $10 US dollars.

Table 1.2: Formula for Cost-Sharing Below and Above Age 70

Panel A. Outpatient Visits

	Below 70				Above70	
	Coinsurance				Coinsurance	
Year	NHI	Employment-based (Employee)	Employment-based (Dep)	Stop-loss	All	Stop-loss
1984	30%[(1)]	10%	30%	51.0	0.4 /mon	-
1987	30%[(1)]	10%	30%	54.0	0.8 /mon	-
1990	30%[(1)]	10%	30%	57.0	0.8 /mon	-
1993	30%[(1)]	10%	30%	63.0	1.0 /mon	-
1996	30%[(1)]	10%	30%	63.0	1.02 /mon	-
1999	30%[(1)]	20%	30%	63.6	0.53 /day [(2)]	-
2002	30%[(1)]	20%	30%	63.6+(TC-318)*0.01	10%	12.0
2005	30%	30%	30%	72.3+(TC-241)*0.01	10%	12.0
2008	30%	30%	30%	80.1+(TC-267)*0.01	10%	12.0

Note: (1) Former employees pay 20% and dependent of former employees pay 30% among the retired (2) Up to 4 times/month. TC stands for total cost per month. All money values without percentage sign are in thousand Yen (roughly 10 US dollar in 2008).

Panel B. Inpatient Admissions

	Below 70				Above70	
	Coinsurance				Coinsurance	
Year	NHI	Employment-based (Employee)	Employment-based (Dep)	Stop-loss	All	Stop-loss
1984	30%[(1)]	10%	20%	51.0	0.4 /day [(2)]	-
1987	30%[(1)]	10%	20%	54.0	0.4 /day	-
1990	30%[(1)]	10%	20%	57.0	0.4 /day	-
1993	30%[(1)]	10%	20%	63.0	0.7 /day	-
1996	30%[(1)]	10%	20%	63.0	0.71 /day	-
1999	30%[(1)]	20%	20%	63.6	1.2 /day	-
2002	30%[(1)]	20%	20%	63.6+(TC-318)*0.01	10%	37.2
2005	30%	30%	30%	72.3+(TC-241)*0.01	10%	40.2
2008	30%	30%	30%	80.1+(TC-267)*0.01	10%	44.4

Note: (1) Former employees pay 20% and dependent of former employees also pay 20% among the retired (2) Up to 2 months. Also see the note above.

Table 1.3: Estimated Out-of-Pocket Medical Expenditure per Month

	Out of Pocket Medical Expenditure (thousand Yen)			% reached stop-loss among insurance claims	
	Below 70	Above70	% reduction	Below 70	Above70
Type of Service	(1)	(2)	((1)-(2))/(3)	(4)	(5)
Outpatient Visits					
	4.0	1.0	74%	0.1%	0.6%
Inpatient Admissions					
	38.0	12.4	67%	14.6%	0.0%

Note: All money values without percentage sign are in thousand Yen (roughly 10 US dollar in 2008).

Table 1.4: RD Estimates at Age 70 on Employment, and Family Structure

	All	By Gender: Male	By Gender: Female	Data: Years Available	Data: Sample Size for “All”
A. Employment related					
(1) Employed	0.3	0.5	0.1	1986-2007	573,104
	(0.4)	(0.5)	(0.5)		
(2) Retired	-0.1	0.8	-0.7	1986-2007	573,104
	(0.5)	(0.7)	(0.6)		
(3) Hours/wk	0.0	0.1	0.0	2004-2007	39,978
	(0.0)	(0.1)	(0.2)		
(4) Family Income (thousand Yen)	-54.9	-212.0	88.1	1986-2007	77,967
	(113.0)	(174.9)	(144.9)		
(5) Income (thousand Yen)	-32.3	-29.9	-34.1	2004-2007	18,757
	(89.8)	(179.9)	(54.3)		
B. Family Structure					
(6) Married Spouse Present	0.5	0.5	0.4	1986-2007	573,104
	(0.5)	(0.5)	(0.7)		
(7) Head of Household	0.0	-0.1	0.1	1986-2007	573,104
	(0.4)	(0.4)	(0.6)		
C. Other					
(8) Receiving Pension	0.3	0.2	0.4	1986-2007	573,104
	(0.3)	(0.4)	(0.4)		
(9) Long Term Care Insurance	-0.1	-0.5	0.2	2001-2007	232,928
	(0.3)	(0.4)	(0.3)		

Note: Estimated regression discontinuities at age 70 are shown, from models that include a quadratic of age, fully interacted with dummy for age 70 or older among people between ages 65-75. The exception is a pension dummy since there is a discrete jump at age 65 for probability of receiving the pension, and thus I limit the sample to age 66-74. Other controls include indicators for gender, region, marital status, birth month, and sample year. I use pooled samples of comprehensive survey of living condition (CSLC) conducted every three year since 1986. Sample sizes differ by variables since some variables are only collected for a shorter period. Note that income is collected for roughly 15 % of all samples. Standard errors (in parentheses) are clustered at the age in month level as this is the most refined version of the age variable available. All regressions are weighted to take into account the stratified sampling frame in the data. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table 1.5: RD Estimates at Age 70 on Outpatient Visits

A.	All	10.3***	F	By Diagnosis	
		(1.8)		Top 5	
B.	By Visit Type			Essential hypertension	8.0***
	First visits	12.7***			(2.4)
		(3.3)		Spondylosis	23.7***
	Repeated visits	10.3***			(3.6)
		(1.9)		Diabetes	1.7
C.	Days from Last Outpatients Visits				(4.4)
	Among Repeated Visits			Osteoarthrosis	25.3***
	1 day	17.9***			(4.2)
		(2.5)		Cataract	12.0**
	2-3 day	16.4***			(4.9)
		(4.4)		Other	
	4-7 day	13.3***		Heart disease	3.0
		(2.8)			(4.6)
	15-30 day	2.8		Cerebrovascular disease	15.2***
		(2.9)			(5.9)
	31-60 day	-1.5		Respiratory disease	14.3***
		(4.3)			(3.6)
D.	By Institution			Ambulatory Care Sensitive Conditions	8.2***
	Hospital	5.1**			(2.3)
		(2.0)		Cancer	6.1
	Clinic	13.8***			(8.0)
		(1.8)		Diseases of nervous and sense organs	10.4***
E.	By Referral				(2.8)
	Without Referral	10.5***		Diseases of genitourinary system	14.9***
		(1.9)			(5.4)
	With Referral	6.4		Diseases of skin	17.4***
		(5.2)			(4.9)
				Diseases of musculoskeletal system	18.6***
					(2.5)

Note: Each cell is the estimate from separate estimated regression discontinuities at age 70. The specification is a quadratic in age, fully interacted with dummy for age 70 or older among people between ages 65-75. Controls are dummies for each survey year and each month of birth. I use pooled samples of 1984-2008 Patient Survey conducted every three years since 1984. Sample size is 1080. Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table 1.6: RD Estimates at Age 70 on Inpatient Admissions

A	All	8.2***	Other		
		(2.6)		Heart disease	11.5**
B	Surgery				(5.7)
	W/o surgery	5.4*		Hypertensive disease	4.8
		(2.9)			(5.5)
	With surgery	10.8***		Ischemic heart disease	14.5**
		(3.8)			(7.1)
C	Type of Surgery			Cerebrovascular disease	10.5***
	Open-head surgery	11.7			(3.9)
		(8.8)		Intracerebral hemorrhage	8.0
	Open-heart surgery	4.1			(6.1)
		(8.5)		Cerebral infarction	12.8***
	Open-stomach surgery	11.4**			(4.6)
		(5.6)		Respiratory Diseases	6.8
	Musculoskeletal surgery	5.6			(4.8)
		(5.0)		Ambulatory Care Sensitive Conditions	7.6
	Endoscopic surgery: stomach	9.3			(5.0)
		(7.3)		Cancer	6.6
	Intraocular lens implantation	19.6***			(4.6)
		(6.2)	E	Location Before Admission	
D	By Diagnosis			Outpatients in Same Hospital	9.7***
	Top 5				(2.9)
	Cataract	22.6***		Other places	1.6
		(6.5)			(5.4)
	Angina pectoris	11.4			
		(7.3)			
	Occlusion of cerebral arteries	13.7***			
		(4.6)			
	Diabetes	7.4			
		(5.8)			
	Malignant neoplasm of stomach	4.9			
		(6.1)			

Note: Each cell is the estimate from separate estimated regression discontinuities at age 70. The specification is a quadratic in age, fully interacted with a dummy for age 70 or older among people between ages 65-75. Controls are dummies for each survey year, each month of birth, and each month of admission. I use pooled samples of 1984-2008 Patient Survey conducted every three year since 1984. Sample size is 3,240 except Panel C, and E. Sample size for C is 1,440 (4 yr, 1999-2008), and sample size for F is 1,800 (5 yrs, 1996-2008) since these information is only collected in the later years. Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table 1.7: RD Estimates at Age 70 on Mortality

		Basic	67-73 yrs	Cubic	LLR
		(1)	(2)	(3)	(4)
A	All	0.0	-0.3	-0.8**	-0.3
		(0.3)	(0.4)	(0.4)	(0.3)
B	By Diagnosis				
	Cancer	-0.5	-1.4***	-2.0***	-0.8
		(0.4)	(0.6)	(0.6)	(0.5)
	Heart disease	0.5	0.5	-0.7	0.1
		(0.8)	(1.0)	(1.0)	(0.9)
	Cerebrovascular disease	0.1	0.3	-0.1	0.3
		(0.8)	(1.1)	(1.2)	(1.0)
	Respiratory diseases	0.5	0.0	0.2	0.4
		(1.3)	(1.6)	(1.7)	(1.5)

Note: Each cell is the estimate from separate estimated regression discontinuities at age 70. The dependent variable is the log of the number of deaths that occurred x days from the person' eligibility to the Elderly Health Insurance See Data Appendix for the ICD codes for each of the categories above. I use pooled 1984-2008 mortality data. LLR (local liner regression) estimates use a triangular kernel and the rule-of-thumb bandwidth selection procedure suggested by Fan and Gijbels (1996). Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table 1.8: RD Estimates at Age 70 on Out-of-Pocket Medical Expenditure

	Out-of-Pocket Expenditure just Below age 70	RD Estimates at Age 70
	(1)	(2)
Mean	152	-52
40th Percentile	30	-14***
Median	52	-24***
60th Percentile	65	-24***
70th Percentile	96	-40***
80th Percentile	139	-49***
90th Percentile	247	-68***
95th Percentile	419	-115***
99th Percentile	1,793	-502*

Note: All money values are thousand Yen in 2007 (roughly 10 US dollar). I omit the 10, 20, and 30 percentile since the out-of-pocket expenditure is zero for those percentiles. ***, **, * denote significance at the 1%, 5% and 10% levels respectively.

Table 1.9: Welfare Gain from Risk Protection

		Distribution adjusted	
		using quantile estimates	"mechanically"
		(1)	(2)
A. At mean			
1. Risk Aversion			
(80% income cap)	1	3	7
	3	**20**	**46**
	5	41	110
2. Cap on percent of income			
(Risk aversion=3)	60	11	22
	90	31	74
B. Distribution			
(80% cap, risk aversion=3)			
	25th percentile	5	11
	Median	13	25
	75th percentile	31	85
	90th percentile	50	112
	95th percentile	63	126
	99th percentile	97	153

Note: All estimates are thousands Yen in year 2007. One thousands Yen is roughly 10 US dollars in 2007.

CHAPTER 2

Supply Induced Demand in Newborn Treatment : Evidence from Japan

with Kiyohide Fushimi

2.1. Introduction

Economists and policy makers have long argued that medical providers "induce" demand of health services by exploiting their informational advantage over patients and providing excessive care of dubious value (Evans 1974, Fuchs 1978, Pauly 1980, Rice 1983).[1] Since medical providers exert a strong influence over the quantity and types of medical care demanded, measuring the size of supply-induced demand (SID) has been a long-standing controversy in health economics (McGuire 2000). While there are numerous empirical studies on SID, they find surprisingly little evidence of SID; the estimated magnitudes are often insignificant or economically small.[2]

[1]See McGuire (2000) for summary of work on supply/physician-induced demand.

[2]For example, a recent study by Grant (2009) showed that a $1000 increase in the reimbursement for performing a Caesarean section would increase the Caesarean section rate by little more than one percentage point.

However, these past studies may underestimate the size of SID for two reasons: First, it is empirically difficult to isolate SID from other confounding hospital behaviors, such as changes in the selection of patients (Ellis and McGuire 1996). Estimates of SID will be biased towards zero if hospitals select unobservably healthier patients for a given treatment intensity. Since it is difficult to control for the severity of patients' conditions, selection bias poses an important empirical challenge in this literature.[3] Second, most of the past literature focuses on medical procedures that carry large risks for both physicians and patients, such as Caesarean sections (Gruber and Owing 1996, Grant 2009) or coronary artery bypass graft surgeries (Yip 1998). SID may be less likely for these high-risk procedures, since physicians face a higher probability of lawsuits if they perform them excessively and must persuade patients to consent.

We overcome these empirical challenges by focusing on specific population: at-risk newborns. Selection is less of a concern for the treatment of newborns, especially low birth weight infants, because the birth weight and severity of the newborns' conditions are difficult if not impossible to predict in advance (Almond et al., 2010), even though we later show suggestive evidence of slight birth weight manipulation. In addition, newborn treatment allows substantial room for demand

[3]One notable exception which suffers less from selection bias is Gruber and Owing (1996); they use a decline in fertility as an income shock, and find that within-state declines in fertility increase within-state Caesarean section rates, since Caesarean sections are more lucrative than normal vaginal deliveries. However, the magnitude is very small; a 10 % fertility drop corresponds to only a 0.97% increase in the probability of a Caesarean section. This increase accounts for only 0.5% of physician's income.

inducement, since the informational advantages of physicians over patients are arguably among the largest.[4] We also focus on a less risky medical procedure: Neonatal Intensive Care Unit (NICU) utilization. NICU utilization is of particular interest since it contains minimal risks for both patients and physicians. Additional days in the NICU do not harm newborns and may even benefit them.[5]

There are two key institutional features that make Japan a nice setting for estimating supply-induced demand of NICU utilization. First, Japan introduced a *partial* prospective payment system (PPS), which made NICU utilization *relatively* more profitable than other procedures, since it was excluded from the per-diem prospective payment and was fully reimbursed. Since hospitals adopted the PPS at different times, we use a difference in difference framework to estimate the effect of relative changes in price on demand. Second, because NICU utilization is costly, the government caps the number of NICU days for which hospitals are reimbursed. Lighter births are allowed longer stays, and the cap changes discontinuously at the birth weight cut-offs of 1000 and 1500 grams.[6] The jump in the cap

[4]Mothers have almost no choice but to conform to directions by physicians, unlike the cases of other common diseases for which patients may have more medical knowledge.

[5]We are not aware of any medical evidence that time in NICU harms infants who do not need to stay in NICU. In fact, NICU is believed to be one of the technological developments that contributed to a decline in infant mortality (Phibbs et al. 2007). The others technological developments include such as pulmonary surfactant replacement therapy, and high-frequency oscillatory ventilation.

[6]Hospitals receive nearly 85,000 Yen (roughly US$944) for each day an infant stays in the NICU. Since infants with birth weights less than 1500 grams stay an average of 43 days in the NICU, hospitals receive an average reimbursement of $40,600 for the NICU utilization of these newborns.

means that hospitals' scope for increasing NICU utilization and their incentives for manipulating birth weights substantially differ by the range of infants' birth weights. Adoption of the PPS, which increases the relative profitability of NICU utilization, only increases their financial incentives for gaming.

Our focus on at-risk newborns and less risky medical procedures uncovers strong evidence of supply-induced demand. First, we find evidence that hospitals manipulate reported birth weights; there is an increased mass of birth weights that are reported just below the cut-offs of 1000 and 1500 grams, which only occurs in hospitals with NICUs and is exacerbated after the introduction of PPS. We run the density test proposed by McCrary (2009) and find that there is statistically significant heaping at just below these cut-offs after the adoption of PPS. We do not have any objective measure of newborns' health besides mortality and cannot link mothers' information to the birth data, so it is difficult to distinguish whether this sorting is the result of benevolence (e.g., physicians mis-recording the birth weights of sicker infants who weigh more than the cut-off so that these infants receive necessary treatment) or gaming (e.g., hospitals mis-recording birth weights to obtain higher reimbursements for NICU utilization). However, since we see exacerbated manipulation after the introduction of PPS, we suspect it is more likely to be gaming.

Second, we find that the hospitals increase NICU utilization in response to the adoption of PPS. Most interestingly, we only find the increase in NICU stays among Very Low Birth Weight infants (VLBW; birth weighing less than 1500 grams) infants, for whom there is more scope for increasing utilization. In fact, we find that after the adoption of PPS, NICU stays of VLBW newborns increase by 4.8 days, an 11.2 % increase relative to NICU stays before the adoption of PPS. This result is robust to a variety of the robustness checks, such as the inclusion of a lead dummy and hospital-specific linear trends. We also rule out alternative explanations than SID. For example, we find no evidence that the characteristics of low-birth weight babies changed or transfers of newborns increased after PPS was introduced.

Finally, there is also little evidence that the induced increase in NICU stay reduced infant mortality, suggesting that the marginal increase in NICU utilization had little impact on newborns' health. The increase in NICU stay translates into an increase in hospitals' reimbursement by roughly 489,000 Yen ($5,400) per VLBW newborn. This increased reimbursement for NICU utilization can result in an additional medical expenditure of as much as 10.6 trillion Yen ($117 million), without any observable improvements in short-term infant health outcomes.[7]

In addition to the literature on supply-induced demand, our paper also contributes to the literature on hospital gaming. Dafny (2005) divides hospital responses to price changes into two categories: *nominal* and *real* responses. Nominal responses correspond to accounting maneuvers (e.g., upcoding diagnoses) while

[7]All figures in the dollar term are measured in 2009 US dollar throughout this paper. All price in Yen is deflated by CPI to 2009 Japanese Yen, and then converted to US dollars by the exchange rate of 90 Yen per US dollar.

real responses correspond to actual increases in the provision of care. Unlike Dafny (2005), which finds evidence of only nominal responses but not real responses to changes in diagnosis-specific prices, we find evidence of both nominal (i.e., manipulation of birth weight) and real (i.e., longer stays in NICU) responses.

Our results can also inform the reimbursement policies for newborn treatment in other countries. Since birth weight is believed to be the best predictor for the treatment intensity of newborns, other countries are increasingly using it to determine the reimbursement level for newborn treatment (Quinn 2008). For example, a few states in the US have already incorporated birth weights into the Diagnosis Related Group (DRG) reimbursement schedules for state Medicaid programs.[8] Our results caution against the use of the birth weights in the reimbursement system, since hospital gaming may become severe under higher stakes. Newborn treatment is also expensive in other countries, and some have questioned the effectiveness of these increasingly intensive treatments (Grumbach, 2002; Goodman et al., 2002) while others have argued that benefits outweigh costs (Cutler and Meara 2000; Almond et al 2010). Our results may suggest that we have reached the "flat-of-curve medicine" in newborn treatment in Japan, where one of the lowest infant

[8]Modified versions of original DRG, such as All Patient DRGs (AP-DRGs), and All Patient Refined DRGs (APR-DRGs) incorporate birth weight in their groupings of diagnoses for reimbursement. For example, the AP-DRGs is used in DC, GA, IN, NY, VA, WA, and the APR-DRGs is used in MD. See Quinn (2008) for details.

mortality rates in the world has been achieved at a relatively low cost (Ikegami and Campbell 1995).[9]

Japan offers a nice empirical setting to examine the existence and size of SID for a number of reasons. First, under universal health insurance, medical providers in Japan are all paid through the same national fee schedule, which is uniformly applied regardless of patients' insurance type. Thus, we can easily measure the monetary size of the SID. Second, there is a little room for cost-shifting in Japan because all citizens are covered by the mandatory universal health insurance.[10] In contrast, in the US, the introduction of DRG/PPS on Medicare led hospitals to charge higher prices for private insurers (Cutler 1998).[11] Third, physicians' incentives are more likely to align with hospitals' incentives since in Japan, physicians in the hospitals are all employed by the hospitals, unlike in the US, where the physicians and hospitals are separate entities. This fact is important to detect SID, since for SID to occur, physicians must be willing to provide excess care, and the administrators must submit these claims for payment.

[9]For example, the life expectancies at birth in Japan and the US are 82.6 and 78.1 in 2006, and the infant mortality rates (deaths per 1,000 live births) in Japan and US are 2.6 and 6.7, receptively (OECD, 2009). On the other hand, the ratio of medical expenditures to GDP in Japan is lowest among the OECD countries. In 2006, the ratio of total expenditures on health to GDP was 8.1 in Japan, 11.0 in France, 10.5 in Germany, 9.0 in Italy, 8.5 in the United Kingdom, and 15.8 in the United States.

[10]Japan achieved universal health insurance policy in 1961. See Kondo and Shigeoka (2011) for more details about introduction of universal health insurance and its impact.

[11]See also McGuire and Pauly (1991) for economic model of physician behavior with multiple payers. This model can be viewed as single payer buying multiple services as in our case.

The rest of the paper is as follows. Section 2.2 describes background on the reimbursement system and the treatment of newborns in Japan. Section 2.3 describes the data and Section 2.4 presents the identification strategy. Section 2.5 shows the birth weight distribution and discusses manipulation of birth weight. Section 2.6 shows the main results on NICU utilization, and Section 2.7 examines health outcomes, and measures the monetary size of the induced demand. Section 2.8 concludes.

2.2. Background

In this section, we briefly describe the reimbursement system and the treatment of newborns in Japan.

2.2.1. Reimbursement system in Japan: FFS to partial PPS

Before the introduction of the PPS, the medical providers in Japan were all paid by a fee-for-service system (FFS). The national fee schedule for procedures was uniformly applied to all Japanese patients, regardless of their insurance type and medical providers.[12] However, medical expenditures in Japan have been rising, largely due to the aging, at a faster rate than any other developed countries.

[12] See Ikegami (1991, 1992) and Ikegami and Campbell (1995) for detailed descriptions of the medical system before the implementation of PPS in Japan. The national schedule is biennially revised by the Ministry of Health, Labor and Welfare (MHLW) through negotiation with the Central Social Insurance Medical Council (CSIMC), which includes representatives of the public, payers, and providers.

To contain the rising medical expenditures, the Japanese government implemented its unique PPS, which bases payment on the patient's Diagnosis Procedure Combination (DPC) (similar to PPS based on Diagnostic Related Group (DRG) in the US), and *partially* replaces the conventional FFS.[13] The PPS based on DPC (DPC/PPS hereafter), is designed as a way of reimbursement for hospitals for acute inpatient care. The government started it in 82 hospitals, mostly university hospitals in April 2003. Since this new payment system is revenue-neutral for each hospital for the time being, it has expanded at a rapid rate to most acute hospitals, even though the participation in the PPS was only mandatory for the first 82 hospitals.[14] Therefore, one potential concern is that the adoption of PPS is endogenous to outcome of interest. While we show that predetermined hospital characteristics explain little of the variation in the timing of adoption, to assuage this concern, we include interactions of these hospital characteristics with time trends in all our regressions to control for differences in trends across hospitals, similar to Acemoglu, Autor and Lyle (2004). We also conducted a variety of robustness checks to account for concerns about potential endogenity of the adoption such as the inclusion

[13]There is extensive empirical literature on the effect of the introduction of PPS for Medicare beneficiaries in the US. See Coulam and Gaumer (1991) and Cutler and Zeckhauser (2000) for reviews.

[14]More precisely, the per-diem fixed payment is multiplied by a hospital update factor, which is unique to each hospital (See Okamura et al. (2005) for details). This hospital update factor is calculated so that the hospital may receive the same revenues as in a prior year as long as hospitals see the same case-mix of patients as a year before. Hospitals were afraid that failure to adopt PPS could jeopardize their status as acute care hospitals. Acute hospitals are considered more advanced and prestigious than chronic disease hospitals, and thus, they attract more patients in Japan. In fact, some hospitals publicize on their websites that they are reimbursed by DPC/PPS.

of a lead dummy and hospital-specific linear trends. We discuss this more in detail in the estimation section.

2.2.2. The Hospital fee and doctor fee

A unique feature of the health care system in Japan is that the physicians who work in hospitals are employed by the hospitals, unlike in the US. Therefore, when the government designed the PPS, it divided medical procedures into two categories, which are referred to the "hospital fee" and "doctor fee." The procedures considered under the "hospital fee" (hereafter, hospital-fee procedures) are paid under a per-diem prospective payment, while procedures considered under the "doctor fee" (hereafter, doctor-fee procedures) are paid by the conventional fee-for-service system.[15] The former includes medical procedures that are relatively standardized across hospitals, such as bed use, diagnostic imaging, injections, and medications. Procedures that reflect technical work by physicians are considered doctor-fee procedures, a major component of which are surgeries.[16] The idea behind this distinction is that hospitals could easily reduce medication expenditures by replacing brand names with generics, but the avoidance of the necessary surgeries may result in huge adverse outcomes.

[15]Because of the per-diem instead of per-admission payment and assignment of the DPC based on types of surgeries, and medication, in addition to diagnosis, this payment system is not completely prospective. But the retrospective nature of diagnosis classifications is also applied to DRG in the US (Zweifel et al, 2009).

[16]Relatively complicated, technological procedures such as endoscopic inspection and anesthesia are also exempted from the per-diem fixed payment.

For newborn treatment, in addition to surgeries, one additional procedure is excluded from the per-diem fixed payment: NICU utilization. NICU utilization is excluded from the hospital-fee procedures since it requires a substantial workload by physicians and because there is a concern that reducing NICU utilization could have adverse effects on at-risk newborns.

This partial PPS can substantially affect the behavior of the hospitals since it makes the doctor-fee procedures relatively more lucrative than the hospital-fee procedures. This substitution effect is substantial since while hospitals are still fully reimbursed for the former, while hospitals need to bear any additional costs incurred for medical treatments for latter. In sum, this partial PPS gives hospitals the financial incentives to perform the doctor-fee procedures, including NICU, intensively while reducing hospital-fee procedures if possible.

2.2.3. Newborns treatment in Japan

A Neonatal Intensive Care Unit (NICU) is a hospital unit that specializes in the care of premature, low birth weight or severely ill newborns. They are developed to provide better temperature and respiratory support, isolation from infection risks, and specialized feeding for vulnerable newborns.[17] The development of NICU has been thought to be one of the main contributors to the decline in death rate of

[17]The Ministry of Health, Labor and Welfare (MHLW) establish requirements for hospitals that claim the reimbursement for NICU utilization. For example, these hospitals must have at least one neonatologist for all day, and possess emergency resuscitation equipment (endotracheal intubation set), a cardio-respiratory monitor, artificial ventilation for infants, micro-infusion device, pulse oximeter, and photoradiation therapy equipment.

at-risk newborns (Lee et al. 1980; Kliegman 1995; Phibbs et al. 1996; and Phibbs et al. 2007).

However, NICU utilization is very costly in Japan as well as other countries. Hospitals are reimbursed 85,000 Yen ($944) for each additional night in the NICU.[18] Since the average NICU stay of newborns with birth weights less than 1500 gram is 43 days, hospitals are reimbursed as much as 3,655,000 Yen ($40,600) per newborn for NICU utilization alone. Since mothers of low birth weight infants (specifically those less than 2000 grams) are mostly exempted from payment under the national policy in Japan, this large charge is almost all paid by the government.

The government acknowledges its concern over over-utilization of NICU. Thus the maximum number of the days that hospital can claim the reimbursement for NICU utilization is set by the birth weight, since birth weight is believed to be the best predictor of requiring NICU utilization. Specifically, these limits are 21 days for newborns above 1500 grams, 60 days for those between 1000 and 1500 grams, and 90 days for those less than 1000 grams. The jump in number of the maximum days is important in our setting since the room for additional claims for NICU utilization substantially differs by the range of birth weight.

Figure 2.1 shows a histogram of the number of days in NICU for each birth weight range (above 1500 grams, between 1000 and 1500 grams, and below 1000 grams, respectively) before the PPS is introduced. Two things are noticeable.

[18]The amount of reimbursement slightly differs by the characteristics of the hospitals but they are very similar.

First, for any birth weight range, there is some bunching at the maximum days that the hospitals can claim the reimbursement for NICU utilization. Second and more importantly, infants with birth weights over 1500 grams have the most bunching at the maximum days. This implies that there is less room for longer claims of NICU for birth weights more than 1500 grams than for birth weights less than 1500 grams.

2.3. Data

2.3.1. Description and sample selection

The main data are the insurance claim data for in-hospital births that are delivered and discharged between April and December 2004-2008.[19] This is the first paper in economics to use this data.[20] Since hospitals that were not acute or chronic care hospitals wanted to join this new payment system, which was designed for the acute care hospitals, the government set an eligibility criterion for hospitals joining after 2006: they were required to submit data for the two years of data prior to joining. Thus, there is no pre-PPS data for the hospitals that adopted PPS before 2004 in our data.

[19]Exception is year 2004 and 2005. For 2004 and 2005, the data was collected from April to October. As a robustness check, we limit the sample to the birth between April and October to be consistent across years, but the main results are quantitatively unchanged. Data submission is only required for these months in early years to reduce hospitals' burden of compiling the data. For the fiscal year of 2010, hospitals had to submit the whole year of data.

[20]This data set has previously been exclusively used for medical research (e.g., Kuwabara and Fushimi (2010)).

Because the national fee schedule sets uniform prices for each procedure, Japanese insurance claim data includes price information for each procedure, and, therefore, we are able to measure the monetary size of any inducement. This is different than the US, where the payment methods used to reimburse hospitals are notoriously complex and frequently incomplete.

We extract the data in the following manner: First, we extract in-hospital births for the 188 hospitals that claimed at least one day of NICU utilization.[21] Second, we merged pre-treatment hospital information from 2002, and dropped the one hospital for which this information was missing, since it opened after 2002. Finally, we limit the sample to the births weighing less than 2000 grams for following reasons: First, under national policy, mothers of infants weighing less than 2000 grams are mostly exempted from payment of newborns treatment, so there are no incentives for mothers to limit over-utilization. Second, we only observe births that are covered by health insurance in our data; while all births weighing less than 2000 grams are covered by health insurance, the only births weighing over 2000 grams that are covered by health insurance are those with severe complications.[22] Thus, births weighing less than 2000 grams in our data

[21]We do not have information on which hospitals have NICUs, so the hospitals that claim at least one day for NICU utilization are regarded as the hospitals with NICU beds. Among 187 hospitals with NICU beds, there are 15 hospitals that have started or stopped claiming NICU utilization during our sample period, and we examine whether the inclusion of these hospitals affects the main results in a robustness check.

[22]While normal vaginal delivery is not covered by health insurance, mothers receive a lump-sum payment set by the government from insurers for birth.

are a more nationally representative sample of births (conditional on weighing less than 2000 grams). Indeed the birth weight distribution for birth weights less than 2000 grams in our sample is very similar to that nationally.[23]

The final sample size for the 187 hospitals we study is 13,408, which is 85.3% of the total number of 15,725 births weighing less than 2000 grams. One concern may be whether births are transferred from non-NICU hospitals. However, only eight percent of births in non-NICU hospitals are transferred to these NICU hospitals, and we also confirmed that the number of transfers from non-NICU hospitals to NICU hospitals did not changed after PPS was introduced.[24]

2.3.2. Outcome variables

The three key outcome variables for NICU utilization are a NICU utilization dummy, which equals one if the hospital claims at least one day for NICU utilization for the newborn; a variable for the number of NICU days claimed, conditional on NICU utilization; and a dummy for whether the newborn reached the

[23]The birth distribution among all births below 2000 gram in 2008 is 14% (less than 1000 gram), 24% (1000-1500 gram), and 63% (1500-2000 gram) according to the national statistics (Ministry of Health, Labor and Welfare 2009). In our sample, the corresponding figures are 13%, 25%, and 62%, which is almost identical to the national statistics. The birth distribution in our data deviates from national statistics above 2000 gram since some of the births in this range are treated as the normal deliveries.

[24]Specifically, we regressed the number of transfers received using equation (2.1) and found no statistically significant change.

maximum number of NICU days allowed for his birth weight.[25] The first variable corresponds to the extensive margin, and the second corresponds to the intensive margin of NICU utilization. For health outcomes, we create a dummy variable for whether the infant dies within 7 days, 28 days, and 90 days. However, the infant mortality rate is quite low in Japan and also because of the small sample size, these health outcomes are not precisely estimated.[26]

Table 2.1 summarizes the key variables in the data. The summary statistics are grouped by the year that hospitals adopted the new payment system. The simple comparison before and after adoption of the PPS for the NICU days shows that hospitals that adopted the new payment system in 2006 and 2008 (hereafter referred to treatment hospitals) increase NICU days by 1.4 days. The NICU days for these hospitals are similar to hospitals that adopted in 2003 and 2004, and slightly shorter than hospitals that adopted in 2009. There is not much difference in total length of stay in hospitals before and after the implementation of PPS among treatment hospitals. The table also shows that roughly one third of the reimbursement comes from hospital-fee procedures and the two thirds come from the doctor-fee procedures, which mainly come from the NICU utilization.

[25] In the US, a NICU utilization variable was just recently added to new birth certificate, which was phased in some states beginning in 2003. Also, HCUP data records the number of days in the NICU for Maryland.

[26] Also we cannot track post-discharge mortality.

Figure 2.2A is the graphical presentation of the regression analysis of NICU utilization. We plot the average length of stay in NICU for newborns which stay at least one day in NICU for each 100 gram interval before and after the PPS. To avoid a composition effect caused by hospitals' differential timing of PPS, we only use one year before and one year after the adoption of PPS for hospitals for which we have both pre and post PPS data. There are two things worthwhile to mention. First, the number of days in NICU differs substantially at the birth cut-offs both pre and post PPS. For example, before PPS is introduced, the jump in NICU days at 1500 grams is roughly 15 days between infants below 1500 grams and above 1500 grams, and that at 1000 grams is a similar magnitude. Due mainly to this increase in NICU days at the birth cut-offs, the total reimbursement measured by national fee schedule also jumps discontinuously at these birth cut-offs in Figure 2.2B.[27] Second, we find a sizable increase in NICU days only for births weighing less than 1500 grams after PPS is adopted. This result is consistent with Figure 2.1 because there is more room for longer claims below 1500 grams, since many of the births above 1500 grams already reach the maximum days.

[27]However these charges are much cheaper than those in the US. Comparing with Figure 3A in Almond et al. (2010), the charges in Japan are less than half of those in the US. Since my figure includes only complicated births which need to stay at least one day in NICU, the gap between Japan and the U.S. should be even larger.

2.4. Estimation

2.4.1. Estimation equation

Since the PPS is introduced at times at the hospital level, I use a difference-in-difference strategy to estimate the effect of PPS on the supply of medical procedures:

$$Y_{iht} = \theta_t + \alpha_h + X_{iht}\beta + Post_{ht}\phi + \epsilon_{iht} \tag{2.1}$$

for newborn i, hospital h, at time t. Y_{iht} is the outcome, such as NICU days, mortality and total reimbursement. θ_t represents a full set of year dummies, and α_h stands for a full set of the hospital fixed effects. X_{iht} is a vector of the newborn characteristics such as birth weight, gestational length and gender. $Post_{ht}$ is a dummy that equals one if hospital h is under the new payment system at time t. Finally, ϵ_{iht} is a random term that captures all omitted variables. The main coefficient of interest is ϕ.

There are five different hospital groups in the data that adopted the new payment system at different times, specifically in 2003, 2004, 2006, 2008, and 2009. Since our data span 2004-2008, the post dummy is always one for hospitals that adopted PPS in 2003 or 2004, and always zero for those that adopted PPS in 2009. Therefore, the identifying variation comes from hospitals that adopted PPS in 2006 or 2008. We cluster the standard error at the hospital level in all specifications to allow for an arbitrary serial correlation within hospitals (Bertrand, Duflo, Mullainathan 2004).

2.4.2. Adoption of PPS

One potential concern is whether the adoption of PPS is exogenous. Participation in the PPS was only mandatory for the first 82 hospitals, mainly university hospitals, that adopted PPS in 2003. Many hospitals followed because the DPC/PPS is revenue-neutral for each hospital. In fact, as of 2009, it has expanded to 1,428 hospitals. Therefore, one potential concern is that participation to the PPS is endogenous with the use of NICU. But it is important to note that if hospitals want to exploit the revenue-neutral nature of the PPS, those hospitals should increase treatment intensity in the year prior to the adoption of the PPS, since this is the base year for which the hospital update factor to guarantee the previous year's revenue is calculated. However, these strategic behaviors would make it more difficult for us to find a result, since hospitals may reduce treatment intensity to increase the profit once PPS is adopted. Also, the hospital update factor is computed based on the hospital-fee procedures only, which do not include NICU utilization.

Nonetheless, anecdotal evidence suggests that government hospitals tended to adopt later, since they often needed approvals from municipal legislature. Table 2.1 shows the hazard of year until adoption of PPS regressed on variety of hospital characteristics from 2002, one year before the implementation of PPS in the first round of 82 hospitals.[28] Consistent with the anecdotal evidence, the governmental hospitals tend to adopt the PPS slower than the non-profit hospitals, and hospitals with fewer beds also tend to implement later.[29] However, the other hospital characteristics explain very little of the variation in the timing of adoption. We view the weakness of this model fit as encouraging to our identification strategy. Nonetheless, in order to control for possible differences in trends across hospitals that are spuriously correlated with the post dummy, all of our regressions include interactions of these 2002 pre-determined hospital characteristics with time trends following Acemoglu, Autor and Lyle (2004), Hoynes and Schanzenbach (2009), and Almond, Hoynes and Schanzenbach (2011). In practice, our results are not sensitive at all to inclusion of these controls. We also conducted a variety of robustness checks to account for concerns about potential endogenity of the adoption such as the inclusion of a lead dummy and hospital-specific linear trends.

2.5. Manipulation of Reported Birth Weight

2.5.1. Distribution of Birth Weight

Figure 2.3 plots the distribution of reported birth weight for 800-2000 gram in the hospitals with NICU beds and without NICU beds. To see the contrast between

[28]The hospitals that were required to adopt PPS in 2003 are excluded from this analysis.

[29]There are no private for-profit hospitals in Japan, since hospitals are not allowed to issue the shares in Japan.

before and after the PPS is introduced, we limit the sample to treatment hospitals, since they have data both before and after the adoption of PPS. Due to the small sample size, we aggregate the frequency within 20 gram intervals for NICU hospitals and 50 grams for non NICU hospitals due to small sample size of the latter group. The two vertical lines correspond to 1000 grams and 1500 grams, where the number of the days that hospitals can claim reimbursement for NICU utilization jumps substantially.

Figure 2.3A shows that there is clearly heaping just below 1000 and 1500 gram cut-offs for both before and after PPS among NICU hospitals, but larger heaping after PPS. We observe the heaping just *left* of 1000 gram and 1500 gram threshold while we observe the heaping just *right* of most of every other 100 grams threshold due to the rounded reporting at every 100 grams, which are included in the *right* bin of these thresholds in the histogram.

More formally, we run the density test proposed by McCrary (2009) using the birth weight as a running variable. We find that there is statistically significant jump at both 1000 gram and 1500 gram after PPS. Figure 2.4 shows the result of the McCrary's test for post PPS among NICU hospitals. We use the pilot bandwidth of 100 gram with the binsize of 10 gram. The log difference in distribution at 1500 grams is -0.84 (t= -2.68), and 1000 grams is -0.45 (t= -2.28) for post PPS. Table 2.3 shows that these results are robust to the different binsize and bandwidth choices. Also we do not see any statistically significant jumps at any of other multiples of 100 grams. These results show that heaping at just below 1000 gram and 1500 gram is not driven mechanically at round numbers as a result of common scale resolutions. For pre PPS, even though we visually see slight heaping, it is not statistically significant at either of the two cutoffs.

2.5.2. Manipulation?

We argue that this heaping is indeed the result of the manipulation of reported birth weight for following reasons. First, Figure 2.3B shows that such heaping is not observed among non NICU hospitals. Second, since we focus only on in-hospital births, this result is not driven by receiving transfers from other hospitals that are just below the birth weight cutoffs or sending transfers to other hospitals that weigh slightly more than the birth weight cutoffs. Third, our results are not driven by uniqueness of our insurance claim data since we find the same heaping among all births in Japan. We obtained the universe of births in Japan for 1995, 2000, and 2005 from the vital statistics. In Figure B.1 in Appendix, we see the obvious heaps even among universe of births for any year of data.[30] The results of McCrary's density test on this data are summarized in Table B.1 in Appendix.[31]

This manipulation *per se* is of particular interest since this result is very different from Almond et al. (2010), which did not find such a sorting at 1500 gram birth

[30]Unfortunately, vital statistics don't have hospital information as well as mother's socioeconomics status to examine the characteristics of the sorting.

[31]We do not observe any statistically significant or economically large heaping at just below 2000 grams for the universe of births either (available from the authors upon request).

cut-off in the US birth data, while they found pronounced reporting heaps at the gram equivalents of one ounce intervals. A recent paper by Barreca et al. (2010) shows that this heaping in birth weight in US data is found to be systematically correlated with socio-economic characteristics. On the other hand, Bharadwaj and Neilson (2011) show that heaping at the 10, 50 and 100 gram intervals in birth data from Chile, which, like Japan, measures birth weights in grams, are not correlated with mother's characteristics.[32]

Since we do not have any objective measures of newborns' health besides mortality[33], and cannot link mothers' information to the birth data, it is difficult to distinguish whether this sorting is the result of benevolence (e.g., physicians mis-recording the birth weights of sicker infants which weigh more than the cut-off

[32]Indeed, we could also potentially use the jump in the NICU days at these birth weight thresholds to examine the effect of NICU on the health outcomes in the RD framework akin to Almond et al. (2010) and Bharadwaj and Neilson (2011). However, if the unobserved quality of the hospitals or physicians is correlated with the manipulation of birth weight, this may violate the identification assumption of RD that birth below and above the cutoff is random. Also more importantly, even without the issue of manipulation, the mortality rate, which is the best objective health outcomes available in our data, is very low in Japan and also the sample size is not large enough to precisely estimate its effect in the RD framework. Therefore we do not seek the approach in this paper.

[33]For example, the diagnosis may be endogenous. Since they are coded by the physicians, they are also potentially driven by the same economic factors determining the NICU utilization. If physicians are going to manipulate the birth weight to reap financial reward, they must indicate a diagnosis that justifies the use of this expensive unit. Many studies have documented such "coding" within the context of Medicare's PPS (Dafny 2005, Silverman and Skinner 2004).

so that these infants receive necessary treatment) or gaming (e.g., hospitals mis-recording birth weights to obtain higher reimbursements for NICU utilization).[34] However, since we see exacerbated manipulation after the introduction of PPS, we suspect the latter story fits better in this setting.

It is important to note, however, that the degree of manipulation does not seem substantial considering the size of the financial reward that hospital can reap.[35] For example, if the hospitals manipulate birth weights that are just above 1500 grams to just below 1500 grams, the maximum number of the NICU days that hospital can claim differs by 39 days (60 minus 21). This difference leads to a maximum of roughly 3,315,000 Yen (85,000 Yen/day*39 days or roughly $36,800) of additional reimbursement for hospitals. But we still see some observations just above these cut-offs.[36] This small magnitude of sorting may indicate the difficulty

[34] We are not aware of any other programs in Japan that uses both 1000 grams and 1500 grams birth weights as cutoffs. One could explore this issue by examining the degree of manipulation by hospital ownership type (Duggan 2000). Due to the small sample size, we cannot detect any differences by ownership types.

[35] Discussions with physicians indicates that it is possible that physicians or nurses weigh the newborns several times and report the lowest birth weight, knowing the differential reimbursement just below the cut-offs. However, they also mentioned that they could manipulate birth weights a maximum of 10-20 grams using this method.

[36] To gauge the rough magnitude of this manipulation, we count the number of births in the range of 10 grams around 1500 gram cut-offs using the universe of births in 2005 (Figure B.1C in the Appendix). The number of births weighing between 1490 and 1499 grams is 181, while between of 1500-1509 gram is 104. If we simply assume that birth from the above the cutoff are moved to below the cut-off within this range, the implied shift of birth is around 43 per year, which is (189-104)/2. Since the additional revenue by shifting one baby is $36,800, the total cost for the government is US $1,582,400. Similarly, the number of births between 990 and 999 grams is 108, and that of 1000-1009 grams is 55, which implies the shift of 27 births.

of the manipulating birth weights.[37] However, it is plausible that the manipulation may become severe if the stakes get high enough. Indeed, Figure B.1 and Table B.1 shows that magnitude of manipulation is larger in later years, especially in 2005 after PPS is introduced. More attention should be paid on manipulation of birth weight, since a handful of states in the US now uses modified version of original DRG that incorporates birth weight in their grouping of diagnosis and reimbursement for states Medicaid plan.

Since the degree of manipulation is not substantial, our difference-in-difference regressions in the following sections are not sensitive to exclusion of births near the birth weight cutoffs. We include them in all the regressions results reported below.

2.6. NICU utilization

2.6.1. Regression results

Our main estimation results on NICU utilization are shown in the Table 2.4. The first three columns present the result for NICU use, a dummy that equals one if the newborn stays at least one day in NICU. Since NICU utilization is high even before PPS, we estimate it using a Probit model.[38] We do not find that hospitals

[37]See also Camacho and Conover (2011), Chetty et al. (forthcoming), and Saez (2010) for other forms of manipulation or bunching.

[38]We also estimated OLS. While the estimates are smaller in magnitude than in Probit, none of them are statistically significant in OLS either.

use NICU more often after the introduction of PPS (column 1). We divide the sample into births weighing more than 1500 grams (column 2) and less than 1500 grams (column 3), but the estimates are not statistically significant in either cases.

The next three columns in Table 2.4 present the results on the NICU days, the number of the days that hospitals claim on the NICU utilization conditional on at least one day in NICU. Column 4 shows that the newborns stay 2.83 days longer after the adoption of the PPS. Column 5 and 6 in Table 2.4 show that the results are largely driven by birth weights less than 1500 grams. While birth weights more than 1500 gram stay 0.56 days longer on average (not statistically significant), birth weights less than 1500 grams stays 4.77 days longer. Since average length of stay in NICU for birth weights less than 1500 gram is 46.0 days on average before the PPS, this increase corresponds to 10.4 % increase of NICU days.

We then investigate whether the probability that the newborns reach the maximum number of the NICU days set by the birth weight range increases in the last three columns in Table 2.4. However, for overall, as well as any birth weight range, the estimates are not statistically significant. For example, birth weighing less than 1500 grams that stay at least one day in NICU is 14.2 percentage points more likely to reach the maximum days after PPS, even though it is far from statistically significant at the conventional level (t= 0.65). These results indicate that hospitals may increase the length of NICU stay for all births.

2.6.2. Robustness checks

In this subsection, we examine robustness of our result to three other explanations. First, we examine further the concern of endogenity of the adoption in PPS; second, we examine the possibility that newborns that are born after the adoption of PPS are sicker. Finally, we investigate whether our result is driven by a mere increase of supply of NICU beds. Overall, none of the alternative explanations is sufficient to account for our results. Since we find the largest effect on NICU days for birth weights less than 1500 grams, we focus on this group in the following analysis. Table 2.5 shows the results. To make the results comparable with our basic results, column 1 in Table 2.5 reports the estimated coefficients from the basic specification.

2.6.2.1. Endogenity of participation. Before estimating a number of specifications, Figure 2.5 shows the results of an event-study analysis where we replace the policy dummy in equation (2.1) by the series of the dummies for each year from the adoption of PPS. Due to data limitations, we only have two years of data before implementation of the PPS. The outcome is NICU days and we focus on births weighing less than 1500 grams. Figure 2.5 shows that there is no pre-trend before the PPS is implemented, and a substantial jump of around five days after the implementation. This result is reassuring since we can rule out the possibility of strategic behavior a year before the implementation of PPS, and mitigates the concerns over the endogenity of the implementation of PPS.

The event study analysis in Figure 2.5 mitigates the concerns of the endogenity of the implementation of PPS since we do not observe any pre-trend. Nonetheless, we further take two different approaches to show that our results may not be driven by the endogenity of the hospital participation in PPS. First, we include a lead dummy which equals one just prior to the year when hospitals join the new payment system. Specifically, it is one for year 2005 for the hospitals that adopted new payment system in 2006, and one for year 2007 for the hospitals that adopted it in 2008, and zero otherwise. This inclusion of the pre-period dummy is often used to investigate the reliability of difference-in-difference estimation (for example, see Acemoglu and Finkelstein 2008), and serves as a specification test to see whether there are any differential trends in the variable of interest before the introduction of policy change. For instance, if the hospitals exploit the revenue-neutral nature of the PPS, hospital would have increased its treatment intensity just prior to the adoption of PPS. The lead dummy should capture such a behavior. Column 2 in Table 2.5 shows that including lead dummy does not change the magnitude of the coefficients from the main result in column 1. Also the size of the "lead" dummy is small in magnitude compared to the main variable of interest, and is not statistically significant at conventional levels.

Second, we include a hospital linear time trend to capture pre-existing time trends that are specific to each hospital. If there is a strategic behavior mentioned above, the hospital specific linear trend may capture it to some extent. This specification is the most stringent form among all specifications since it leaves little variation in the variables of interest. Column 3 shows that the coefficient on post dummy is still statistically significant at the 10% level even in this most stringent form of regression, and in fact the magnitude of the coefficient gets even larger (7.00 days).[39] Overall, there seems a little concern that endogenity of the participation in PPS is driving our results on NICU utilization.

2.6.2.2. Sicker newborns. Another interpretation for our results is that the newborns after the PPS are sicker and thus these newborns need more intensive care. It is hard to imagine a sudden change in the distribution of the birth weight and severity among low birth newborns since the birth weight and severity of the newborns is not easy to predict in advance and, the number of low birth weight newborns does not change drastically within a few years. Nonetheless, there is a possibility that the PPS induce the hospitals to focus on the treatment of diagnosis that hospitals have highest cost efficiency (Dranove 1987). If this leads to the concentration of sicker babies in the hospitals that adopted the PPS specifically in 2006 and 2008, our results could be spurious.

To examine the possibility of change in birth distribution, we collapse the full data (before extracting the births weighing less than 2000 gram) at the hospital-year level, and we regress the number of births weighing less than 2000 gram as

[39]We also created two different control groups to examine the robustness of our results. The hospitals that adopted in early years such as 2003 and 2004 may be different from hospitals that adopted in later years. Thus, we use the hospitals that adopted early in 2003 and 2004 (early adopters) and hospitals that adopted in 2009 (late adopters) as a two control groups. Using different control groups gives similar coefficient as the main result (available from the author upon request).

well as the ratio of the births weighing less than 2000 gram among all the births observed in our dataset on the post PPS dummy. The estimate on the number of births is -1.66 (p-value 0.431), and ratio is 0.0037 (p-value 0.719).[40] These results show that the distribution of the low birth weight newborns did not change within hospitals after the adoption of PPS.

Also even though we are aware of that there is a potential concern that diagnosis coding is endogenous (Dafny 2005, Silverman and Skinner 2004), we include the three main ICD10 diagnosis (short gestation, respiratory distress syndrome (RDS), and birth asphyxia), and three main complications (retinopathy of prematurity, patent ductus arteriosus, and nutritional deficiency) in Column 5 and Column 6. The coefficient on post dummy does not change much. We also run the same estimation using the birth weight and the gestational length, which are observable birth characteristics in our data, as outcomes, to examine whether the newborns are different after adoption of PPS. However, there is little evidence on that newborns are sicker in terms of observable characteristics (available from the author upon request). As a supplement to the analysis here, Table B.2 in the Appendix examines the delivery method in the sample hospitals among all births as well as births with less than 34 weeks of gestation, which corresponds to the mean gestational length for birth less than 1500 grams since we don't have birth

[40] The standard errors are clustered at the hospital level.

weight in the delivery data.[41] It is reassuring that we don't see any increase in the delivery method that can be associated with the high risk of newborns such as emergency Caesarean sections.

2.6.2.3. Increase in Supply of NICU beds. Finally, it is possible that increase in NICU days merely reflects an increase in the availability of NICU beds, which Pauly (1980) calls an "availability effect." Unfortunately, we do not have yearly data on the number of the NICU beds.[42] However, we can identify hospitals that opened or closed the NICU beds from the information on whether the hospitals claim NICU utilization at least one day in the year. It is plausible to assume that if the hospitals have NICU beds, the beds should be utilized. By this method, we find that there are 8 hospitals in our data that started claiming for the reimbursement for NICU utilization and 7 hospitals that stopped it. We exclude these 15 hospitals and run the same main specification. Column 7 shows that coefficient on post dummy does not change. Therefore our results are not driven by the mere change in supply of the hospitals that opened or closed the NICU beds. Also it is important to note that hospitals cannot easily increase or decrease the number of NICU beds. The hospitals also need to increase the equipment and staff to meet government requirements for NICUs.

[41] Since the unit of observation in this data is delivery of mother instead of the infants born, we only have gestational length but not the birth weight of infants.

[42] We only have data for NICU beds in 2008 for 144 out of 188 hospitals. The number ranges from 3 to 36. Both the median and mean number of beds is 9.

2.7. Health outcomes and the size of the induced demand

The remaining two questions are whether longer stays in NICU had any observable health impacts, and what would be the monetary value of these medical procedures.

2.7.1. Health outcomes

Table 2.6 investigates the first question by examining mortality. We should take the mortality results with caution, though, because the mortality rate is quite low in Japan and due to the small sample size, the effect on mortality is not precisely estimated.[43] Also, we cannot track post-discharge mortality, so at best these are short-term health outcomes. Table 2.6 shows the results for 7-day, 28-day, and 90-day mortality, but we do not find any evidence as expected that mortality changes after PPS is adopted, partly because mortality rate is very low in Japan.[44]

On the other hand, there is suggestive evidence that a longer stay in NICU may not affect the mortality obtained by comparing the results between length of stay in NICU and the total length of stay in the hospital. Total length of stay (TLOS) in hospital may serve as the summary measure of the sickness of the newborns, and hence treatment intensity. The first row in Table 2.7 shows that TLOS increases

[43]For example, see Itabashi et al. (2009) for very low mortality rate in Japan among Extremely Low Birth Weight Infants (births weighing less than 1000 grams).

[44]We also look at deaths within 24 hours, 3-day, and 60-day, but none of them are statistically significant.

by 1.82 day for all births weighing less than 2000 grams and 1.93 days for births weighing less than 1500 grams, but neither of these is statistically significant. This indicates that while the length of stay in NICU increases by roughly five days for birth weighing less than 1500 grams, TLOS does not increase as much. This result is plausible, since staying longer in normal hospital beds may not be profitable to hospitals, especially beyond the national average of TLOS, since the average per-diem fixed payment is declining as the newborns stay longer.[45] In the second row in Table 2.7, we also examine the number of surgeries, which may also serve as a measure of treatment intensity. It is important to note that even simple procedures such as blood transfusions and tapping of the lungs are recorded as "surgeries" in Japan, since they require any skills of physicians. We excluded blood transfusions from our measure of surgeries. We do not observe any statistically significant changes in this variable either.

2.7.2. Size of the induced demand

As this induced demand results in no observable improvement on infants' health, the next question is what is the size of this induced demand? Since we have data

[45]The per-diem fixed payment is three step declining function, which is designed so that if the patients stays for the national average of days in the hospitals, the hospital receive the national average of the reimbursement (See Matsuda et al 2009 for detail). An alternative interpretation of this result is that because of the per-diem nature of PPS in Japan, payment of three step declining function did not lead to a decline in TLOS, which is much longer on average than most of the OECD countries. For example, the average length of stay for acute care in Japan is by far longest among the OECD countries. In 2006, they are 19.2 (days) in Japan, 5.3 in France, 7.9 in Germany, 6.7 in Italy, 7.6 in the United Kingdom, and 5.6 in the United States (OECD, 2009).

on price for each procedure calculated under national fee schedule, we can run the main specification using the reimbursement for NICU as a dependent variable. Table 2.8 presents the size of inducement for NICU utilization. For births weighing less than 1500 grams, the increase in reimbursement for longer stays in NICU is 489,000 Yen ($5,400).[46] Since the average reimbursement for births weighing less than 1500 grams before the introduction of PPS was 5,176,700 Yen ($57,500), this increase corresponds to 9.5% increase in reimbursement.

We also investigated whether there is any change in treatment intensity of other medical procedures, as measured by price of national fee schedule. It is possible that increase in NICU is the result of a reduction in necessary medical procedures included in the hospital-fee procedures. Also, it is possible that we may observe an increase in surgeries, another major component that are excluded from the hospital-fee procedures. Table 2.9 investigates these possibilities. In sum, we did not see any change in other medical procedures. This result may indicate that hospitals or physicians cannot reduce or increase the medical procedures that can potentially lead to adverse health outcomes.

Since the reimbursement under the PPS is designed to be revenue-neutral for each hospital for hospital-fee procedures (so in principle, the government cannot save money on hospital-fee procedures), this increase in the reimbursement for NICU utilization in doctor-fee procedures can be taken as the magnitude of the

[46]Even though we don't have cost information, if the cost stays constant before and after the adoption of PPS, the change in revenue is equivalent to change in profit.

additional reimbursement incurred by the implementation of new payment system. Or taken differently, this amount is an additional cost to society which does not result in any observable improvement in the short-term health outcomes of infants. If all the other hospitals behave the same way as the hospitals observed in this data, this increase in the reimbursement for the NICU utilization can result in additional medical expenditure of as much as 10.6 trillion Yen ($117 million).[47]

2.8. Conclusion

The title of Phelps's (1986) "Induced demand: can we ever know its extent?", still remains as a question today. In this paper, we focus on at-risk newborns to examine evidence on the size of the supply-induced demand. At-risk newborns are less subject to selection bias, since the birth weight and severity of newborns' conditions are often difficult to predict in advance. Also, we focus on NICU utilization, which is arguably less harmful to patients than previously studied procedures, like Caesarean sections.

We find that the hospitals increase the number of days in NICU in response to a policy change that makes NICU utilization more profitable than other medical procedure. Also, we did not find any evidence that this marginal increase in NICU utilization had an impact on newborns health. This increase in inducement

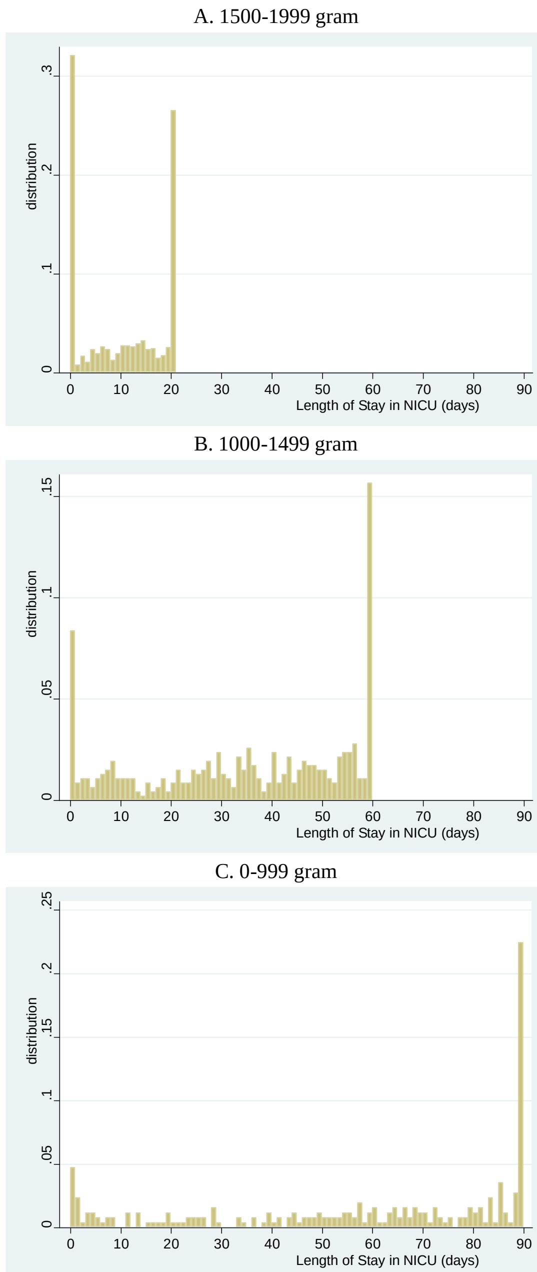
[47]In 2008, the number of births weighing less than 1500 grams in Japan is 21,667. 489,000 Yen times 21,667 is 10,595,163,000 Yen. Since we cannot determine whether manipulation of birth weight is due to benevolence or gaming, we did not take the cost of the manipulation into our social cost calculations.

translates into additional reimbursement of 489,000 Yen (\$5,400) per newborn for births weighing less than 1500 grams. If we take this figure literally, the increase in reimbursement can lead to an additional social cost of 10.6 trillion Yen (\$117 million) without any observable improvement in health outcomes.

Even though our results may be only applied to a specific case of at-risk newborns, this research may indicate that we may observe much larger supply-induced demand if we could mitigate the selection bias and focus on less risky medical procedures such as NICU.

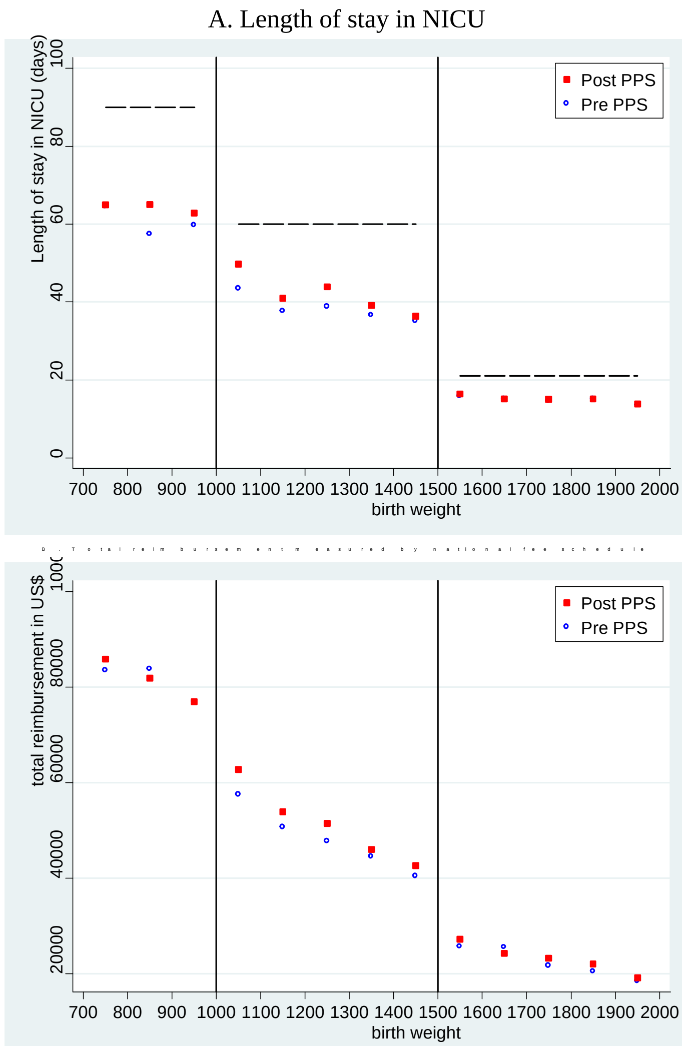
One limitation of this paper is that since we only focus on particular patients (at-risk newborns), our results do not capture the overall response of the hospitals to the introduction of the PPS. For example, it is plausible that hospitals may make profits on newborn treatment to compensate for losses in the treatment of other diagnoses, such as cancer, in response to adoption of PPS. While it is difficult to compare the severity of the patients across hospitals and thus to identify diagnoses for which hospitals have the highest cost efficiencies (Dranove 1987), future research should examine whether hospitals devote resources differentially to such diagnoses.

Figure 2.1: Length of Stay in NICU by Birth Weight Range



Note: The maximum days for birth more than 1500 gram, more than 1000 but less than 1500 gram, and less than 1000 gram, are 21, 60 and 90 days, respectively. The data used here are newborns during pre PPS period at the hospitals that adopted the PPS in 2006 and 2008 (treatment hospitals).

Figure 2.2: Pre and Post PPS



Note: The two vertical lines correspond to 1000 and 1500 grams, where the maximum number of the days hospitals can claim reimbursement for NICU utilization differs. The maximum days for birth more than 1500 grams, between 1000 and 1500 grams, and less than 1000 grams, are 21, 60 and 90 days, respectively. The three horizontal dotted lines in Figure A are these maximum days for each birth range. The total reimbursement in Figure B is calculated based on national fee schedule. We converted Yen to US 2009 dollar to make it comparable with Figure 3A in Almond et al. (2010), which draws similar graph for the US. Exchange rate of 90 Yen per US dollar is used. To avoid a composition effect from hospitals that adopted the PPS at different timings, this graph uses data from one year before and one year after the adoption of PPS for hospitals that adopted in 2006 and 2008. The bin size is 100 g. There are 4,684 observations in total.

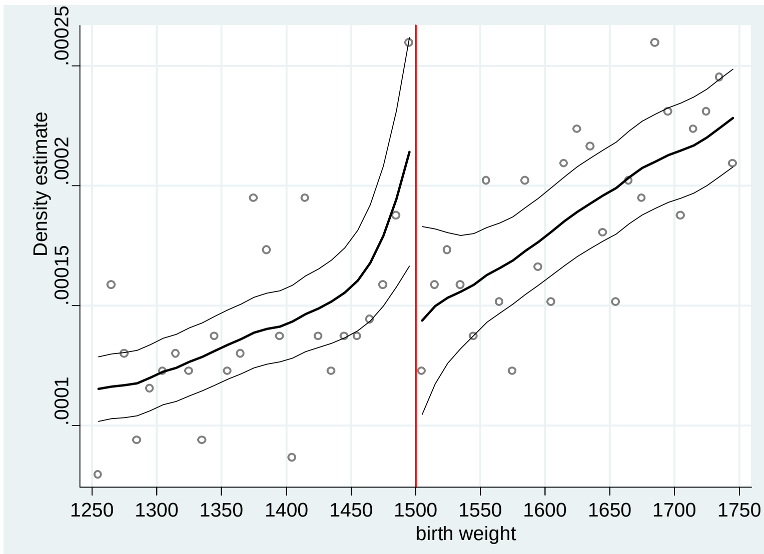
Figure 2.3: The Birth Distribution Pre and Post PPS

A. Hospitals with NICU beds

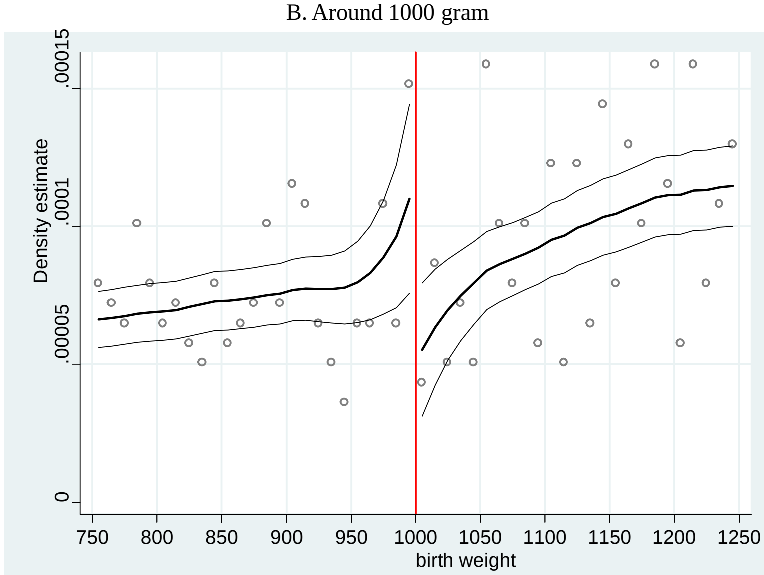
Note: This histogram uses only births at the hospitals that adopted Prospective Payment System (PPS) in 2006 and 2008, which have both pre and post data. The two vertical lines correspond to 1000 and 1500 grams, where the maximum number of the days that hospitals can claim reimbursement for NICU utilization differs. The dotted lines corresponds to every 100 grams. The bin size is 20 grams for upper graph, and 50 grams for lower graph due to small sample size.

Figure 2.4: McCrary's density test (NICU hospitals post PPS)

A. Around 1500 gram

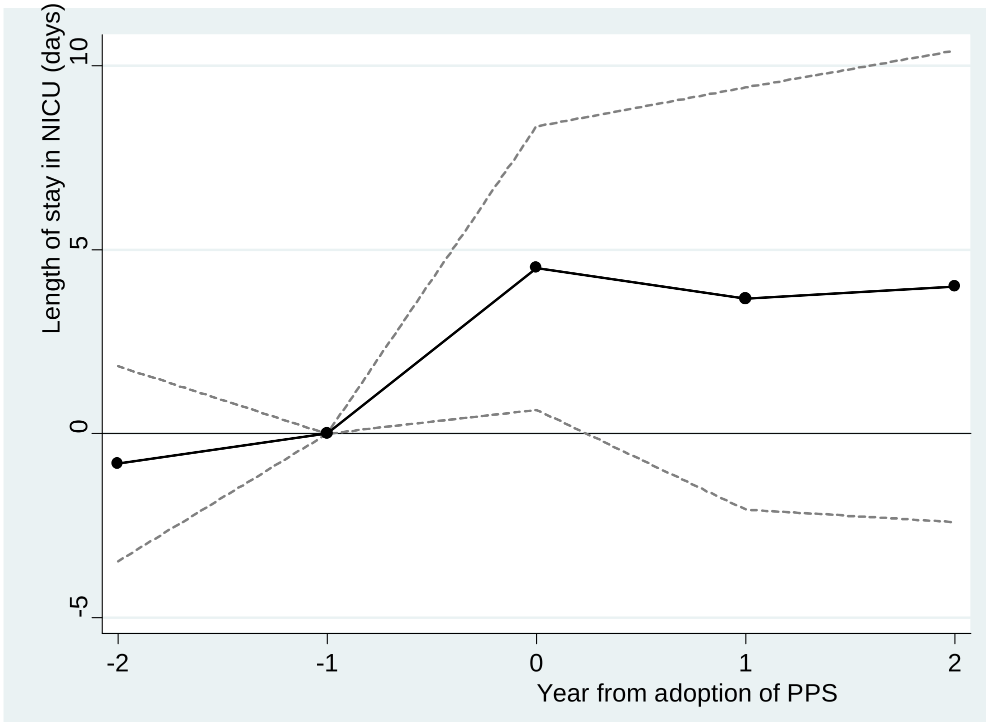


B. Around 1000 gram



Note: This graph uses the same sample used for Post PPS in Figure 3A.I use the pilot bandwidth of 100 gram with the binsize of 10 gram. The log difference in distribution at 1500 gram is -0.84 (t= -2.68), and 1000 gram is -0.45 (t=-2.28) for post PPS. Thin line corresponds to the 95 % confidence interval. For pre PPS, none of the estimates are statistically significant.

Figure 2.5: Event-study Analysis: Change in Length of Stay in NICU



Note: Year zero is when PPS is adopted. Dashed line corresponds to the 95 % confidence interval.
The sample focuses on the birth less than 1500 grams.

Table 2.1: Hazard analysis: Year to adoption of PPS

Dep: Year to adoption	Hazard Rate
Number of beds	1.001**
	[0.021]
Ownership: semi-public	0.627*
	[0.099]
Ownership: government	0.412***
	[0.001]
Teaching hospital	1.238
	[0.726]
Care level: secondary care	2.311
	[0.162]
Care level: tertiary care	1.697
	[0.400]
Have ER section	0.637
	[0.375]
Have mandatory hosp within same HSA	0.996
	[0.984]
Log Likelihood	-520.23
Sample size	124

The hazard rate is reported, and the p-value is reported in blanket. Significance level * p<0.10, ** p<0.05, *** p<0.01. All explanatory variables are predetermined hospital characteristics collected in 2002, before first implementation of PPS in 2003. The omitted ownership type is non-profit hospitals. The omitted care level is primary care. HSA stands for hospital service area.

Table 2.2: Summary Statistics by hospital groups

Variables	Year when PPS is adopted			
	2003/2004	2006/2008		2009
	Post only	Pre	Post	Pre only
A Birth characteristics				
Birth weight (grams)	1,468.1	1,502.0	1,468.2	1,491.8
Gestational length (weeks)	31.9	31.9	31.6	32.0
Male	0.50	0.49	0.51	0.51
B NICU				
Utilization	0.80	0.78	0.81	0.81
Length of stay in NICU (days)	30.5	30.0	31.4	28.6
Fraction of maximum stay in NICU	0.18	0.20	0.18	0.14
C Health measures				
Death within 7 days	0.03	0.01	0.02	0.02
Death within 28 days	0.04	0.02	0.03	0.02
Death within 90 days	0.05	0.02	0.03	0.03
D Treatment Intensity				
Total length of stay (days)	52.7	52.7	53.1	52.8
Total number of surgeries (times)	0.45	0.36	0.45	0.45
E Reimbursement (thousand Yen)				
Total payment per patient ((1)+(2))	3,267	3,021	3,249	3,200
(1) Doctor-fee procedures	2,217		2,314	
(2) Hospital-fee procedures	1,051		935	
Number of hospitals	74	72		41
Number of Observations	6,455	1,725	2,959	2269

Note: The sample is composed of births weighing less than 2,000 grams in the hospitals that already have NICU beds. The data span 2004-2008. Hospitals that adopted prospective payment system (PPS) in 2003 and 2004 only have post PPS data and hospitals that adopted in 2009 only have pre PPS data.

Table 2.3: Density Test

binsize	10	10	20	20
bandwidth	50	100	100	200
Cutoff (grams)				
800	-0.26	-0.36	-0.39	-0.17
	(0.42)	(0.31)	(0.31)	(0.23)
900	0.43	0.16	0.11	0.06
	(0.37)	(0.28)	(0.27)	(0.21)
1000	-0.98***	-0.84***	-0.61***	-0.35*
	(0.42)	(0.31)	(0.30)	(0.21)
1100	0.61	0.06	-0.09	0.06
	(0.44)	(0.28)	(0.29)	(0.20)
1200	-0.52	-0.39	-0.36	-0.27
	(0.38)	(0.25)	(0.24)	(0.18)
1300	0.20	-0.03	-0.07	0.03
	(0.35)	(0.25)	(0.24)	(0.17)
1400	-0.29	-0.32	-0.28	-0.15
	(0.33)	(0.22)	(0.22)	(0.15)
1500	-0.72***	-0.45***	-0.42***	-0.27*
	(0.28)	0.20	(0.20)	(0.15)
1600	-0.11	0.06	0.06	0.09
	(0.29)	(0.20)	(0.20)	(0.14)
1700	-0.29	-0.11	-0.11	-0.12
	(0.25)	(0.18)	(0.17)	(0.13)
1800	0.16	0.25	0.23	0.18
	(0.23)	(0.16)	(0.16)	(0.12)
1900	-0.30	-0.04	-0.01	-0.05
	(0.20)	(0.15)	(0.15)	(0.11)
2000	-0.36	-0.22	-0.21	-0.13
	(0.21)	(0.15)	(0.16)	(0.11)

Significance level * p<0.10, ** p<0.05, *** p<0.01. See McCrary (2009) for methodological details.

Table 2.4: NICU Utilization

	NICU use dummy			Length of stay in NICU			Probability of maximum stay in NICU		
	All	>=1500 grams	<1500 grams	all	>=1500 grams	<1500 grams	all	>=1500 grams	<1500 grams
	Probit	Probit	Probit	OLS	OLS	OLS	Probit	Probit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)	(4)	(5)	(6)
Post	-0.012	-0.012	0.011	2.83**	0.878	4.77**	0.122	0.075	0.142
	(0.16)	(0.19)	(0.28)	(1.16)	(0.61)	(2.07)	(0.14)	(0.17)	(0.22)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Hospital FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
2002 HC*linear time	Y	Y	Y	Y	Y	Y	Y	Y	Y
R2/ Persudo R2	0.31	0.31	0.19	0.59	0.35	0.36	0.20	0.25	0.23
Sample size	11,856	6,512	3,631	9,915	4,897	5,018	9,610	4,686	4,429
Mean	0.80	0.70	0.93	30.3	14.1	46.1	0.18	0.21	0.14

Note: Standard errors (in parentheses) are clustered at the hospital level. Probit estimation reports the marginal effect. Significance levels * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Post is a dummy that equals one if hospital is under the PPS and zero otherwise. All specifications include the year fixed effects and hospital fixed effects. Controls are birth weight, gestational length, and male dummy. In addition to fixed effects and controls, we include 2002 hospital characteristics (number of beds, ownership of the hospital, a dummy for teaching hospital, level of hospital care (primary, secondary and tertiary), a dummy that takes the value of one if hospitals have an ER section, and a dummy that takes the value of one if hospitals have mandatory hospital within the same Health Service Area) each interacted with a linear time trend.

Table 2.5: Robustness checks for length of stay in NICU

	Baseline	Lead dummy	Hospital linear time	With 3 main diagnosis	With 3 main complication	Availability effect
	(1)	(2)	(3)	(4)	(5)	(6)
Post	4.77**	5.39*	7.00*	4.89**	3.84*	4.07*
	(2.07)	(2.88)	(3.95)	(2.09)	(2.25)	(2.11)
Lead		0.787				
		(2.40)				
Short gestation				4.28***		
				(1.62)		
RDS				0.875		
				(1.05)		
Birth asphyxia				-3.87		
				(2.72)		
Retinopathy of prematurity					9.90***	
					(0.98)	
Patent ductus arteriosus					2.87***	
					(1.07)	
Nutritional deficiency					3.56***	
					(0.95)	
Year FE	Y	Y	Y	Y	Y	Y
Hospital FE	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y
2002 HC*linear time	Y	Y	Y	Y	Y	Y
R-squared	0.36	0.36	0.40	0.36	0.39	0.36
Sample size	5,018	5,018	5,018	5,018	5,018	4,795

Note: Standard errors (in parentheses) are clustered at the hospital level. Significance levels * p<0.10, ** p<0.05, *** p<0.01. Post is a dummy that equals one if hospital is under the new payment system and zero otherwise. Lead is a dummy that equals one a year prior to the adoption of the PPS, and zero otherwise. All specifications include the year fixed effects and hospital fixed effects. Controls are birth weight, gestational length, and a male dummy. In addition to fixed effects and controls, we include 2002 hospital characteristics (number of beds, ownership of the hospital, a dummy for teaching hospital, level of hospital care (primary, secondary and tertiary), a dummy that takes the value of one if hospitals have an ER section, and a dummy that takes the value of one if hospitals have mandatory hospital within the same Health Service Area) each interacted with a linear time trend.

Table 2.6: Mortality

	all birth	>=1500 gram	<1500 gram
Death within 7 days	-0.003	-0.002	-0.005
	(0.005)	(0.004)	(0.012)
Death within 28 days	-0.002	-0.001	-0.000
	(0.006)	(0.005)	(0.014)
Death within 90 days	-0.003	-0.001	-0.000
	(0.007)	(0.005)	(0.016)
Sample size	12,406	6,981	5,425

Note: Each row corresponds to separate regression of OLS. The estimate on *post* is reported. *Post* is a dummy that equals one if hospital is under the new payment system and zero otherwise. All specifications include the year fixed effects and hospital fixed effects. Controls are birth weight, gestational length, and male dummy. In addition to fixed effects and controls, we include 2002 hospital characteristics (number of beds, ownership of the hospital, a dummy for teaching hospital, level of hospital care (primary, secondary and tertiary), a dummy that takes the value of one if hospitals have an ER section, and a dummy that takes the value of one if hospitals have mandatory hospital within the same Health Service Area) each interacted with a linear time trend. Standard errors (in parentheses) are clustered at the hospital level. Significance levels * p<0.10, ** p<0.05, *** p<0.01.

Table 2.7: Treatment Intensity

	all birth	>=1500 gram	<1500 gram
Total length of stay (days)	1.82	0.711	1.93
	(1.44)	(1.03)	(2.85)
Total number of surgeries (times)	0.078	0.071	0.079
	(0.05)	(0.04)	(0.08)
Sample size	12,406	6,981	5,425

Note: Each row corresponds to separate regression of OLS. The estimate on *post* is reported. See Table 5 for detail.

Table 2.8: The size of the inducement

	all	>=1500 gram	<1500 gram
	(1)	(2)	(3)
Post	297.4**	107.3	489.5**
	(127.4)	(70.4)	(227.8)
Year FE	Y	Y	Y
Hospital FE	Y	Y	Y
Controls	Y	Y	Y
2002 HC*linear time	Y	Y	Y
R2	0.57	0.41	0.32
Sample size	12,406	6,981	5,425

Note: In thousands Yen (90 Yen/US$). Estimate on *post* is reported. See Table 5 for detail.

Table 2.9: Medical spending on other procedures

	all birth	>=1500 gram	<1500 gram
hospital-fee procedures			
inspection	3.3	0.0	6.5
	(7.8)	(8.9)	(8.5)
diagnostic imaging	1.7	1.0	3.4
	(2.1)	(1.4)	(3.6)
medicine	1.5	-1.6	7.0
	(3.2)	(3.7)	(4.7)
injection	13.0	9.0	18.0
	(9.7)	(6.4)	(21.0)
doctor-fee procedures			
surgery	18.8	9.5	40.4
	(16.5)	(14.1)	(30.9)
anesthesia	-1.2	-2.2	2.2
	(4.6)	(4.6)	(8.1)
Sample size	12,406	6,981	5,425

Note: In thousands Yen (90 Yen/US$). Estimate on *post* is reported. See Table 5 for detail.

CHAPTER 3

Effects of Universal Health Insurance on Health Care Utilization, Supply-Side Responses, and Mortality Rates: Evidence from Japan

with Ayako Kondo

3.1. Introduction

Most developed countries have implemented some form of universal public health insurance to ensure that their entire population has access to health care. Even the United States, which has been a rare exception, is moving towards near-universal coverage through health care reform.[1] Despite the prevalence of universal health care, most studies on the impact of health insurance coverage on utilization and health have been limited to specific subpopulations, such as infants and children, the elderly, or the poor.[2] Estimates from a policy that focuses on the elderly (e.g., Medicare in the United States) may differ from the average impact of health

[1]The Patient Protection and Affordable Care Act, passed in March 2010, imposes a mandate for individuals to obtain coverage or pay a penalty.

[2]Examples of studies that examine specific populations include Currie and Gruber (1996a,b), Hanratty (1996), and Chou et al. (2011), on infants and children; Finkelstein (2007), Card et al. (2008, 2009), and Chay et al. (2010), on the elderly; and Finkelstein et al. (2011), on the poor.

insurance for an entire population if the price elasticity of the elderly differs from that of the younger population.[3]

This paper studies the impact of a large expansion in health insurance coverage on utilization and health by examining the case of Japan, which achieved in 1961 universal coverage for its entire population. We identify the effect of health insurance by exploiting regional variations in health insurance coverage prior to the full enforcement of universal coverage. In 1956, roughly one-third of the population was not covered by any form of health insurance, and the portion of the population uninsured ranged from almost zero to almost half, across prefectures. Our empirical strategy identifies changes in outcome variables in a prefecture in which the enforcement of universal coverage had a large impact, relative to a prefecture in which the impact was smaller.

This study also has several other advantages, compared to those in the existing literature. Since universal health insurance was achieved as early as 1961 in Japan, we can examine the impact of health insurance expansion in the long term. Since the effects incurred by such a large policy change may emerge with lags, it is important to examine the long-term impact, in order to capture the overall implication of a large policy change. Also, we provide a more detailed analysis of

[3] An important exception is Kolstad and Kowalski (2010), who examine the impact of the introduction of universal health insurance in Massachusetts in 2006; however, they are unable to explore long-term effects, because their data cover only the three years following the policy change.

supply-side responses to large demand shocks by investigating the several outcomes not explored extensively in previous studies, such as the number of physicians.[4]

We have three key findings. First, we find that the expansion of health insurance coverage results in large increases in health care utilization, measured in terms of admissions, inpatient days, and outpatient visits to hospitals. For example, our estimates imply that the introduction of universal health insurance increased impatient days by 7.3 percent and outpatient visits by 12.6 percent from 1956 to 1961. The long-term impact is even larger: the estimated increases in inpatient days and outpatient visits from 1956 to 1966 are 11.6 percent and 25.1 percent, respectively. Our estimate of the effect on outpatient visits is roughly four times larger than the estimate from the RAND Health Insurance Experiment (hereafter RAND HIE), which explores the effects of individual-level changes in insurance status.

Second, we find that supply-side responses to demand shocks differ across the types of services supplied. While the expansion of health insurance coverage did not increase the numbers of clinics and nurses even in the long term, the number of beds increased immediately in response to the expansion in health insurance coverage. Our results vis-à-vis the numbers of hospitals and physicians are mixed and sensitive to the way in which we control for regional time trends. It is not

[4] For example, Finkelstein (2007) finds a large increase in hospital employment in response to the introduction of Medicare in the United States, but her data do not include most of physicians, because physicians in the United States are not directly employed by the hospital. On the other hand, our data cover all physicians who were working at hospitals in Japan.

surprising that we observe a robust positive effect only on the number of beds, because it is less costly for existing hospitals to add beds than for new hospitals and clinics to pay large fixed costs to enter the market. Also, the total supply of physicians and nurses is generally limited by the capacity of medical and nursing schools. Furthermore, we find that even the number of beds increased at a slower rate than increases in health care utilization.

Third, despite massive increases in utilization, we find little evidence of effects on health, measured in terms of age-specific mortality. In addition to analysis that relies on prefecture-level variation, we conduct an event study using the municipality-level variation in Ibaraki prefecture and confirm that there was no effect on short-term mortality. This lack of short-term effects may be because individuals with acute life-threatening and treatable health conditions had already sought care at hospitals, despite having a lack of health insurance. As suggestive evidence, we find no change in the number of deaths from treatable diseases at that time (e.g., pneumonia), which should have fallen if universal health insurance coverage enabled some formerly untreated patients to have access to hospitals or clinics.

Taken together, our empirical results show that a large expansion in health insurance coverage increases health care utilization, without there being any observable short-term improvement in health; the magnitude of the effect on utilization is much larger than the prediction from individual-level changes in insurance status.

Another implication is that a slow supply-side response can constrain attempts to meet the demand increases induced by large policy changes.

This paper is related to several strands of literature. The first relevant body of literature comprises studies on the effect of health insurance on utilization and expenditure. The pioneering works of the RAND HIE (Manning et al. 1987; Newhouse 1992) typically find modest effects of individual-level changes in health insurance on health care utilization and expenditure. In contrast, Finkelstein (2007) examines the impact of the introduction of Medicare in 1965, and finds a much larger effect on aggregate spending than those predicted by the RAND HI by virtue of individual-level changes in health insurance. Finkelstein (2007) attributes this larger effect to a shift in supply induced by market-wide changes in demand. While we find mixed evidence of such increases in the market entries of hospitals and clinics, the magnitude of our estimates on utilization is closer to that of Finkelstein (2007) than to estimates from the RAND HIE.

The second related strand of literature comprises studies that examine whether health insurance improves health. Existing studies show evidence of the positive effects of health insurance coverage vis-à-vis infant health in Canada (Hanratty 1996), in low-income households in the United States (Currie and Gruber 1996b) and Thailand (Gruber et. at. 2012), and in farm households in Taiwan (Chou et al 2011). Studies on Medicare also tend to show that Medicare eligibility has a modest positive effect on the health of the elderly (Chay et al. 2010; Card et al. 2009).[5] Our results show that, at least in the case of Japan in the 1960s, the expansion of health insurance seems to have no short-term health effects.[6]

Finally, a growing body of literature examines the effect of a large health insurance coverage expansion on various outcomes in less-developed countries such as Mexico, Colombia, Thailand, and Taiwan.[7] Under significant credit constraints in less-developed countries, health care utilization without insurance can be inefficiently low (Miller et al. 2009). Japan's per-capita gross domestic product (GDP) in 1956 was about one-quarter of that of the United States at that time.[8] Thus, our estimates may be more relevant to developing countries that are currently

[5]Chang (2011) finds that the introduction of Taiwan's National Health Insurance led to a reduction in mortality among the elderly there, while Chen et al. (2007) find no such evidence.

[6]Although Finkelstein and McKnight (2008) find no discernible impact of Medicare expansion on mortality among the elderly, this is probably because the effect of Medicare on mortality is not large enough to be identified with regional-level aggregate data but is detectable with a regression discontinuity design with individual-level data, as employed by Chay et al. (2010) and Card et al. (2009). While the same issue may apply to our case, we supplement our prefecture-level analysis with event-study analysis at the municipality level to support our results.

[7]For example, see King et al. (2009) for Mexico; Miller et al. (2009) for Colombia; Cataife and Courtemanche (2011) for Brazil; Dow and Schmeer (2003) for Costa Rica; Hughes and Leethongdee (2007), Damrongplasi and Melnick (2009), and Gruber et al. (2012) for Thailand; and Chen and Jin (2010) for China. There are a considerable number of studies on Taiwan; see, for example, Cheng et al (2007), Chang (2011), and Chou et al. (2011). Studies on Taiwan also examine the effect of the introduction of universal health insurance; however, the empirical strategy of those studies mostly relies on difference-in-difference approaches, by comparing those previously covered to those newly covered. Such a strategy may not be able to capture the effects through market entry, as argued in Finkelstein (2007)—unlike our case, which relies on prefecture-level hospital data.

[8]Countries whose per-capita GDP is about one-quarter of the United States today include, for example, Chile and Turkey. Also, Japan's average life expectancy at that time was 66, whereas that of the United States was 70.

considering a massive expansion in health insurance coverage, than those of existing studies on developed countries such as the United States.[9] Our results show that countries planning to expand health insurance coverage drastically need to set aside enough financial resources for the anticipated surge in health care expenditures, which will be much larger than that predicted from individual-level changes in insurance status.

The rest of the paper is organized as follows. Section 3.2 describes the institutional background of the implementation of universal health insurance in Japan. Section 3.3 describes the data we use, and Section 3.4 presents the identification strategy. Section 3.5 shows the main results for utilization. Section 3.6 analyzes the supply-side responses to changes in demand, and Section 3.7 examines health. Section 3.8 concludes the paper.

3.2. Background

This section briefly reviews the history of Japan's universal health insurance system, up to the 1960s.[10] Japan's public health insurance system consists of two parallel subsystems: employment-based health insurance and the National Health Insurance (hereafter, NHI). Combining the two subsystems, Japan's health

[9]Of course, the technology available at that time was quite different from that available now. However, the major causes of death in Japan around this time were not much different from the causes of death in less developed countries now (e.g., pneumonia, bronchitis, gastritis, and duodenitis).

[10]The discussion in this section draws heavily from Yoshihara and Wada (1999).

insurance program is one of the largest in the world today, as it covers nearly 120 million people, making it almost three times larger than Medicare in the United States, which covers 43 million people (The Centers for Medicare and Medicaid Services 2010).

Employment-based health insurance is further divided into two forms: employees of large firms and government employees are covered by union-based health insurance, whereas employees of small firms are covered by government-administered health insurance. In both cases, employers must contribute about half of the insurance premiums, and the other half is deducted from employee salaries. Enrollment in the government-administered health insurance program was legally mandated to all employers with five or more employees, unless the employer has its own union-based health insurance program. If the household head enrolls in an employment-based health insurance program, his or her dependent spouse and children are also covered by employment-based health insurance.

The NHI is a residential-based system that covers anyone who lives in the covered area and does not have employment-based health insurance. Therefore, the NHI mainly covers employees of small firms (i.e., fewer than five employees), self-employed workers in the agricultural and retail/service sectors and their families, the unemployed, and the retired elderly. An important feature for our identification strategy is that the decision to join the NHI system is left to municipalities, not individuals, and individuals living in covered municipalities cannot opt out.

Both health insurance programs offer similar benefits, and cover outpatient visits, admissions, diagnostic tests, and prescription drugs. However, different coinsurance rates are applied, depending on the type of insurance; also, the rates changed several times. When universal health insurance was achieved in Japan in 1961, the coinsurance rate for NHI was 50 percent for both household heads and other family members, while that of employment-based health insurance was nearly zero for employees and 50 percent for family members. The coinsurance rate for NHI for household heads was reduced to 30 percent in 1963, and then that for other NHI enrollees was reduced to the same rate in 1968. In 1973, the coinsurance rate of employment-based health insurance for family members was also reduced to 30 percent.[11]

The history of Japan's public health insurance system goes back to the 1920s. First, in 1922, enrollment to employment-based health insurance was mandated to blue-collar workers in establishments with ten or more employees. In 1934, mandatory enrollment was expanded to workers in establishments with five or more employees. Then, to address the lack of health insurance among people excluded from employment-based health insurance, the NHI was introduced in 1938.

During World War II, the wartime government rapidly expanded the NHI, and by 1944, universal health insurance had seemingly been achieved. However, in reality, coverage was far from universal: the medical system was not fully functioning

[11]The cap on the maximum limit on out-of-pocket expenditures was not introduced until 1973.

owing to budgetary constraints incurred by the war. Furthermore, after defeat in the war, hyperinflation and other disruptions caused a serious breakdown in the health insurance system.

The Japanese government, with the support of General Headquarters, started to restore the health insurance system immediately following the war. However, even in 1956, roughly one-third of the population (i.e., 30 million people)—mainly the self-employed, employees of small firms, the unemployed, and the retired elderly—were still not covered by any form of health insurance. Those without any health insurance had to bear the full cost of health care utilization. This lack of coverage was partly because a nonnegligible number of municipalities had not yet rejoined the NHI system. Therefore, in 1956, the Advisory Council on Social Security recommended that all municipalities should join the NHI system. Given this recommendation, the Four-year Plan to achieve universal coverage by 1961 was proposed in 1957 by the Ministry of Health and Welfare.[12] By April 1961, all municipalities had joined the NHI, and universal health insurance was achieved.

Figure 3.1 shows the time series of health insurance coverage by the NHI, employment-based health insurance, and all types of insurance combined. The figure also includes a linear trend extrapolated from data prior to 1956. Two vertical lines indicate 1956, which is the reference year before the start of Four-year Plan, and 1961, the year in which universal health insurance was achieved. The

[12]In 1959, an amendment to the National Health Insurance Act legally prescribed the mandatory participation of all municipalities in the NHI, by April 1961.

number of individuals covered by both employment-based health insurance and the NHI gradually increased until the mid-1950s, and there was a sharp increase, especially in NHI coverage, in the late 1950s. During the four years immediately preceding 1961, around 30 percent of the total population became newly covered by health insurance.

Crowding-out from employment-based health insurance due to the introduction of NHI seems to have been negligible. The insured were likely to have preferred employment-based health insurance, because it offered lower coinsurance rates and the employer contributed to the premium. In theory, the NHI expansion could have increased the number of self-employed workers by insuring them, in the absence of employment-based health insurance.[13] Another possible implication of crowding-out is that the introduction of the NHI could have induced firms to reduce their size to fewer than five employees, in order to be exempt from contributing to employment-based health insurance. Appendix Section C.1 assesses both possibilities. We find no strong evidence of either type of crowding-out.[14]

There are a few important institutional features of Japan's health insurance system, from the supply-side perspective. First, its detailed fee schedules are set by centralized administration, and reimbursement from the health insurance system

[13]See, for example, Madrian (1994) on the job-lock effects of employment-based health insurance.

[14]The proportion of self-employed workers in the labor force declined just as quickly in prefectures that experienced a large NHI expansion as those prefectures that experienced a small expansion. Also, changes in the fraction of establishments with fewer than five employees do not seem to correlate systematically with NHI coverage in 1956. See Appendix Section C.1 for details.

to medical providers follows these schedules strictly.[15] Until 1963, each medical institution was able to choose one schedule from two options, but it had to apply the same schedule to all patients. Thus, there was little room for each hospital or physician to charge differential fees for specific types of patients, as seen in the United States (Cutler 1998). Furthermore, from 1963, fee schedules are integrated into a unified schedule that is applied nationwide.[16] Second, there was no effective legal obligation for physicians or hospitals to provide cares to uninsured patients.[17] Public aid for the uninsured was limited to patients quarantined with tuberculosis and other diseases specified the in Infectious Deceases Prevention Act and those who lived on welfare.

[15] According to Ikegami (1991, 1992) and Ikegami and Campbell (1995), the national schedule is usually revised biennially by the Ministry of Health, Labor and Welfare through negotiations with the Central Social Insurance Medical Council, which includes representatives of the public, payers, and providers.

[16] This stringent fee control is considered one of the primary reasons why Japan had been able to keep a relatively low total medical expenditures-to-GDP ratio (Ikegami and Campbell 1995). The ratio of total medical expenditures to GDP had been slightly higher than 3 percent throughout the 1950s. Although it gradually increased during the early 1960s, it leveled off at around 4 percent in the mid-1960s until 1973, when healthcare services were made free for elderly. In addition, there is no trend break in per-capita medical expenditures until 1973, either.

[17] Article 19 of the Medical Practitioners Act stipulates that a physician cannot refuse to diagnose and treat without a legitimate reason. However, this Act was not very effective, because the lack of ability to pay the fee was considered a legitimate reason. There was no legal obligation equivalent, for example, to the Emergency Medical Treatment and Labor Act in the present-day United States, which mandates that hospitals must provide stabilizing care and examination for people who arrive at an emergency room for a life-threatening condition, without consideration of whether a person is insured or has the ability to pay.

In contrast to the strict price control, entry and expansion of private hospitals had been left unrestricted until the upper limit of the number of beds in each region was introduced in 1985. In the 1950s and 1960s, the government attempted to increase the supply of medical institutions in regions with short supplies, but the effect of this move seems to have been limited. Construction of public institutions is of course guided by the government, but its impact is small compared to the increase in private hospitals.[18] Regarding private institutions, Medical Care Facilities Financing Corporation was founded in 1960 to facilitate the finance of private medical institutions. This financing alleviates the credit constraints of potential entrants, but whether to enter the market or expand, and where to build hospitals, are still voluntary decisions.

The supply of physicians and nurses is constrained by the capacity of medical schools and nursing schools. However, their mobility was not controlled by the national government. Although medical schools had some power to control the choice of hospitals at which their alumnus work, there seemed to be no coordinated system to allocate physicians or nurses across prefectures.

[18]The ratio of public hospitals to the total number of hospitals was 33 percent in 1956; the number of public hospitals increased by only 6 percent by 1965, whereas that of private hospitals increased by 48 percent. Consequently, the share of public hospitals fell to 27 percent in 1965. Admittedly, however, since public hospitals tend to be larger than private ones, the share in terms of the number of beds was larger: 55 percent in 1956. Nonetheless, the expansion speed of private hospitals was faster. The number of beds in public hospitals increased by 34 percent during the 1956–65 period, whereas that in private hospitals increased by more than 100 percent. Since we are not aware of any prefecture-level data on the number of hospitals by ownership, we are not able to examine separately the effect by ownership type.

3.3. Data

Our data derive from various sources. Although the decision to join the NHI was made at the municipality level, municipality-level data are not available for most of the outcomes and explanatory variables. Thus, our unit of observation is the prefecture year, except for a supplemental event study using municipality-level data from Ibaraki prefecture.[19] In Section 3.7.2, we explain the data from Ibaraki in detail. We mainly focus on the 1950–70 period, although some specifications use a shorter time period, owing to the limited availability of data pertaining to variables of interest.[20] Appendix Table C.1 describes the definition, data sources, and available periods for each variable. All expenditure variables are converted to real terms at 1980 price levels, using a GDP deflator.

3.3.1. Health Insurance Coverage Rate

We construct the rate of health insurance coverage for each prefecture at year 1956, the year before the implementation of the Four-year Plan, as follows. First, the population covered by the NHI in prefecture p in 1956 (NHI_p) is obtained from the Social Security Yearbook. Second, the population covered by employment-based

[19]It is important to note that our analyses at the prefecture level can capture the effects through hospital entry and exit, unlike studies that rely on hospital-level data. In all, 46 prefectures, excluding Okinawa, returned to Japan in 1973.

[20]We do not extend our data beyond 1970, because some prefectures started to provide free care for the elderly in the early 1970s; this could have confounded our results. See Shigeoka (2011) for details on health care for the elderly in Japan. Also, attenuation bias caused by migration among prefectures would become more severe as the sample period grows longer.

health insurance is imputed from nationwide, industry-level coverage rates and the industry composition of each prefecture's workforce.[21] Note that, owing to data limitations, we need to assume that the coverage rate within each industry does not vary across prefectures (i.e., the variation of employment-based health insurance across prefectures is attributable solely to variation in industry composition).[22] Then, for each year and prefecture, the coverage rate of each industry is weighted by the ratio of household heads in the industry. We use this weighted sum of industry-level coverage rates as the coverage rate of employment-based programs in each prefecture.[23]

Specifically, let E_CovR_j denote the ratio of households covered by employment-based health insurance, among those with a household head who works in industry j, in 1956. Let W_{pj} denote the population living in prefecture p with a household head who works in industry j in 1956. Then, the imputed population covered by employment-based health insurance in 1956 in prefecture p can be written as $\sum_j W_{pj} * E\_CovR_j$ where E_CovR_j is available from the Comprehensive Survey

[21]Specifically, the population was divided into the following 13 categories: agriculture, forestry and hunting, fishing, mining, construction, manufacturing, whole sale and retail trades, finance and real estate, transportation and other utility, service, government sector, unknown (employed), and non-employed.

[22]Although some prefecture-level tables of employment-based insurance have been published, most of these tables show the location of *employers*, not the residence of employees.

[23]A potential bias arising from omitting heterogeneity in the coverage rate within each industry across prefectures is that the ratio of population without health insurance may be overestimated for prefectures that have larger firms. Larger firms are much more likely to offer employment-based health insurance, and they tend to be located in either Tokyo or Osaka. Thus, as a robustness check, we estimate the case without Tokyo and Osaka from the sample.

of the People on Health and Welfare.[24] W_{pj} is calculated as linear interpolations from the 1955 and 1960 Censuses.

Lastly, the total population of each prefecture, pop_p, is taken from the Statistical Bureau's website.[25] Then $CovR_p$, the ratio of prefecture p's population covered by any kind of health insurance in 1956, is estimated as follows:

$$CovR_p = [NHI_p + \sum_j W_{pj} * E\_CovR_j]/pop_p \tag{3.1}$$

We define the impact of the health insurance expansion, $impact_p$, as the proportion of the population *without* health insurance in prefecture p in 1956:

$$impact_p = 1 - CovR_p \tag{3.2}$$

Figure 3.2 shows the regional pattern of $impact_p$, the proportion of people without health insurance in 1956, one year before the implementation of the Four-year Plan. The figure shows substantial regional variation in the health insurance coverage rate. Most of the variation in this coverage rate comes from variation

[24]Note that the Comprehensive Survey of the People on Health and Welfare classifies a household as being covered by an employment-based program if at least one of the household members is covered by an employment-based program. Although this is a sensible approach given that most employment-based insurance also cover spouses and children, it may also overstate the coverage rate of employment-based programs if some of the other household members are covered by the national program. Thus, as a robustness check, we tried replacing with zero the coverage rate of employment-based program for households in the agricultural sector, because most agricultural workers were self-employed in Japan at that time. Result remained virtually unchanged.

[25]These data seem to be interpolated from the Population Census by the Statistics Bureau, and the value is as of October 1. Thus, we take the average of 1955 and 1956, thus deriving the population as of April 1, 1956.

in the NHI coverage rate. Indeed, the coverage rate of employment-based health insurance tends to be high in prefectures with a low total coverage rate; thus, the NHI coverage rate varies more than that of the sum of employment-based health insurance and NHI.[26]

The proportion of the population without health insurance coverage ranged from almost zero in several prefectures (including Yamagata and Niigata) to a high of 49 percent in Kagoshima. The proportion of the population without health insurance was relatively high in the southwest prefectures and low in the northeast prefectures. Additionally, prefectures with large populations—such as Tokyo and Osaka—tended to have low coverage rates, given the additional time needed to build a health insurance tax-collection system and to reach agreements between local governments and medical providers in cities with larger numbers of physicians (Yoshihara and Wada 1999).

It is difficult to know *a priori* whether average income positively or negatively correlates with the initial health coverage rate. On one hand, affluent prefectures tended to have a high rate of employment-based health insurance coverage. On the other hand, poorer prefectures may have tried to restore the NHI earlier, to insure the poor. Figure 3.2 suggests that the latter effect dominated the former given that the northeast part of Japan is on average poorer than the southwest. Figure 3.3

[26] $Var(CovR_p)$ can be decomposed into the variances of the coverage rates by the NHI, that by employment-based insurance, and the covariance between them. The variance of NHI coverage rates is 0.037, which is larger than $Var(CovR_p) = 0.031$. The variance of employment-based insurance is as small as 0.004, and the covariance between coverage rates of two types is –0.005.

shows the correlation between changes in per-capita gross national product (GNP) and $impact_p$. The figure clearly shows that larger increases in the health insurance coverage rate were not driven by income growth; on the contrary: increases in the coverage rate may slightly *negatively* correlate with the growth rate of per-capita GNP in the long term. Section 3.4 discusses how we address the fact that the distribution of the initial health insurance coverage rate may not be completely random.

3.3.2. Outcome and Explanatory Variables

Our main outcome variables are divided into three categories: utilization; capital and labor inputs, as the supply-side response; and mortality rates. The three measures for utilization are admissions, inpatient days, and outpatient visits. Admissions represent the number of admissions to hospitals in each prefecture per calendar year. Inpatient days are the sum of the days in hospitals among all inpatients, while outpatient visits are visits to hospitals for reasons not requiring hospitalization. Note that these variables are limited to the utilization of hospitals (defined in Japan as medical institutions with 20 or more beds), because clinics (institutions with no more than 19 beds) are excluded from the survey.[27]

[27]Unlike in the United States, direct outpatient visits to hospitals are a common practice in Japan, since there are no restrictions on the patients' choice of medical provider. Therefore, an increase in the number of outpatient visits may simply reflect the fact that people are switching from clinics to hospitals for outpatient visits. However, almost all admissions occur at hospitals, and thus our data capture the universe of admissions and inpatient days in Japan.

From a number of different sources, we can also obtain the numbers of hospitals, clinics, beds, physicians, and nurses, in order to explore supply-side response to expansions in health insurance coverage. As a measure of health outcomes, we compute the age-group-specific mortality rate (i.e., number of deaths per 1,000 individuals) for the age groups aged 0–4, 5–9, 50–54, 55–59, and 60–64 years.[28] We do not report the results for the age group 10–49 years old, as the mortality rate is too low for this group. We also exclude the elderly (i.e., those aged 65 and over), to prevent our results from being confounded by the effects of welfare benefits paid to elderly persons not covered by the employment-based pension plan, which was introduced in 1961 as a part of the National Pension Plan.[29]

Figures 3.4–3.6 present the time-series patterns for each outcome variable used in this study; they also compare the prefectures whose ratio of uninsured population was greater than the median (27.5 percent) in 1956 (i.e., high-impact prefectures), as well as the others (i.e., low-impact prefectures). Figure 3.4 describes the utilization measures (admission, inpatients, and outpatients) per capita. Health care utilization in high-impact prefectures seems to have started rising more quickly

[28]We also examined gender-specific mortality rates, and found the results to be the same for both men and women.

[29]This benefit was a bail-out measure for those who were already elderly when the National Pension Plan was enacted. The benefit was paid for disabled people aged 65 or older and non-disabled people aged 70 years or older; it was funded by national taxes, not pension premiums. This benefit was not paid for people with other income sources, including employment-based pension benefits. Given that employment-based pensions are often provided with employment-based health insurance, the impact of this welfare benefit is likely to correlate with our measure of the impact of universal health insurance.

than that in low-impact prefectures, following the introduction of universal health insurance; however, the pattern is not very clear. Figure 3.5 shows the supply-side variables (hospitals, clinics, beds, bed occupancy rates, physicians, and nurses); as in Figure 3.4, all variables except the bed occupancy ratio rate increased during the sample period. The bed occupancy rate declined in the late 1950s and increased in the 1960s after the achievement of universal health insurance, probably owing to an increase in the number of inpatients. Also, high-impact prefectures had, on average, more clinics and physicians before 1956 than did the low-impact ones. These two figures underscore the importance of controlling for pre-existing differences across prefectures. Figure 3.6 plots age-specific mortality rates. All age groups experienced a substantial decline in mortality rate over the study period. Also, low-impact prefectures had, on average, higher mortality rates.

Table 3.1 reports the summary statistics of all outcome variables. The mean represents the weighted average of outcomes where population figures are used as weights, as in the regression analysis. We also show the mean for 1956, the reference year, and those of low-impact and high-impact prefectures. Importantly, prefectures whose initial coverage rates were lower (i.e., high-impact prefectures) tended to be more affluent, have more medical resources, and have lower mortality rates prior to the implementation of universal coverage. Thus, any bias on the estimated positive effects of health insurance expansion is likely to be *downward*, because the convergence of economic growth works against finding positive effects.

3.4. Identification Strategy

Our identification strategy is akin to that of Finkelstein (2007). We exploit variations in health insurance coverage rates across prefectures in 1956—which is to say, one year prior to the start of the Four-year Plan to achieve universal coverage by 1961. The basic idea is to compare changes in outcomes in prefectures where the implementation of universal coverage led to a larger increase in health insurance coverage, to prefectures where it had a smaller effect.

Health insurance coverage prior to universal health insurance may not be random. For example, income levels in 1956 tended to be higher in prefectures with more uninsured people. Therefore, it is essential to control for unobserved components that potentially correlate with both the initial coverage rate of health insurance and health care utilization as well as with health outcomes. In fact, Japan experienced a rapid economic growth during the period studied: the speed and timing of such economic growth may have been different across prefectures.[30] We control for differences in the levels of the outcome variables by controlling for prefecture-level fixed effects. Furthermore, we divide the 46 prefectures into 10 regions and control for region–year effects; we also control for convergence of the

[30]The average real GDP growth rate during the 1956–70 period is as high as 9.7 percent. As people became more affluent, their nutrition and sanitary conditions improved. Also, the Tuberculosis Prevention Act enacted in 1951 effectively suppressed tuberculosis, which had been one of the main causes of death in Japan until the early 1950s.

growth rates by including the interaction terms of the initial value of the outcome variable and year dummies.[31]

The basic estimation equation is as follows:

(3.3)

$$Y_{pt} = \alpha_p * 1(pref_p) + \delta_{rt} * 1(year_t) * 1(pref_p \in region_r) + \kappa_t * Y_{p1956} * 1(year_t) \\ + \sum_{t \neq 1956} \lambda_t(impact_p) * 1(year_t) + X_{pt}\beta + \varepsilon_{pt}$$

Subscript p indicates prefecture and t indicates year. α_p represents a prefecture fixed effect; δ_{rt} represents a region-specific year effect; κ_t is meant to capture the differences in the growth of Y due to differences in the initial value; and $impact_p$ is the percentage of the population 1956 in prefecture p without health insurance, as defined in (3.2).

Our parameters of interest are the λ_t's, which represent the coefficients of the interaction terms between year dummies and the percentage of the population without health insurance in 1956. A plot of λ_t's over t shows the flexibly estimated pattern over time in the changes in Y in prefectures where the enforcement of universal coverage had a larger impact on the insurance coverage rate relative to prefectures where it had a smaller impact. If the trend of these λ_t's changes around the 1957–61 period—the phase-in period of universal coverage—such a change

[31]We divide 46 prefectures into the following 10 regions, as defined by the Statistics Bureau: Hokkaido, Tohoku, Kitakanto-Koshin, Minamikanto, Hokuriku, Tokai, Kinki, Chugoku, Shikoku, and Kyushu.

in trend is likely to be attributable to an expansion in health insurance. It is important to note that the equation (3.3) does not impose any *ex ante* restrictions on the timing of the structural trend break; we therefore allow the data to show when changes in the time pattern actually occur.

The covariate X_{pt} controls for potential confounding factors that might have been changing differentially over time across different prefectures. In our basic regression over the 1950–70 period, only the log of the total population and the ratio of the population over 65 years are included, because data pertaining to many of the other control variables are not available for years prior to 1956. As a robustness check, we restrict the sample to the 1956–70 period and include the log of the population, the log of real GNP per capita, local governments' revenue-to-expenditure ratio, and the log of local governments' per-capita real expenditures on health and sanitation. Also, to control for changes in coinsurance rates applied only to the NHI in 1963 and 1968, we add interaction terms between the ratio of population covered by the NHI in the year prior to these changes and dummy variables indicating the period after these changes.

As another robustness check, we include prefecture-specific linear trends in the equation (3.3) for outcome variables whose data are available at least back to 1952. However, note that we have only four to six observations prior to the base year and that the change in insurance coverage was gradual and took place over a four-year period. Thus, the estimated prefecture-specific linear trend might be overfitted; that is, it might pick up part of the effect of the policy change of interest. Given this possibility for overfitting, we do not include prefecture-specific linear trends in our main specification.

Furthermore, following Finkelstein (2007), we take the following two approaches into account for pre-existing trends. First, we calculate changes in λ_t during the first five years following 1956—the year in which the Four-year Plan started—and take the differences with the changes in λ_t in the five years prior to 1956; we calculate $(\lambda_{61} - \lambda_{56}) - (\lambda_{56} - \lambda_{51})$ and their estimated standard errors, to see whether they are statistically significantly distinct from zero. We also estimate $(\lambda_{66} - \lambda_{61}) - (\lambda_{56} - \lambda_{51})$; that is, we repeat the same exercise for the 1961–66 period—the second five-year period following the expansion—to examine long-term effects. A drawback of this approach, however, is that it relies on only three years' worth of data, and thus results can vary, depending on which year is chosen for point-to-point comparisons.

To utilize all available information efficiently, we also estimate the following deviation-from-trend model:

(3.4)

$$Y_{pt} = \alpha_p * 1(pref_p) + \delta_{rt} * 1(year_t) * 1(pref_p \in region_r) + \kappa_t * Y_{p1956} * 1(year_t)$$

$$+ \gamma_{pre} * year_t * impact_p + \gamma_{mid} * 1(year_t \geq 1956) * (year_t - 1956) * impact_p$$

$$+ \gamma_{after} * 1(year_t \geq 1961) * (year_t - 1961) * impact_p + X_{pt}\beta + \varepsilon_{pt}$$

γ_{pre} captures any pre-existing trends that are correlated with health insurance coverage rates in 1956. γ_{mid} represents any trend breaks caused by the massive expansion in health insurance, starting in 1956; and γ_{after} is meant to capture further trend breaks after the achievement of universal coverage. That is, we allow the slope to differ between the expansion period (1956–61) and the lagged period (1961–70). A disadvantage of this approach is that we need to impose *ex ante* restrictions on the timing of trend breaks.

We use prefecture-specific population as a weight in all regressions, to account for substantial variations in population size. We also cluster the standard errors at the prefecture level, to allow for possible serial correlation within prefectures, over time.

Lastly, it is important to clarify how much and in which direction migration could bias our results. First, during the 1950–70 period, there were substantial inflows of working-age people from rural areas to industrialized cities, especially Tokyo and Osaka. Since large cities tended to have low coverage rates in 1956, the prefectures that had a large increase in insurance coverage from 1956 to 1961 also had an increase in its proportion of younger individuals during the same period. Given that younger individuals are less likely to use health care services than older ones, any bias caused by inter-prefecture migration would drive estimates towards zero. Furthermore, as a robustness check, we present results that exclude Tokyo and Osaka from the sample. If inter-prefecture migration were to cause substantial biases, the results excluding Tokyo and Osaka should be different from the results including them; however, as presented in the next section, excluding Tokyo and Osaka does not affect the results. Second, it is possible that sicker people would migrate from a municipality without NHI coverage to one with NHI coverage, within the same prefecture. If so, actual changes in health insurance status might have been larger among healthier people, and thus the impact on health care utilization and health outcomes might be smaller than would have been the case without such migration.

3.5. Results Regarding Utilization

3.5.1. Basic Results

Figure 3.7 plots the estimated λ_t's from equation (3.3) without prefecture-specific linear trends for the following three dependent variables, which serve as measures of health care utilization: log of admissions, inpatient days, and outpatient visits. Because 1956 is the reference year, λ_{56} is set to 0 by definition. Therefore, the coefficient in each year can be interpreted as the relative change in outcomes from 1956 that would have resulted if the expansion in health insurance coverage had increased the coverage ratio by 100 percent, compared to a prefecture where the coverage ratio did not change.

The upper left-hand graph in Figure 3.7 shows the results regarding hospital admissions. Until 1956, there is no pre-existing trend in the λ_t's; at that point, the number of admissions started to grow more quickly in areas in which health insurance expansion had had a larger impact. The estimated λ_{61} and λ_{66} are 0.290

and 0.548, respectively.[32] Given that roughly 28 percent of the total population did not have any health insurance as of 1956, these estimates imply that admissions increased by 8.5 percent ($= exp[0.290 * 0.28] - 1$) over 5 years and 16.6 percent over 10 years, owing to the enforcement of universal health insurance. Inpatient days and outpatient visits show trends very similar to those of admissions: both graphs increase sharply in the late 1950s and stay high until the late 1960s. The magnitude is larger for outpatient visits than for either admissions or inpatient days. The estimated λ_{61} and λ_{66} imply 7.3 and 11.6 percent increases for inpatients days and 12.6 and 25.1 percent increases for outpatient visits by 1961 and by 1966, respectively, both due to the enforcement of universal health insurance.

It is informative to compare our estimates to those from the RAND HIE, although we need to pay considerable attention to differences in the coinsurance systems and other relevant factors between Japan in the 1950s and the United States in the 1970s.[33] Given that the coinsurance rate of the Japan's NHI in Japan was 50 percent at that time, the most comparable case in the RAND experiment HIE is the change in the coinsurance rate from 95 to 50 percent. Manning et al. (1987)

[32] Hereafter, we focus mainly on λ_{61}, that is, changes up to the full achievement of universal health insurance, and λ_{66}, that is, changes within the 10 years following the reference year. The estimated coefficients and standard errors for 1950–70 are available from the authors, upon request.

[33] An important difference is that the RAND HIE set limits on the maximum out-of-pocket expenditures (MDE) that the individual should pay, whereas there was no MDE limit in our case. Since this limit on maximum payment should cause medical utilization to be higher than would otherwise be the case, the estimates from RAND HIE may overestimate the size of the medical expenditures, compared to those in our case.

showed that an individual who moved from 95 to 50 percent coinsurance would increase his or her annual number of face-to-face visits by 11 percent (i.e., from 2.73 to 3.03 visits).[34] Therefore, the RAND HIE suggests that the effect of moving 28 percent of the population from no insurance to having insurance is tantamount to increasing the number of outpatient visits (i.e., face-to-face visits in hospitals) by 3.1 percent (11*0.28). Our estimates show that outpatient visits increased by 12.6 percent in the five years following 1956. Thus, our estimates are about four times larger than individual-level changes in health insurance would have predicted.

3.5.2. Robustness Checks

Table 3.2 presents the robustness checks for our utilization results. To save space, we report only estimates for the interaction terms of 1961 and 1966. To make the results comparable to our basic results, rows (1) and (5) repeat the results from the basic specification.

First, to check whether our results are driven by prefectures with large populations, we exclude Tokyo and Osaka, the two largest prefectures, which together comprised 15 percent of Japan's total population in 1956. Rows (2) and (6) indicate that our results are not driven by these prefectures. Second, to control for other confounding factors that may affect the outcomes, we add the following time-varying variables: the log of the real GNP per capita, converted to 1980 yen;

[34]These figures are taken from Table 2 of Manning et al. (1987). The same figures are presented in Table 3.2 in Newhouse et al. (1993).

the ratio of local governments' revenue to expenditures; and local governments' per-capita real expenditures on health and sanitation. Also, to control for changes in coinsurance rates applied only to the NHI in 1963 and 1968, we add interaction terms between the ratio of population covered by the NHI in the year prior to these changes and dummy variables indicating after these changes. Because most of our additional control variables are available only after 1956, in this specification, we limit the sample to 1956–70.[35] As seen in rows (3) and (7), adding these controls does not significantly change the estimated coefficients. Lastly, rows (4) and (8) show results with prefecture-specific linear trends. Although some of the point estimates change, all λ_t's remain statistically significant.

Furthermore, to check the robustness to pre-existing trends, we compare changes in λ_t during a fixed length of time following the expansion of health insurance coverage, relative to the change in λ_t during the same length of time before the expansion. We do not perform this test for admissions, because data for 1951 are not available. In the first row of Table 3.3, we take a five-year difference in change in the outcome. The increases in both inpatient days and outpatient visits were statistically significant after 1956. The second row in Table 3.3 repeats the same five-year test for 1961–66—namely, the next five-year period—using the same reference period (1951–56). None of the coefficients are statistically significant, although they are all positive. These results indicate that the effect of the expansion

[35]Limiting the sample to 1956–70, in itself, has no impact on the estimated coefficients.

of health insurance on utilization is concentrated in the period when the health insurance coverage was expanding.

Rows (3)–(5) in Table 3.3 show the estimated coefficients of the two slopes in the deviation-from-trend model as equation (3.4). The slope prior to 1956 is not statistically significant and is close to zero for all three outcomes. The coefficients for difference in the slopes before and after 1956 (row (4)) are positive for all three utilization measures, and indicated changes are in the same order as the estimates from other specifications. For example, the coefficient on the first slope for the admissions is interpreted as a 14.7-percent increase ($= exp[0.098 * 5 * 0.28] - 1$) by 1961.[36] In contrast, the estimated coefficients for the second slopes (row 4) are all negative but the magnitude is smaller than the absolute value of the first slopes, which is consistent with positive but flatter slopes after 1961 in Figure 3.7.

3.6. Results vis-à-vis Supply-Side Response

Given the increase in utilization in response to the expansion of health insurance coverage in Japan, the next question is whether the supply side adequately accommodated a drastic increase in the demand for health care. Understanding this supply-side response is particularly important, since one of the major concerns

[36]Note that the estimated coefficient provides only a one-year effect, and roughly 28 percent of Japan's total population had no health insurance coverage as of 1956.

regarding a massive health insurance expansion is a shortage of human capital, including physicians and nurses.[37]

The supply-side response is also interesting from a theoretical perspective. Finkelstein (2007) argues that a market-wide change in health insurance coverage may have larger effects than those implied by individual-level changes in health insurance coverage—especially if the expansion of health insurance coverage sufficiently increases the aggregate demand, so as to induce medical providers to incur the fixed costs associated with building new institutions.

Thus, we begin by testing this hypothesis by estimating the effects of health insurance expansion on the number of medical institutions. The upper left-hand graph of Figure 3.8 plots the estimated λ_t's in equation (3.3) with the log of the number of hospitals as the dependent variable. The estimates for 1961 and 1966 are 0.229 and 0.578, respectively, and both are statistically significant at the conventional level. Therefore, this graph may lead one to believe that the hospitals had increased in size in the areas where utilization had increased.

However, the graph also shows a strong pre-existing trend before 1956. Indeed, as shown in rows (4) and (8) in Table 3.4, once prefecture-specific linear trends are included, the estimated coefficients are no longer significantly positive. Table 3.5 also reports that any positive effects on the number of hospitals disappear when

[37]For example, one of the major concerns related to the Patient Protection and Affordable Care Act in the United States is the shortage of physicians (Association of American Medical College 2010).

pre-existing trends are controlled; therefore, the positive association between the increase in health insurance coverage and the number of hospitals may not be a causal link.

We repeat the same analysis for clinics; the results are shown in the upper right-hand graph in Figure 3.8, as well as the second column in Table 3.4. As shown in the graphs, λ_t's are not estimated very precisely. Moreover, none of the estimates presented in Table 3.4 are statistically significant. We cannot control for any pre-existing trend, because clinic data are available only from 1954; thus, rows (4) and (8) in Table 3.4 are blank and Table 3.5 does not contain a column for clinics. Overall, the response of the number of clinics is small.

Next, we explore the other supply-side response, measured in terms of the supply of beds, physicians, and nurses. The rest of Figure 3.8 shows the estimated λ_t's for the following four outcomes: log of the number of beds, bed occupancy rate, log of the number of physicians, and log of the number of nurses[38]

The graphs in the middle row of Figure 3.8 show that the number of beds in Japan started to increase in the mid-1950s. Compared to 1956, the expansion of health insurance increased the number of beds by 3.4 percent by 1961 and 10.9 percent by 1966.[39] The bed occupancy rate also increased substantially in the late

[38]Because data regarding admissions, inpatient days, and outpatient visits are from hospitals only, we use the number of beds and the physicians and nurses working in hospitals, for the sake of consistency. We have confirmed that the results do not change much if we expand our data to all beds, physicians, and nurses in hospitals and clinics.

[39]Note that the increase in the number of beds at that time was mainly driven by the entry and expansion of private hospitals. It is true that public hospitals also increased its supply of

1950s and then declined in the early 1960s. This pattern suggests that although the number of beds increased in response to an expansion in health insurance coverage, the surge in the number of patients exceeded the increase in the supply of beds during the late 1950s. Unlike the case with the number of hospitals, we do not observe a discernible pre-existing trend for the number of beds. The third column in Table 3.4 and the second column in Table 3.5 confirm that the results are not sensitive to the inclusion of prefecture-specific linear trends or controls for pre-existing trends.

The bottom two graphs in Figure 3.8 show the estimated λ_t's for the number of physicians and nurses. The graph in Figure 3.8 of the number of physicians shows an increase at a slightly slower pace than that of beds, although the estimated λ_t's are not always statistically significant.[40] Pre-1956 data for the number of physicians are available only from 1953; thus, we do not control for prefecture-specific linear trends or pre-existing trends, and rows (4) and (8) in Table 3.4 are blank and Table 3.5 does not contain a column for physicians. The response of the number of nurses is noisier and apparently weak.

beds by 48 percent during the 1956–65 period; yet, the increase rate of beds in private hospitals was in excess of 100 percent in the same period. As pointed by Ikegami (1992), there had been no restrictions on the capital development of private hospitals until 1985, when a ceiling on the number of hospital beds by region was imposed. In contrast, the supply of physicians and nurses are inevitably constrained by the capacity of medical and nursing schools.

[40]This result implied that patient per physician has decreased. While we cannot directly explore this possibility, time spent with each patient may have decreased as well (Garthwaite, 2011).

To recapitulate our results: we find no robust evidence of increases in the number of the hospitals and clinics in response to Japan's expansion of health insurance, but we find evidence of increases in the number of beds. The effect on the number of physicians seems to be positive but noisier than that on beds, whereas the effect on the number of nurses is negligible. These various results are plausible, since it is less costly for existing hospitals to increase their capacity by adding beds than for new hospitals to pay large fixed costs to enter the market. Also, not surprisingly, it is not as easy to increase the numbers of physicians and nurses as to add beds, because the total supply of physicians and nurses are constrained by the capacities of medical and nursing schools.[41]

3.7. Results vis-à-vis Mortality Rates

3.7.1. Basic Results

To complete the picture of the impact of expansion in Japan's health insurance coverage, this section explores whether health insurance benefits the health of insured individuals. On one hand, cheaper access to health care services may improve health outcomes;[42] on the other, if some people are receiving medical care because

[41]In theory, it is also possible that there was excessive capacity before the expansion of health insurance coverage, or that the economics of scale enhanced efficiencies in the provision of medical services, and hence it was not necessary to build new institutions or hire new physicians and nurses.

[42]Another potential benefit to patients is the lower risk of unexpected and high out-of-pocket medical spending. However, we cannot explore this benefit, because data regarding the variance in individual household health care expenditures are not available. Appendix Section A2 shows

of the expansion of health insurance but are not severely ill—or if the expansion of health insurance increases the volume of "unnecessary" treatments (i.e., an *ex post* moral hazard)—there may be no effects on health outcomes. Therefore, the impact of health insurance on health outcomes is *a priori* ambiguous. As a measure of health outcomes, we use age-specific mortality rates.

Figure 3.9 presents the estimated λ_t's in equation (3.3), with the mortality rates of five age groups as the dependent variables. The expansion of health insurance coverage does not reduce the mortality rate among any of the age groups studied. As shown in Table 3.6, the results do not change after excluding Tokyo and Osaka and adding more controls.

However, row (8) in Table 3.6 shows that when prefecture-specific linear trends are controlled, statistically significant negative effects emerge in the late 1960s, except with the 5–9 years age group. At the same time, Table 3.7 shows that controlling for pre-existing trends does not yield any statistically significantly negative effects. Thus, while we cannot conclude from our analysis whether the expansion of health insurance coverage has long-term negative effects on mortality, at least in the short term, there do not seem to be any effects.

that, at least on average, the introduction of universal health insurance did not affect out-of-pocket medical expenditures.

3.7.2. An Event Study: Ibaraki Prefecture

Unlike the other outcome variables, some prefectures publish mortality rates at the municipality level. Since the NHI was introduced at the municipality level, we exploit municipality-level data from Ibaraki prefecture to conduct an event-study analysis. We choose Ibaraki because among the prefectures whose municipality-level mortality data are available, it had a relatively low coverage rate as of 1956 (i.e., 59 percent). A low initial coverage rate means that many municipalities introduced the NHI along with the implementation of universal coverage. Ibaraki is located northeast of Tokyo in the Kanto area, and in 1956, it had a relatively low per-capita GNP (37th among 46 prefectures) and high mortality rates (about the 5th to 15th-largest, depending on the age group).

The data are taken from the Ibaraki prefecture Statistical Book, which provides the number of NHI enrollees, population figures, and the number of deaths in each municipality. We exclude municipalities that merged during the 1956–61 period, because these mergers make it difficult to identify the year in which the NHI was introduced or fully implemented; such excluded municipalities include Mito city, the capital city of the prefecture. Then, for the remaining 73 municipalities, we consider the year of full NHI implementation as the year in which the number of NHI enrollees exceeded 90 percent of the number of 1961 enrollees. Forty-one municipalities implemented the NHI fully during the 1956–61 period.

We define the mortality rate as the number of deaths per 1,000 people. Although data on NHI participation are available from 1955, the number of deaths and the population of each municipality are available only from 1957. Thus, we limit our analysis to the 1957–65 period.[43] We then estimate the following equation:

$$Y_{mt} = \alpha_m + \sum_{T=-4}^{8} \pi_T(\tau_{mt} = T) + \gamma_m t + \varepsilon_{mt} \tag{3.5}$$

where Y_{mt} is the mortality rate of municipality m in year t. τ_{mt} is time to the year when municipality m fully implemented the NHI measured by years, and π_T is the changes in the mortality rate, relative to the year in which the municipality fully implemented the NHI.[44] Furthermore, α_m represents municipality fixed effects and γ_m represents municipality-specific linear trends.[45] Standard errors are estimated with clustering by municipality so that ε_{mt} can be correlated within municipalities across time.

Figure 3.10 plots the estimated π_T's. It shows that there was no change in mortality as a result of the full NHI implementation. Therefore, we conclude

[43]Although data after 1965 are available, we do not extend our data period, because across-municipality mobility would attenuate the estimates more severely as we move farther from the base year.

[44]Using the year in which the NHI was introduced (but not necessarily in which it was fully implemented) yields almost the same results, except that five municipalities are excluded because they had partially introduced NHI before 1956.

[45]We have also tried prefecture-wide year dummies instead of municipality-specific linear trends. The results are qualitatively the same.

that although there might have been some modest effects emerging with a lag of approximately 10 years, the expansion in health insurance coverage in Japan did not affect the mortality rate, at least within the several years following its implementation.

3.7.3. Cause-specific Mortality

Neither the basic specification using prefecture-level data nor the event study using municipality-level data show any short-term decline in mortality rates. This lack of decline in mortality in the short term may be because individuals with acute, life-threatening, and treatable health conditions had previously sought care at hospitals, even when they lacked health insurance and thus incurred costs at their own expense. Although there was no public aid for the uninsured, mutual aid from blood relatives and the local community could have supported poor, uninsured patients.

To examine such a possibility, we examine the cause-specific mortality of diseases that were considered treatable at that time, including pneumonia, bronchitis, gastritis, and duodenitis.[46] If those who could have been saved with appropriate treatment did not have access to care owing to a lack of health insurance coverage, the mortality rates of these treatable diseases should have fallen more in the prefectures that were more greatly affected by health insurance expansion. However,

[46] At that time, hospitals could effectively treat only these short-term, acute illnesses, rather than chronic illness such as cancer and cardiovascular disease.

as shown in Figure 3.11, we find no statistically significant reduction in the number of deaths as a result of these treatable diseases.[47]

3.8. Conclusion

We have estimated the impact of a massive expansion in Japan's health insurance program on health care utilization and health outcomes in that country. We find substantial increases in health care utilization—increases much larger than those implied by micro-level estimates from the RAND HIE, among others. We then investigate why we find such larger effects, and differential supply-side responses, as argued in Finkelstein (2007). While we do not find that the expansion of health insurance induced the market entries of hospitals and clinics—which would necessarily incur large fixed costs among those facilities—we find increases in the number of beds, which may be less costly than market entries.

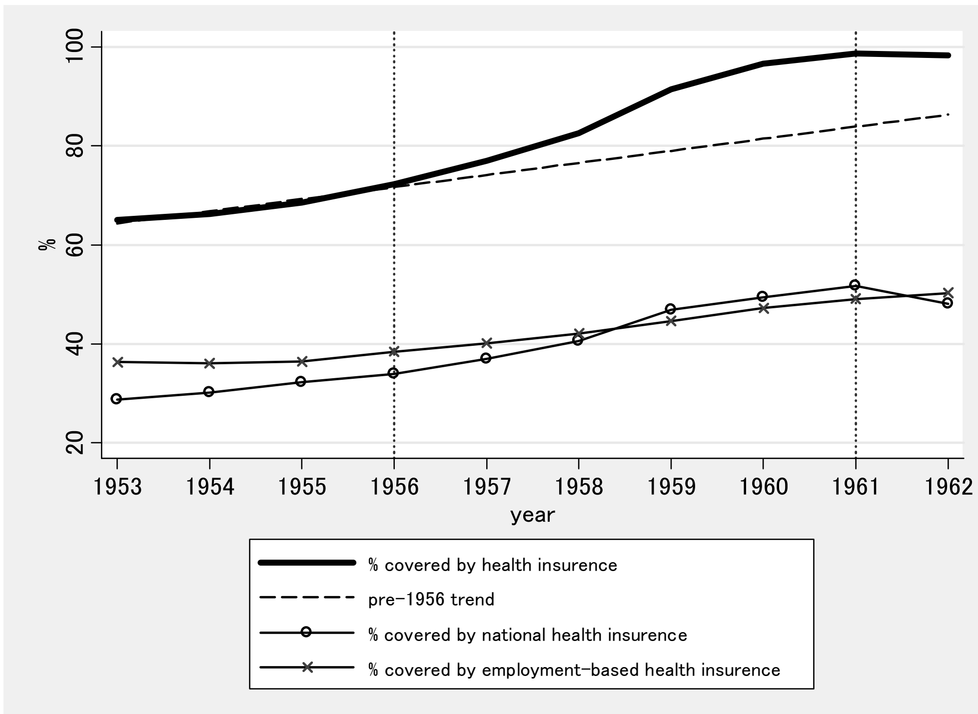
Despite the increase in health care utilization in Japan during the period under examination, we find no strong evidence of improved health outcomes, at least in the short term. Admittedly, our results vis-à-vis health outcomes are limited to mortality, and thus it is possible that the introduction of universal health insurance reduced the morbidity of nonfatal diseases. Nonetheless, universal health insurance

[47]Another possibility is that the sudden increase in demand lowered the quality of health care services. Because health care utilization increased dramatically—whereas the number of physicians and nurses did not fully "catch up"—the expansion of health insurance might have reduced the number of physicians and nurses per patient. Although we cannot directly measure the quality of medical treatment, this overcrowding may have lowered the overall quality of health care services.

is unlikely to be the main factor explaining Japan's drastic improvement in life expectancy in the 1960s, at least in the short term.

Another limitation of the current study is that we cannot conclude from our results that universal health insurance does not improve social welfare. Our limited data do not allow us to explore the decline in the risk of sudden out-of-pocket medical expenditures, which is another important benefit from health insurance. Rather, the takeaway from our empirical results is that a large expansion in health insurance coverage will increase health care utilization, regardless of whether it improves health outcomes, and the magnitude of the effect will be much larger than that predicted from individual-level changes in insurance status. Therefore, countries planning to introduce universal health insurance need to set aside sufficient financial resources for the anticipated surge in health care expenditures. Also, our results may indicate that slow supply-side response may constrain the ability of the health care system to meet the increased demand resulting from expansions in coverage.

Figure 3.1: National Time Series of Health Insurance Coverage Rates



Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved.
Source: Social Security Year Book (1952-57) and Annual Report on Social Security Statistics (1958-1964).

Figure 3.2: % of Population without Any Health Insurance as of April 1956

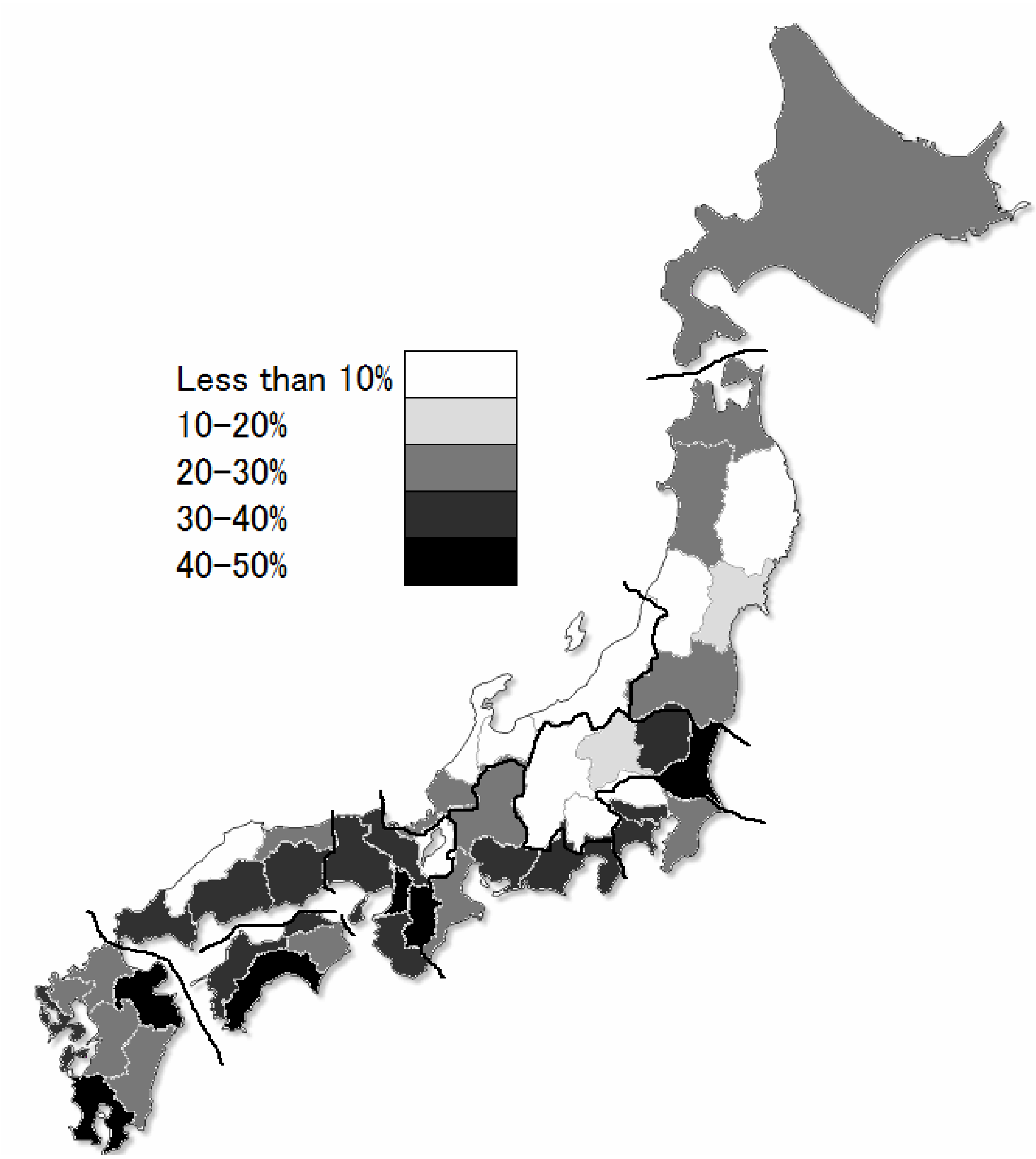


Figure 3.3: Scatter Plots of Changes in Per Capita GNP and Health Insurance Coverage Rate

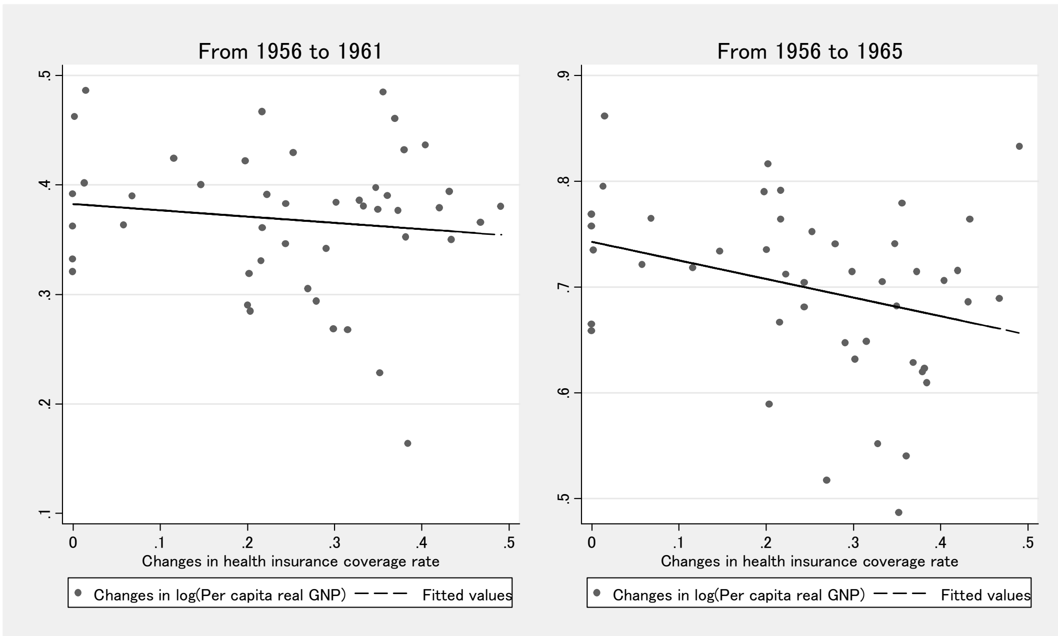
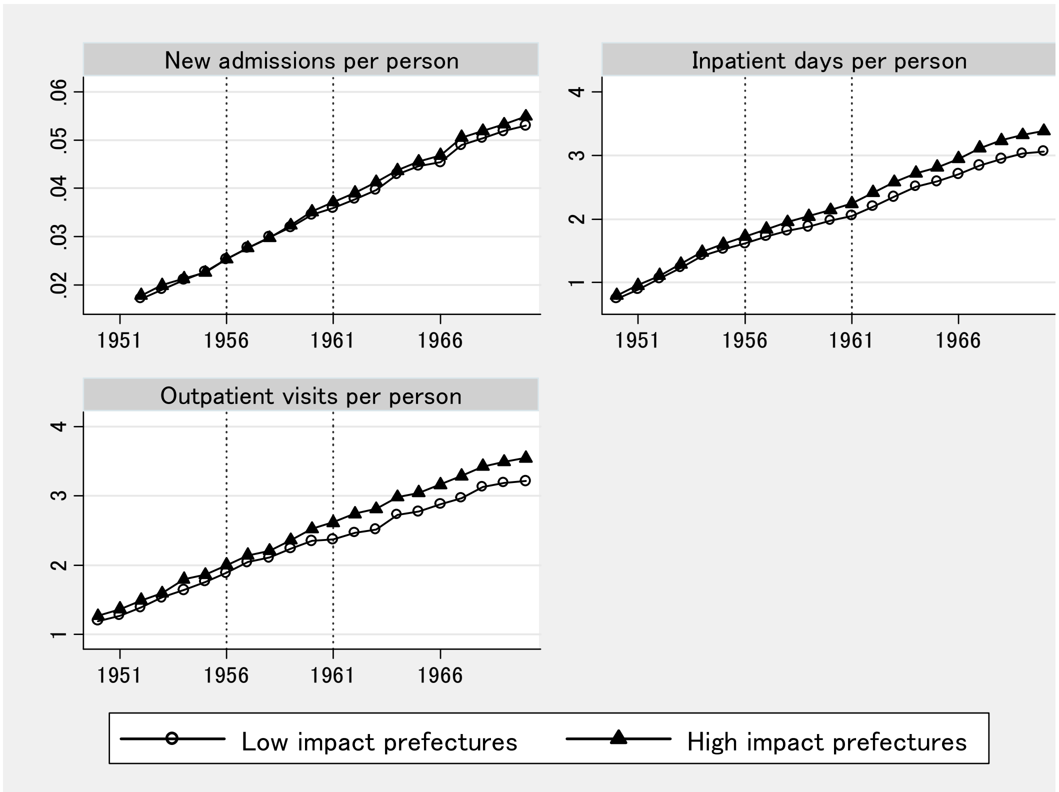
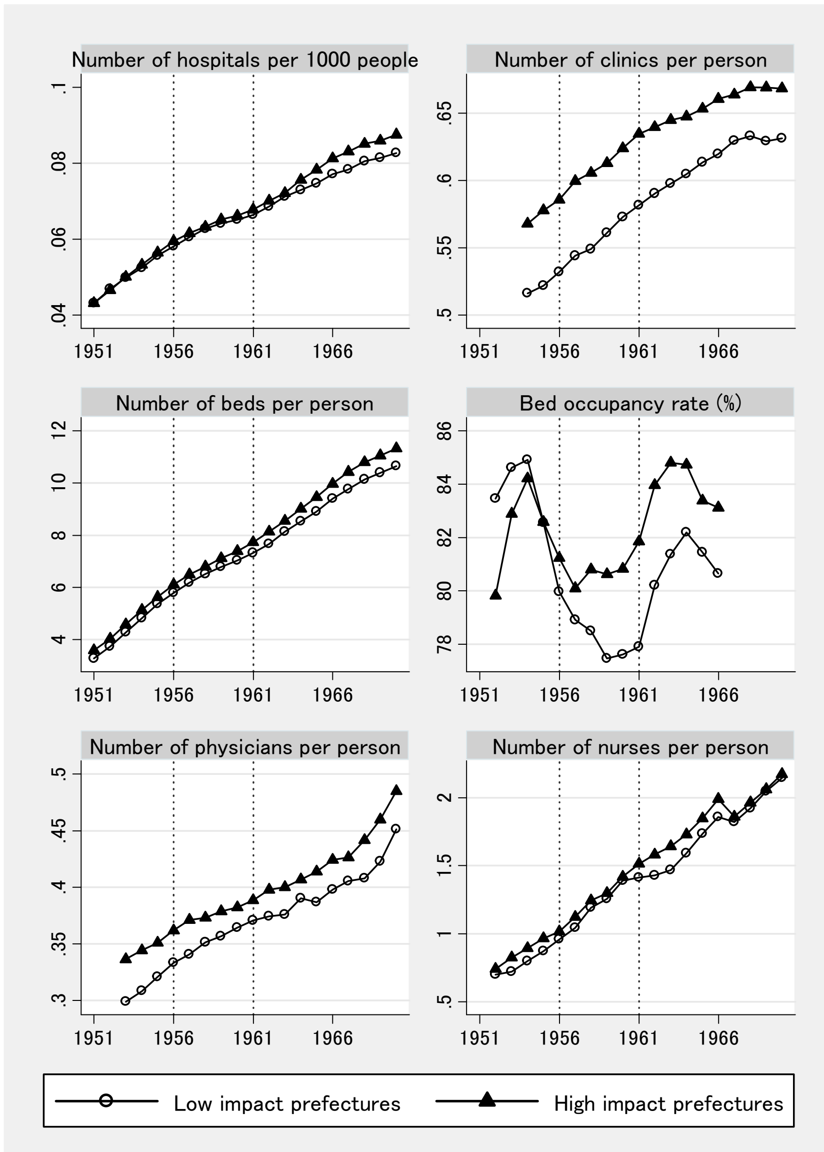


Figure 3.4: Time Series of Health Care Utilization



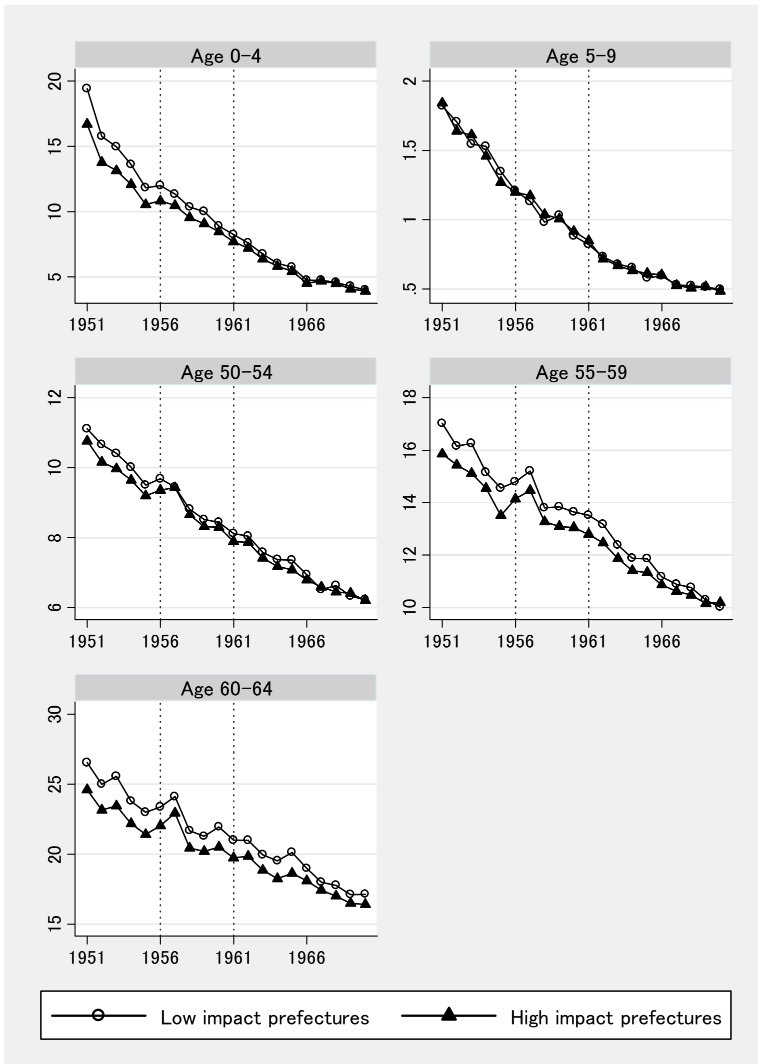
Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Low impact prefectures are prefectures whose rate of uninsured population was less than 27.5% in 1956, i.e. lower than the median.

Figure 3.5: Time Series of Per Capita Supply of Health Care



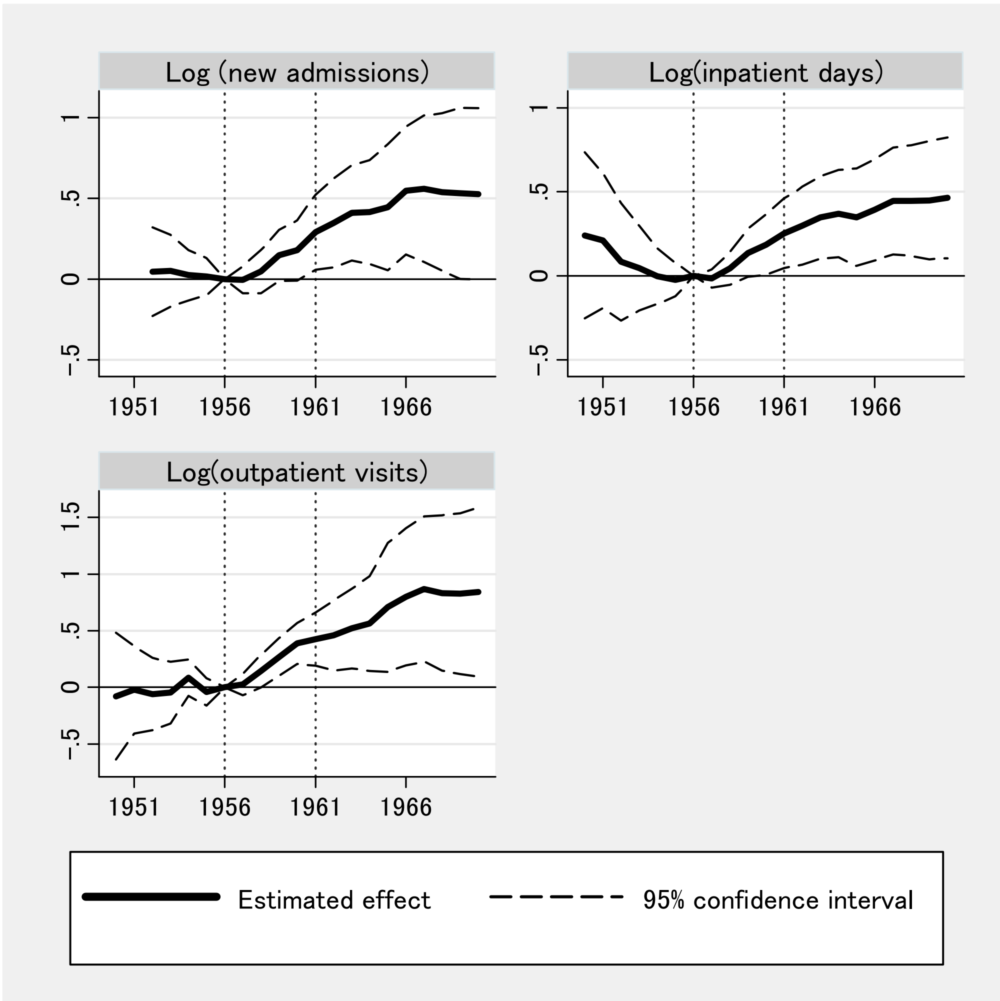
Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Low impact prefectures are prefectures whose rate of uninsured population was less than 27.5% in 1956, i.e. lower than the median.

Figure 3.6: Time Series of Age Specific Mortality Rates



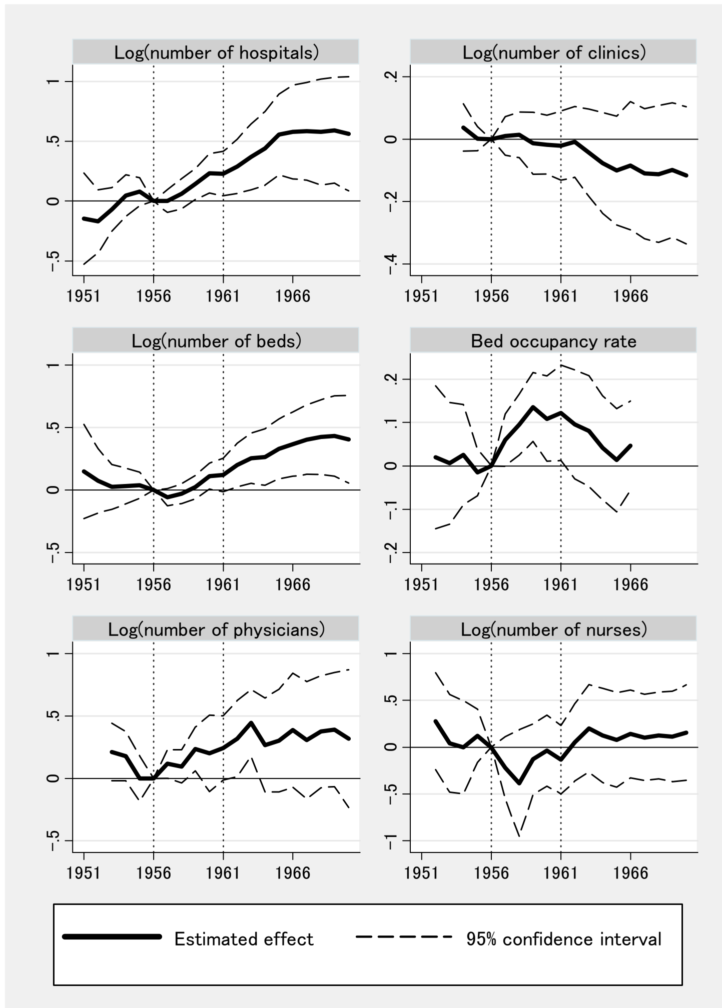
Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Low impact prefectures are prefectures whose rate of uninsured population was less than 27.5% in 1956, i.e. lower than the median.

Figure 3.7: Effect of Health Insurance Coverage on Healthcare Utilization



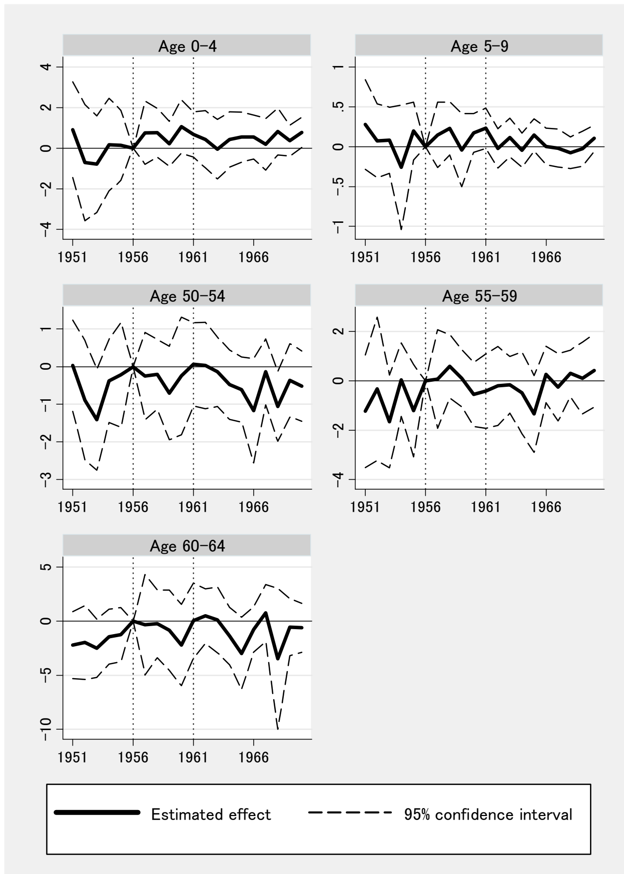
Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Regressions on which these graphs are based include prefecture-fixed effects, region-specific year effects, interactions between year dummies and the value of the dependent variable as of 1956, log population and the ratio of over 65 in population. Standard errors are clustered by prefecture.

Figure 3.8: Effect of Health Insurance Coverage on Supply of Health Care



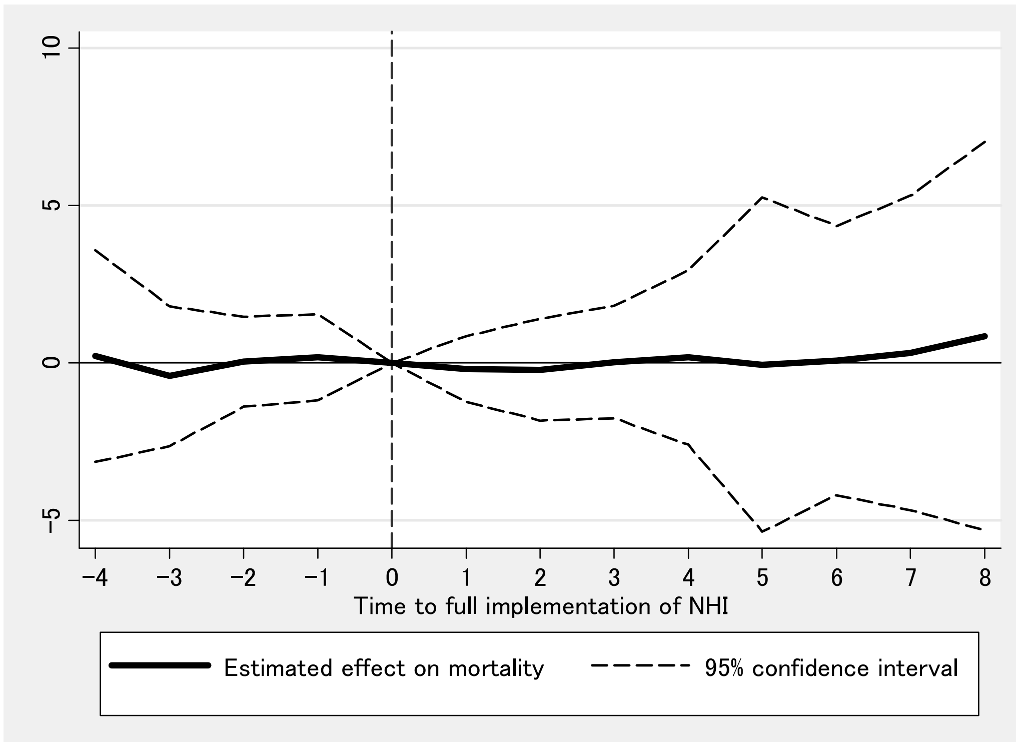
Note: Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Regressions on which these graphs are based include prefecture-fixed effects, region-specific year effects, interactions between year dummies and the value of the dependent variable as of 1956, log population and the ratio of over 65 in population. Standard errors are clustered by prefecture.

Figure 3.9: Effect of Health Insurance Coverage on Age-Specific Mortality Rates



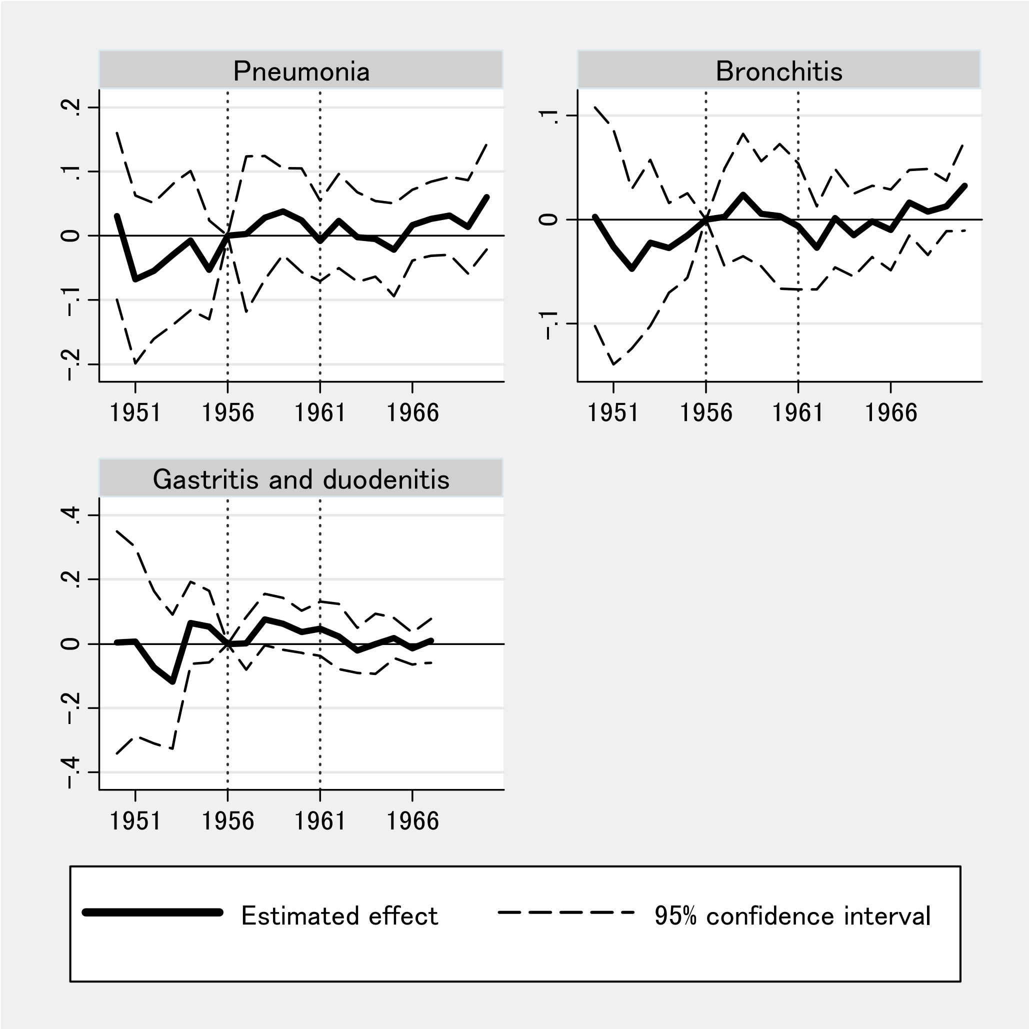
Note: Mortality rate is number of deaths per 1000 population. Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Regressions on which these graphs are based include prefecture-fixed effects, region-specific year effects, interactions between year dummies and the value of the dependent variable as of 1956, log population and the ratio of over 65 in population. Standard errors are clustered by prefecture.

Figure 3.10: Mortality Rates by Time to Full Implementation of the NHI



Note: The sample includes 41 municipalities in Ibaraki prefecture that fully implemented NHI during the period of 1957-1961. Regressions on which these graphs are based include municipality fixed effects and municipality-specific linear trends. Standard errors are clustered by municipalities.

Figure 3.11: Effect of Health Insurance Coverage on Mortality Rates by Treatable Diseases



Note: Mortality rate is number of deaths per 1000 population. The data for gastritis and duodenitis are not available after 1968 because of the changes in classification. Two vertical lines indicate 1956, the reference year, and 1961, the year in which universal health insurance was achieved. Regressions on which these graphs are based include prefecture-fixed effects, region-specific year effects, interactions between year dummies and the value of the dependent variable as of 1956, log population and the ratio of over 65 in population.

Table 3.1: Mean of Dependent and Control Variables

Variable	Obs	Available period	Whole period	All prefectures in 1956	High impact prefectures in 1956	Low impact prefectures in 1956
Admission (thousands)	874	1952-70	148.5	91.5	118.4	48.0
Inpatient days (thousands)	966	1950-70	7517.1	5610.1	7087.2	3224.9
Outpatient visits (thousands)	966	1950-70	9744.5	7322.9	9388.6	3987.3
Hospitals	920	1951-70	215.4	180.9	223.3	112.5
Clinics	782	1954-70	2455.6	1911.7	2494.0	971.4
Number of beds in hospitals	828	1951-70	27619.7	19439.1	24420.5	11395.3
Bed occupancy rate (%)	690	1952-66	82.1	81.1	81.6	80.2
Number of physicians in hospitals	828	1953-70	1516	1349.7	1739.1	720.9
Number of nurses in hospitals	874	1952-70	5884.6	3649.9	4774.8	1833.4
Mortality rate: age 0-4	920	1951-70	8.1	10.6	9.9	11.8
Mortality rate: age 5-9	920	1951-70	0.9	1.2	1.1	1.2
Mortality rate: age 50-54	920	1951-70	8.2	9.6	9.4	9.8
Mortality rate: age 55-59	920	1951-70	12.9	14.5	14.2	15.0
Mortality rate: age 60-64	920	1951-70	20.5	22.8	22.3	23.7
Population (thousands)	966	1950-70	3325.8	2939.6	3607.4	1861.2
Population over 65 (%)	966	1950-70	4.9	3.9	3.8	4.2
Real GNP per capita (1980 thousand yen)	736	1955-70	700.7	378.9	415.0	320.5
Real local gov. expenditure on health and sanitation (1980 thousand yen)	690	1956-70	5.6	1.8	1.9	1.5
Local gov. expenditure to revenue ratios	690	1956-70	1.03	1.02	1.03	1.01
Real medical expenditures per person by NHI (1000 yen in 1980 price)	644	1957-70	20.1	6.7 (in 1957)	6.8 (in 1957)	6.6 (in 1957)

Note: Mortality rate is the number of deaths per 1000 population. High impact prefectures are prefectures whose uninsured rate was 27.5% or higher in 1956. Low impact prefectures are prefectures whose uninsured rate was lower than 27.5% in 1956. 27.5% is the median uninsured rate in 1956.

Table 3.2: Robustness Checks for Utilization Outcomes

	λ in 1961		
Dependent variable:	Log(admissions)	Log(inpatient days)	Log(outpatient visits)
(1) λ shown in Figure 7	0.290**	0.253**	0.426***
	[0.116]	[0.103]	[0.116]
(2) Excluding Tokyo and Osaka	0.267**	0.218**	0.389***
	[0.116]	[0.108]	[0.132]
(3) More controls	0.279**	0.265**	0.412***
(sample period: 1956-1970)	[0.105]	[0.104]	[0.130]
(4) Prefecture specific	0.192**	0.449***	0.409***
linear trends	[0.073]	[0.064]	[0.110]
	λ in 1966		
Dependent variable:	Log(admissions)	Log(inpatient days)	Log(outpatient visits)
(5) λ shown in Figure 7	0.548***	0.392**	0.800**
	[0.196]	[0.150]	[0.301]
(6) Excluding Tokyo and Osaka	0.459**	0.302*	0.637**
	[0.195]	[0.157]	[0.294]
(7) More controls	0.567***	0.412***	0.884***
(sample period: 1956-1970)	[0.188]	[0.149]	[0.272]
(8) Prefecture specific	0.403***	0.786***	0.748***
linear trends	[0.066]	[0.089]	[0.077]

Note: Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3.3: Controlling for Pre-existing Trend: Utilization Outcomes

Dependent variable:	Log(admissions)	Log(inpatient days)	Log(outpatient visits)
$(\lambda_{61}-\lambda_{56})-(\lambda_{56}-\lambda_{51})$	--	0.462**	0.481**
	--	[0.203]	[0.158]
$(\lambda_{66}-\lambda_{61})-(\lambda_{56}-\lambda_{51})$	--	0.349	0.353
	--	[0.221]	[0.261]
Slope prior to 1956	-0.028	-0.048	0.006
	[0.032]	[0.042]	[0.043]
(Slope in 1956-1961) - (Slope prior to 1956)	0.098**	0.117**	0.085**
	[0.046]	[0.048]	[0.038]
(Slope in 1961-1970) - (Slope prior to 1961)	-0.038*	-0.044*	-0.038
	[0.022]	[0.023]	[0.048]

Note: Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. The first two rows for Log(admissions) are blank because the data for 1951 are not available.

Table 3.4: Robustness Checks for Supply of Health Care

	λ in 1961					
Dependent variable:	Log(hospitals)	Log(clinics)	Log(beds)	BOR	Log(physicians)	Log(nurses)
(1) λ shown in Figure 8	0.229**	-0.021	0.121*	0.122**	0.243*	-0.132
	[0.092]	[0.055]	[0.067]	[0.055]	[0.128]	[0.181]
(2) Excluding Tokyo and Osaka	0.183*	-0.031	0.085	0.115**	0.241*	-0.24
	[0.092]	[0.053]	[0.070]	[0.057]	[0.130]	[0.261]
(3) More controls (sample period: 1956-1970)	0.205**	-0.013	0.130*	0.128**	0.168	-0.102
	[0.091]	[0.055]	[0.069]	[0.061]	[0.118]	[0.186]
(4) Prefecture specific	-0.017	--	0.075*	0.351***	--	-0.146
linear trends	[0.060]	--	[0.039]	[0.059]	--	[0.218]
	λ in 1966					
Dependent variable:	Log(hospitals)	Log(clinics)	Log(beds)	BOR	Log(physicians)	Log(nurses)
(5) λ shown in Figure 8	0.578***	-0.085	0.368***	0.047	0.387*	0.142
	[0.194]	[0.102]	[0.128]	[0.051]	[0.226]	[0.232]
(6) Excluding Tokyo and Osaka	0.509**	-0.096	0.304**	0.022	0.387*	-0.068
	[0.201]	[0.096]	[0.135]	[0.061]	[0.225]	[0.250]
(7) More controls (sample period: 1956-1970)	0.622***	-0.083	0.384***	0.070	0.372*	0.182
	[0.164]	[0.092]	[0.127]	[0.061]	[0.203]	[0.233]
(8) Prefecture specific	0.145	--	0.299***	0.443***	--	0.157
linear trends	[0.089]	--	[0.065]	[0.083]	--	[0.225]

Note: BOR stands for bed occupancy rate. Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. Rows (4) and (8) for log(clinics) and log(physicians) are left blank because available data prior to 1956 are limited to less than 4 years for these two outcomes.

Table 3.5: Controlling for Pre-existing Trend: Supply of Health Care

Dependent variable:	Log(hospitals)	Log(beds)	BOR	Log(nurses)
$(\lambda_{61}-\lambda_{56})-(\lambda_{56}-\lambda_{51})$	0.084	0.270	--	--
	[0.195]	[0.199]	--	--
$(\lambda_{66}-\lambda_{61})-(\lambda_{56}-\lambda_{51})$	0.204	0.397*	--	--
	[0.273]	[0.212]	--	--
Slope prior to 1956	0.030	-0.037	0.003	-0.098*
	[0.031]	[0.035]	[0.022]	[0.058]
(Slope prior to 1956) - (Slope in 1956-1961)	0.023	0.078*	0.019	0.125
	[0.037]	[0.045]	[0.020]	[0.081]
(Slope prior to 1961) - (Slope in 1961-1970)	-0.011	-0.005	-0.046***	-0.003
	[0.037]	[0.021]	[0.016]	[0.059]

Note: BOR stands for bed occupancy rate. Clinics and physicians are excluded from the analyses because of the lack of pre-1956 data. Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively. The first two rows for BOR and log(nurses) are blank because the data for 1951 are not available.

Table 3.6: Robustness Checks for Age Specific Mortality

	λ in 1961				
Dependent variable:	Age 0-4	Age 5-9	Age 50-54	Age 55-59	Age 60-64
(1) λ shown in Figure 9	0.681	0.231*	0.057	-0.422	0.042
	[0.552]	[0.126]	[0.550]	[0.747]	[1.728]
(2) Excluding Tokyo and Osaka	0.222	0.311**	0.286	0.154	1.505
	[0.563]	[0.150]	[0.558]	[0.740]	[1.504]
(3) More controls (sample period: 1956-1970)	0.614	0.200	0.068	-0.358	-0.253
	[0.548]	[0.132]	[0.580]	[0.860]	[1.790]
(4) Prefecture specific linear trends	-0.485	0.213	-0.239	-1.219	-1.119
	[0.783]	[0.168]	[0.606]	[0.816]	[1.588]
	λ in 1966				
Dependent variable:	Age 0-4	Age 5-9	Age 50-54	Age 55-59	Age 60-64
(5) λ shown in Figure 9	0.547	0.003	-1.175*	0.260	-0.758
	[0.538]	[0.114]	[0.689]	[0.572]	[1.037]
(6) Excluding Tokyo and Osaka	0.176	0.075	-0.797	0.724	0.102
	[0.579]	[0.125]	[0.637]	[0.494]	[0.997]
(7) More controls (sample period: 1956-1970)	0.675	-0.048	-1.041	0.487	-0.912
	[0.527]	[0.116]	[0.719]	[0.585]	[1.083]
(8) Prefecture specific linear trends	-1.490**	-0.063	-1.832***	-1.279*	-3.028***
	[0.636]	[0.182]	[0.615]	[0.758]	[0.886]

Note: Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table 3.7: Controlling for Pre-existing Trend: Age Specific Mortality

Dependent variable:	Age 0-4	Age 5-9	Age 50-54	Age 55-59	Age 60-64
$(\lambda_{61}-\lambda_{56})-(\lambda_{56}-\lambda_{51})$	1.594	0.510	0.082	-1.655	-2.168
	[1.514]	[0.328]	[0.917]	[1.244]	[2.551]
$(\lambda_{66}-\lambda_{61})-(\lambda_{56}-\lambda_{51})$	0.779	0.051	-1.207	-0.551	-3.010
	[1.054]	[0.309]	[0.830]	[1.449]	[1.829]
Slope prior to 1956	0.067	-0.019	0.071	0.247	0.406
	[0.240]	[0.041]	[0.125]	[0.220]	[0.317]
(Slope prior to 1956) - (Slope in 1956-1961)	0.004	0.037	-0.047	-0.346	-0.444
	[0.310]	[0.055]	[0.159]	[0.283]	[0.426]
(Slope prior to 1961) - (Slope in 1961-1970)	-0.078	-0.035	-0.082	0.168	-0.027
	[0.188]	[0.037]	[0.163]	[0.166]	[0.258]

References

[1] **Acemouglu, Daron, David Autor, and David Lyle.** "Women War and Wages: The Impact of Female Labor Supply on the Wage Structure at Mid-Century," *Journal of Political Economy*, 112(3): 497–551, 2004.

[2] **Acemoglu, Daron, and Amy Finkelstein.** "Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector," *Journal of Political Economy*, Vol.116, No.5, pp837-880, 2008.

[3] **Almond, Douglas, Joseph J. Doyle Jr., Amanda E. Kowalski, and Heidi Williams.** "Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns", *Quarterly Journal of Economics*, Vol. 125, No. 2: 591-634, 2010.

[4] **Almond, Douglas, Hilary W. Hoynes and Diane Whitmore Schanzenbach.** "Inside the War on Poverty: The Impact of Food Stamps on Birth Outcomes", *Review of Economics and Statistics*, 93:2, 387-403, 2011.

[5] **Anderson, Gerard F. and Peter Sotir Hussey.** "Population Aging: A Comparison among Industrialized Countries," *Health Affairs*, 19, No.3:191-203, 2000.

[6] **Anderson, Michael L., Carlos Dobkin, and Tal Gross.** "The Effect of Health Insurance Coverage on the Use of Medical Services," *American Economic Journal: Economic Policy*, forthcoming.

[7] **Association of American Medical College.** "Physician Shortages to Worsen Without Increases in Residency Training," www.aamc.org/download/150584/data/physician_shortages_factsheet.pdf (accessed May 31, 2011), 2010.

[8] **Barreca, Alan, Melanie Guldi, Jason M. Lindo, and Glen R. Waddell.** "Heaping-Induced Bias in Regression-Discontinuity Designs," *NBER Working Paper* No. 17408, 2011.

[9] **Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan.** "How Much Should We Trust Differences-in-Differences Estimates?," *Quarterly Journal of Economics*, 119(1), pp 249-275, 2004.

[10] **Bharadwaj, Prashant, and Christopher Neilson.** "Early Life Health Interventions and Academic Achievement," *Mimeo*, 2011.

[11] **Bhattacharya, Jayanta, William B. Vogt, Aki Yoshikawa, and Toshitaka Nakahara.** "The Utilization of Outpatient Medical Services in Japan," *Journal of Human Resources*, Vol. 31, No. 2, 450-476, 1996.

[12] **Bitler, Marianne, Jonah Gelbach, and Hilary Hoynes.** "What Mean Impacts Miss: Distributional Effects of Welfare Reform Experiments," *American Economic Review*, 96 (4), 988-1012, 2006.

[13] **Camacho, Adriana, and Emily Conover.** "Manipulation of Social Program Eligibility." *American Economic Journal: Economic Policy*, 3(2): 41–65, 2011.

[14] **Card, David, Carlos Dobkin, and Nicole Maestas.** "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence from Medicare," *NBER Working Paper* No. 10365, 2004.

[15] **Card, David, Carlos Dobkin, and Nicole Maestas.** "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare," *American Economic Review*, 98(5), 2242-2258, 2008.

[16] **Card, David, Carlos Dobkin, and Nicole Maestas.** "Does Medicare Save Lives?" *Quarterly Journal of Economics*, 124(2): 597-636, 2009.

[17] **Card, David, David Lee, and Zhuan Pei.** "Quasi-Experimental Identification and Estimation in the Regression Kink Design," *Industrial Relations Section Working Paper* No. 553, 2009.

[18] **Carneiro, Pedro, Katrine V. Loken, and Kjell G. Salvanes.** “A Flying Start? Long Term Consequences of Maternal Time Investments in Children During Their First Year of Life,” *IZA Discussion Paper* No. 5362, 2010.

[19] **Carpenter, Christopher, and Carlos Dobkin.** "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age," *American Economic Journal: Applied Economics*, 1(1), 164-182, 2009.

[20] **Cataife, Guido and Charles J. Courtemanche.** "Is Universal Health Care in Brazil Really Universal? ", *NBER Working Paper* No. 17069, 2011.

[21] **Centers for Medicare and Medicaid Services.** *Medicare Enrollment Reports.* www.cms.gov/MedicareEnRpts/Downloads/10All.pdf (accessed December 28, 2011), 2010.

[22] **Centers for Medicaid and Medicare Services.** *National Health Expenditures.* www.cms.hhs.gov/NationalHealthExpendData/
04_NationalHealthAccountsAgePHC.asp. (accessed Aug 21, 2011), 2005.

[23] **Chandra, Amitabh, Jonathan Gruber, and Robin McKnight.** “Patient Cost-Sharing, Hospitalization Offsets, and the Design of Optimal Health Insurance for the Elderly,” *NBER Working Paper* No. 12972, 2008.

[24] **Chandra, Amitabh, Jonathan Gruber, and Robin McKnight.** “Patient Cost-Sharing and Hospitalization Offsets in the Elderly,” *American Economic Review*, 100:1, 193-213, 2010.

[25] **Chay, Kenneth Y., Daeho Kim, and Shailender Swaminathan.** “Medicare, Hospital Utilization and Mortality: Evidence from the Program’s Origins," *Mimeo*, 2010.

[26] **Chang, Simon.** "The Effect of Taiwan’s National Health Insurance on Mortality of the Elderly: Revisited", *Health Economics*, 2011.

[27] **Chen L, Yip W, Chang MC, Lin HS, Lee SD, Chiu YL, Lin YH.** "The Effects of Taiwan’s National Health Insurance on Access and Health Status of the Elderly", *Health Economics*, 16(3): 223–242, 2007.

[28] **Chen, Yuyu and Ginger Zhe Jin.** "Does Health Insurance Coverage Lead to Better Health and Educational Outcomes? Evidence from Rural China", *NBER Working Paper* No. 16417, 2010.

[29] **Chetty, Raj, John N. Friedman, Tore Olsen and Luigi Pistaferri**. "Adjustment Costs, Firm Responses, and Labor Supply Elasticities: Evidence from Danish Tax Records," *Quarterly Journal of Economics,* forthcoming.

[30] **Chou, S.-Y., M. Grossman, and J.-T. Liu.** "The Impact of National Health Insurance on Birth Outcomes: A Natural Experiment in Taiwan," *NBER Working Paper* No. 16811, 2011.

[31] **Congressional Budget Office.** *CBO's 2011 Long-Term Budget Outlook*. www.cbo.gov/ftpdocs/122xx/doc12212/06-21-Long-Term_Budget_Outlook.pdf. (accessed Nov 1, 2011), 2011.

[32] **Coulam, Robert, and Gary Gaumer**."Medicare's Prospective Payment System: A Critical Appraisal." *Health Care Financing Rev.* (suppl.): 45-77, 1991.

[33] **Cunningham, P. J., and O'Malley, A.,S.** "Do Reimbursement Delays Discourage Medicaid Participation by Physicians?" *Health Affairs*, Web Exclusives, Vol. 28, No. 1-2, W17-W28, 2009.

[34] **Currie, J., and J. Gruber.** "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." *Quarterly Journal of Economics,* 111: 431-466, 1996a.

[35] **Currie, J., and J. Gruber.** "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women." *Journal of Political Economy,* 104: 1263-1296, 1996b.

[36] **Cutler, David M.** "Cost Shifting or Cost Cutting?: The Incidence of Reductions in Medicare Payments," *Tax Policy and the Economy*, Vol. 12, 1-27, 1998.

[37] **Cutler, David M, and Ellen Meara.**"The Technology of Birth: Is It Worth It?" In Frontiers in Health Policy Research, Volume 3. , ed. Alan M. Garber, 33-68, 2000.

[38] **Cutler, David M, and Richard Zeckhauser.** "The Anatomy of Health Insurance." *In Handbook of Health Economics*, vol. 1A, edited by Anthony Culyer and Joseph Newhouse. Amsterdam: Elsevier, 2000.

[39] **Dafny, Leemore S.** "How Do Hospitals Respond to Price Changes?," *American Economic Review*, Vol. 95, No. 5, pp. 1525-1547, 2005.

[40] **Damrongplasit, Kannika and Glenn A. Melnick.** "Early Results From Thailand's 30 Baht Health Reform: Something to Smile About", *Health Affairs*, 28, no.3:w457-w466, 2009.

[41] **Dong, Yingying.** "Jumpy or Kinky? Regression Discontinuity without the Discontinuity," *Mimeo*, 2010.

[42] **Dow, W. and K. Schmeer.** "Health Insurance and Child Mortality in Costa Rica," *Social Science and Medicine*, 57(6): 975-986, 2003.

[43] **Dranove, David**. "Rate-Setting by Diagnosis Related Groups and Hospital Specialization," *RAND Journal of Economics*, 18(3), pp. 417-27, 1987.

[44] **Duggan, Mark G.** "Hospital Ownership and Public Medical Spending," *Quarterly Journal of Economics*, 115(4), pp. 1343-73, 2000.

[45] **Ellis, Randall.** "Rational Behavior in the Presence of Coverage Ceilings and Deductibles," *The RAND Journal of Economics*, 17(2):158-175, 1986.

[46] **Ellis, Randall P. and McGuire, Thomas G.** "Hospital Response to Prospective Payment: Moral Hazard, Selection, and Practice-Style Effects," *Journal of Health Economics*, 15(3), pp. 257-77, 1996.

[47] **Engelhardt, Gary V., and Gruber, Jonathan.** "Medicare Part D and the Financial Protection of the Elderly," *American Economic Journal: Economic Policy*, 3(4): 77-102, 2011.

[48] **Evans R.** "Supply-induced Demand: Some Empirical Evidence and Implications", in: M.Perlman, ed., *The Economics of Health and Medical Care*, Macmillan, London, pp.162-173, 1974.

[49] **Fan, Jianqing, and Irene Gijbels.** *Local Polynomial Modeling and its Applications.* London: Chapman and Hall, 1996.

[50] **Finkelstein, Amy.** "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare," *Quarterly Journal of Economics*, vol. 122, no. 3, pp. 1-37, 2007.

[51] **Finkelstein, Amy, and Robin McKnight.** "What Did Medicare Do? The Initial Impact of Medicare on Mortality and Out of Pocket Medical Spending," *Journal of Public Economics*, 92(7), 1644-1668, 2008.

[52] **Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and The Oregon Health Study Group.** "The Oregon Health Insurance Experiment: Evidence from the First Year," *Quarterly Journal of Economics*, forthcoming.

[53] **Feenberg, Daniel, and Jonathan Skinner.** "The Risk and Duration of Catastrophic Health Expenditures," *Review of Economics and Statistics*, 76(4): 663-647. 1994.

[54] **Frandsen, Brigham, Markus Frolich, and Blaise Melly.** "Quantile Treatment Effects in the Regression Discontinuity Design," *Mimeo*, 2010.

[55] **French, Eric, and John Bailey Jones.** "On the Distribution and Dynamics of Health Care Costs," *Journal of Applied Econometrics*, 19(6): 705-721, 2004.

[56] **Frolich, Markus, and Blaise Melly.** "Quantile Treatment Effects in the Regression Discontinuity Design," *Mimeo*, 2010.

[57] **Fuchs, Victor.** "The Supply of Surgeons and the Demand for Operations." *Journal of Human Resources*,13: 35-56, 1978.

[58] **Gans, Joshua S., and Andrew Leigh.** "Born on the First of July: An (Un)natural Experiment in Birth Timing," *Journal of Public Economics*, Volume 93, Issues 1-2, 246-263, 2009.

[59] **Garthwaite, Craig L.** "The Doctor Might See You Now: The Supply Side Effects of Public Health Insurance Expansions," *NBER Working Paper* No.17070, 2011.

[60] **Goodman, David, Fisher, Elliott, Little, George, Stukel, Therese, Hua Chang, Chiang, and Schoendorf, Kenneth.** "The Relation between the Availability of Neonatal Intensive Care and Neonatal Mortality." The *New England Journal of Medicine*, 346(20), 1538–1544, 2002.

[61] **Grant, Darren.** "Physician Financial Incentives and Cesarean Delivery: New Conclusions from the Healthcare Cost and Utilization Project", *Journal of Health Economics*, 28(1), 244-250, 2009.

[62] **Grossman, Michael.** "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, No. 2, 223-255, 1972.

[63] **Gruber, Jonathan, Nathaniel Hendren, and Robert Townsend.** "Demand and Reimbursement Effects of Healthcare Reform: Health Care Utilization and Infant Mortality in Thailand," *NBER Working Paper* No. 17739, 2012.

[64] **Gruber, Jonathan, and Maria Owings.** "Physician Financial Incentives and Cesarian Section Delivery," *RAND Journal of Economics*, 27(1): 99-123, 1996.

[65] **Grumbach, Kevin.** "Specialists, technology, and newborns – Too much of a good thing?" *The New England Journal of Medicine*, 346(20), 1574–1575, 2002.

[66] **Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw.** "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1): 201-209, 2001.

[67] **Hai Fang, and John A. Rizzo.** "Competition and Physician-enabled Demand: The Role of Managed Care," *Journal of Economic Behavior and Organization*, 72(1), 463-474, 2009

[68] **Hanratty, M. J.** "Canadian National Health Insurance and Infant Health," *American Economic Review,* 86: 276-284, 1996.

[69] **Hashimoto, Hideki, Naoki Ikegami, Kenji Shibuya, Nobuyuki Izumida, Haruko Noguchi, Hideo Yasunaga, Hiroaki Miyata, Jose M Acuin, and Michael R Reich.** "Cost Containment and Quality of Care In Japan: Is There a Trade-off?" *The Lancet*, 30, 2011.

[70] **Hoynes, Hilary W., and Diane Whitmore Schanzenbach**. “Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program,” *American Economic Journal: Applied Economics*, 1, 109–139, 2009.

[71] **Hughes, David and Songkramchai Leethongdee.** "Universal Coverage In The Land of Smiles: Lessons From Thailand’s 30 Baht Health Reforms," *Health Affairs*, 26, no.4:999-1008, 2007.

[72] **Ikegami Naoki.** “Japanese Health Care: Low Cost through Regulated Fees,” *Health Affairs*, 10(3):87-109, 1991.

[73] **Ikegami, Naoki**. "Japan: maintaining equity through regulated fees," *Journal of Health Politics, Policy and Law*, 17:689-713, 1992.

[74] **Ikegami Naoki, and Campbell JC.** “Medical Care in Japan,” *New England Journal of Medicine*, 333:1295-1299, 1995.

[75] **Ikegami, Naoki, Byung-Kwang Yoo, Hideki Hashimot, Masatoshi Matsumoto, Hiroya Ogata, Akira Babazono, Ryo Watanabe, Kenji Shibuya, Bong-Min Yang, Michael R Reich, and Yasuki Kobayashi.** “Japanese Universal Health Coverage: Evolution, Achievements, and Challenges,” *The Lancet*, 30, 2011.

[76] **Ikeda, Nayu, Eiko Saito, Naoki Kondo, Manami Inoue, Shunya Ikeda, Toshihiko Satoh, Koji Wada, Andrew Stickley, Kota Katanoda, Tetsuya Mizoue, Mitsuhiko Noda, Hiroyasu Iso, Yoshihisa Fujino, Tomotaka Sobue, Shoichiro Tsugane, Mohsen Naghavi, Majid Ezzati, and Kenji Shibuya.** “What Has Made the Population of Japan Healthy?” *The Lancet*, 30, 2011.

[77] **Itabashi, Kazuo, Takeshi Horiuchi, Satoshi Kusuda, Kazuhiko Kabe, Yasufumi Itani, Takashi Nakamura, Masanori Fujimura, and Masafumi Matsuo**. “Mortality Rates for Extremely Low Birth Weight Infants Born in Japan in 2005”, *Pediatrics*, 123:2 445-450, 2009.

[78] **Imbens, Guido W., and Thomas Lemieux.** “Regression Discontinuity Designs: A Guide to Practice,” *Journal of Econometrics*, 142(2): 615-35, 2008.

[79] **Jorgenson, Dale W., and Kun-Young Yun.** *Investment, Volume 3, Lifting the Burden: Tax Reform, the Cost of Capital, and U.S. Economic Growth.* Cambridge, MA: MIT Press, 2001.

[80] **Kawaguchi, Daiji.** "Actual Age at School Entry, Educational Outcomes, and Earnings," *Journal of the Japanese and International Economies*, Volume 25, Issue 2, 64-80, 2011.

[81] **Keeler, Emmett B., Joseph P. Newhouse, and Charles E. Phelps.** "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule Under Uncertainty," *Econometrica*, 45(3):641-656, 1977.

[82] **Keeler, Emmett, and John E. Rolph.** "The Demand for Episodes of Treatment in the Health Insurance Experiment," *Journal of Health Economics*, 7(4):337-367, 1988.

[83] **King, Gary, Emmanuela Gakidou, Kosuke Imai, Jason Lakin, Ryan T Moore, Clayton Nall, Nirmala Ravishankar, Manett Vargas, Martha María Téllez-Rojo, Juan Eugenio Hernández Ávila, Mauricio Hernandez Avila, and Hêtor Hernández Llamas.** "Public Policy for the Poor? A Randomised Assessment of the Mexican Universal Health Insurance Programme", *The Lancet*, Vol. 373, Issue 9673, pp. 1447-1454, 2009.

[84] **Kliegman, R.M.** "Neonatal Technology, Perinatal Survival, Social Consequences, and the Perinatal Paradox," *American Journal of Public Healt*h, 85, no. 7: 909-913, 1995.

[85] **Kolstad Jonathan T., and Amanda E. Kowalski.** "The Impact of Health Care Reform on Hospital and Preventive Care: Evidence from Massachusetts," *NBER Working Paper* No. 16012, 2010.

[86] **Kondo, Ayako, and Hitoshi Shigeoka.** "Effects of Universal Health Insurance on Health Care Utilization, Supply-Side Responses and Mortality Rates: Evidence from Japan," *Mimeo*, 2011.

[87] **Kopczuk, Wojciech.** "Tax Bases, Tax Rates and the Elasticity of Reported Income," *Journal of Public Economics*, 89(11-12), 2093-2119, 2005.

[88] **Kowalski, Amanda E.** "Estimating the Tradeoff between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance," *Mimeo*, 2011.

[89] **Kuwabara, Hiroyo, and Kiyohide Fushimi.** "The impact of a new payment system with case-mix measurement on hospital practices for breast cancer patients in Japan," *Health Policy*, 92(1), 65-72, 2009.

[90] **Lee, David S.** "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, 142(2) 675-697, 2008.

[91] **Lee, David S., and David Card.** "Regression Discontinuity Inference with Specification Error," *Journal of Econometrics*, 142(2): 655-674, 2008.

[92] **Lee, David S., and Thomas Lemieux.** "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, 48: 281-355, 2010.

[93] **Lee, David S., and Justin McCrary.** "The Deterrence Effect of Prison: Dynamic Theory and Evidence," *Industrial Relations Section Working Paper* No.550, 2009.

[94] **Lee K S, N Paneth, L M Gartner, M A Pearlman, and L Gruss**. "Neonatal Mortality: An Analysis of the Recent Improvement in the United States," *American Journal of Public Health*, 70, no. 1: 15-21, 1980.

[95] **Madrian, Brigitte C.** "Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock?" *Quarterly Journal of Economics*, Vol. 109, No. 1, pp. 27-54, 1994.

[96] **Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz.** "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review* 77(3), pp251-277, 1987.

[97] **Matsuda Shinya, Kohichi B Ishikawa, Kazuaki Kuwabara, Kenji Fujimori, Kiyohide Fushimi and Hideki Hashimoto**. "Development and use of the Japanese case-mix system", *Eurohealth,* Vol 14, No 3, pp25-30, 2008.

[98] **McClellan, Mark, and Jonathan Skinner.** "The Incidence of Medicare," *Journal of Public Economics*, 90 (1-2), 257-276, 2006.

[99] **McGuire, Thomas G.** "Physician agency." In: Culyer, A.J., Newhouse, J.P. (Eds.), *Handbook of Health Economics*. North Holland, 2000.

[100] **McGuire, Thomas G. and Mark V. Pauly**. "Physician Response to Fee Changes with Multiple Payers," *Journal of Health Economics*, 10, 385-410, 1991.

[101] **Miller, Grant, Diana M. Pinto, and Marcos Vera-Hernández.** "High-Powered Incentives in Developing Country Health Insurance: Evidence from Colombia's Régimen Subsidiado," *NBER Working Paper* No. 15456, 2009.

[102] **Ministry of Health, Labour and Welfare.** *Estimates of National Medical Care Expenditure*. www.e-stat.go.jp/SG1/estat/List.do?lid=000001058043 (accessed Aug 21, 2011), 2010.

[103] **Newhouse, Joseph, and the Insurance Experiment Group.** *Free for All? Lessons from the RAND Health Insurance Experiment*, Cambridge, MA: Harvard University Press, 1993.

[104] **Okamura, Shinichi, Ryota Kobayashi, and Tetsuo Sakamaki.** "Case-mix Payment in Japanese Medical Care," *Health Policy*, 74, 282-286, 2005.

[105] **Organisation for Economic Co-operation and Development.** *OECD Health Data*. www.oecd.org/document/16/0,2340,en_2649_34631_2085200_1_1_1_1,00.html (accessed August 20, 2011), 2011.

[106] **Pauly, M.V.** *Doctors and Their Workshops: Economic Model of Physician Behavior*, University Chicago Press, Chicago, 1980.

[107] **Phelps, C.E.** "Induced Demand: Can We Ever Know Its Extent?" *Journal of Health Economics* 5:355-365, 1986.

[108] **Phibbs, Ciaran S., Janet M. Bronstein, Eric Buxton, and Roderic H. Phibbs**. "The Effects of Patient Volume and Level of Care at the Hospital of Birth on Neonatal Mortality." *Journal of the American Medical Association*, 276(13): 1054-1059, 1996.

[109] **Phibbs, Ciaran S., Laurence C. Baker, Aaron B. Caughey, Beate Danielsen, Susan K. Schmitt, and Roderic H. Phibbs.** “Level and Volume of Neonatal Intensive Care and Mortality in Very-Low-Birth-Weight Infants." *New England Journal of Medicine*, 356(21): 2165-2175, 2007.

[110] **Pierdomenico SD, Nicola MD, Esposito AL, Mascio RD, Ballone E, Lapenna D, and Cuccurullo F.** “Prognostic Value of Different Indices of Blood Pressure Variability in Hypertensive Patients,” *American Journal of Hypertension*, 22: 842-847, 2009.

[111] **Porter, Jack.** “Estimation in the Regression Discontinuity Model,” *Mimeo*, 2003.

[112] **Poterba, James.** “Government Intervention in the Markets for Education and Health Care: How and Why?” in *Individual and Social Responsibility* Victor Fuchs (ed). University of Chicago Press, 1996.

[113] **Quinn, Kevin.** "New directions in Medicaid payment for hospital care.", *Health Affairs*, 27(1), 269–280, 2008.

[114] **Rice, T.** "The Impact of Changing Medical Care Reimbursement Rates on Physician-Induced Demand", *Medical care*, vol 21, n8, p803-815, 1983.

[115] **Saez, Emmanuel**. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy*, 2(3): 180–212, 2010.

[116] **Shigeoka, Hitoshi.** “The Effect of Patient Cost-Sharing on Utilization, Health and Risk Protection: Evidence from Japan,”,*Mimeo*, 2011.

[117] **Shigeoka, Hitoshi, and Kiyohide Fushimi.** “Supply-Induced Demand in Newborn Treatment: Evidence from Japan,” *Mimeo*, 2011.

[118] **Silverman, Elaine M. and Skinner, Jonathan S.** "Medicare Upcoding and Hospital Ownership," *Journal of Health Economics*, 23(2), pp. 369-89, 2004.

[119] **Swartz, Katherine.** “Cost-Sharing: Effects on Spending and Outcomes,” *Robert Wood Johnson Foundation Research Synthesis Report* No. 20, 2010.

[120] **Weissman, Joel S., Constantine Gatsonis, and Arnold M. Epstein.** "Rates of Avoidable Hospitalization by Insurance Status in Massachusetts and Maryland," *Journal of American Medical Association*, 268(17): 2388-2394, 1992.

[121] **Yip, W.** "Physician Responses to Medical Fee Reductions: Changes in the Volume and Intensity of Supply of Coronary Artery Bypass Graft (CABG) Surgeries in the Medicare and Private Sectors," *Journal of Health Economics*, 6 (17), 675-699, 1998.

[122] **Yoshihara, Kenji and Masaru Wada.** *History of health insurance system in Japan* (Nihon Iryo Hoken Seidoshi), Toyo Keizai Shinbum Sha, in Japanese, 1999.

[123] **Zeckhauser, Richard.** "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, Vol 2: 10-26, 1970.

[124] **Zweifel, Peter, Breyer, Friedrich H. J., and Kifmann, Mathias**. *Health Economics*. Second edn. Oxford University Press, 2009.

APPENDIX A

The Effect of Patient Cost-sharing on Utilization, Health and Risk Protection: Evidence from Japan

A.1. Derivation of Out-of-Pocket Health Expenditures

This section in the appendix describes how I convert the cost-sharing formula in Table 1.2 into the actual monthly out-of-pocket health expenditures in Table 1.3. It is ideal if we have information on actual out-of-pocket expenditures at the individual level, such as Medical Expenditure Panel Survey (MEPS) in the US. In the absence of such data, I derive this myself.

Fortunately, I know the exact formula for cost-sharing (Table 1.2) and have individual level insurance claim data, which is the monthly summary of medical expenditures claimed for insurance reimbursement to medical institutions (called the Survey of Medical Care Activities in Public Health Insurance). Since a portion of this monthly total medical expenditure is paid as patient cost-sharing, using the formula in Table 1.2, I can compute the average out-of-pocket medical expenditures at each age for each survey year of the Patient Survey.[1]

[1]The rest of medical expenditures are paid by insurance societies. The source of the money is a fund of the pooled premiums of insured members and assistance from the government.

The insurance claim data is monthly since reimbursements to the medical institutions are conventionally paid monthly in Japan. Thus the stop-loss is set by monthly rather than annually unlike the US. The age of patients is measured in years in this data.

The steps I compute the average monthly out-of-pocket expenditures are as follows. Note that cost-sharing formula differs by outpatient visits and inpatient admissions; since inpatient admissions are more expensive and put more financial burden on patients, the coinsurance rate of inpatient admissions tend to be set lower than those of outpatient visits.

Those below age 70

First, I compute the average monthly out-of-pocket health expenditures for 69-year-old patients. For those below age 70, the coinsurance rate is determined by the type of health insurance: NHI, employees in employment-based health insurance, and dependent of employees in employment-based health insurance. Among those in NHI, the coinsurance rate differs among those who are still employed, retired former employees, and dependents of retied employees. I use information from the CSLC to compute the rate of those employed among NHI recipients. Also, assuming that males who are not employed are retired former employees and females who are not employed are dependents of retied employees, I compute the weighted average of the coinsurance rate for NHI. This assumption does not make any major differences for this computation, since the fraction of retired former employee is quite small. In fact, the coinsurance rate for only outpatient visits during

1984-2002 differs by 10 percent between retired former employees and dependents of retied employees, and the computed weighted coinsurance rate for NHI is around 28 percent, which is very close to the coinsurance rate for the employed and dependents of retired employees among NHI (30 percent). For inpatient admissions, this assumption plays no role, since the coinsurance rate for inpatient admissions is the same (20 percent) for retired former employees and dependents of retired employees.

Then, actual out-of-pocket medical expenditures, AM_{ipt}, for individual i whose health insurance plan p (p=1-3, where 1: NHI, 2: employees in employment-based health insurance, and 3: dependent of employees in employment-based health insurance), and types of services use j (j=1-2, where 1: inpatient admissions, 2: outpatient visits) in survey year t, is given as follows:

$$AM_{ipt} = min(EM_{ijpt}, SL_{jpt})$$

where EM_{ijpt} is the expected payment without stop loss (or maximum amount of out-of-pocket expenditures), and SL_{jpt} is stop-loss for each plan p for each service use j in survey year t.

Suppose there is an individual whose total medical expenditures for inpatient use in June 2008 is 1,000,000 Yen, and the coinsurance rate is 30 percent. This indicates that EM_{ijpt} of 300,000 Yen. On the other hand, SL_{jpt} is 87,430, which is 80,100+(1,000,000-267,000)*0.01, according to the formula in Table 1.2. Since

SL is smaller than EM, AM is 87,430 Yen. I compute AM for each individual level claim data, and take the simple average to compute the average expenditure AM_{jpt}, by each plan type p, for each service j in survey year t.

Finally, I take a weighted average of each insurance type W_{pt}, obtained from the CSLC. Therefore, the average monthly out-of-pocket medical expenditure AM for age 69 is:

$$AM_{jt}(age69) = \sum_{p=1}^{3}(W_{pt} * AM_{jpt})$$

for use of type j in each survey year t of Patient Survey. I take W_{pt} for each year t, from the CSLC in year $t-1$ since CSLC is conducted a year before the Patient Survey. The exception is the Patient Survey year of 1984, when the fraction from 1987 of the CSLC is used as a weight since it is the closest year of information available. The majority of 69 year-olds (roughly 70-80 percent) belongs to NHI, and the rest belongs to employment-based health insurance.

Those above age 70

Next, I compute the average out-of-pocket health expenditures for 70-year-old patients, who all receive Elderly Health Insurance. Since utilization is endogenous (i.e. observed out-of-pocket medical expenditure already reflects the change in cost-sharing), I compute a counterfactual out-of-pocket expenditure for 70-year-old patient if they had the same amount of utilization as the average 69-year-old. I compute the average monthly frequency of visits for outpatient visits, and average length of stay for inpatient admissions for age 69, and applied the formula for age 70 to compute the monthly average out-of-pocket medical expenditures, in the same manner as those for age 69 described above.

Finally, the overall out-of-pocket medical expenditure in Table 1.3 is the weighted average of the out-of-pocket medical expenditure across all survey years for out-patient visits and inpatient admissions respectively, using the population of age 69 in each survey year as weights. For reference, Appendix Table A.8 shows the estimated out-of-pocket medical expenditure for each survey year.

It is worth mentioning that these figures I compute is a rough estimates of actual out-of-pocket medical expenditures since the actual cost-sharing is a little bit more complicated than this simple exercise. For example, different coinsurance rates are applied to specific populations, and there is another way to reduce out-of-pocket medical expenditures. For example, in October 2002, the coinsurance rate for those over age 70 with high income – 7 percent according to Ikegami et al. 2011 - was raised from 10 percent to 20 percent. Also for all ages, the stop-loss is set lower for very low-income people. There is a stop-loss at the household level, instead of individual level, where family members are allowed to aggregate their medical spending. Nonetheless, since most of the patients are under the basic cost-sharing formula, the cost-sharing I estimate should be within an acceptable range.

Table: Summary of the Datasets Used in this Study

	Name of Dataset	Period	Interval
1	Patient Survey	1984-2008	Every three year (9 rounds in total)
2	Survey of Medical Institutions	1984-2008	Every three year (9 rounds in total)
3	Comprehensive Survey of Living Conditions	1986-2007	Every three year (8 rounds in total)
4	Survey of Medical Care Activities in Public Health Insurance	1984-2008	Every year
5	Vital Statistics: Mortality data	1987-1991	Every year

A.2. Data Apendix

In this study, I use a variety of datasets collected mainly by the Ministry of Health, Labour and Welfare. A brief description of each dataset is provided in this data appendix. The English-Japanese crosswalks of the name of the datasets can be found at the following website from Ministry of Health, Labour and Welfare. http://www.mhlw.go.jp/toukei/itiran/eiyaku.html

A.2.1. Patient Survey

Detail: http://www.mhlw.go.jp/english/database/db-hss/dl/sps_2008_06.pdf

The Patient Survey is a national sample survey of hospitals and clinics that has gathered information on the utilization of medical institutions in Japan since 1948. The comprehensive version of the current Patient Survey is conduced every three years since 1984. It covers roughly 2000-7000 hospitals and 3000-6000 clinics per survey year. It collects information on ICD code, patients' principal sources of payment, and the limited socio-demographic characteristics such as gender and patients' place of living. The individual patient level microdata files are available starting from 1984.

There are two datasets in the Patient Survey, outpatient data, which I use to examine outpatient visits, and discharge data, which I use to examine inpatients admissions.

A.2.1.1. Outpatient data. The outpatient data in the Patient Survey is conducted one day in middle of the October (normally a weekday in the second week), and collects information on all patients that visit hospitals or clinics for outpatients reasons (i.e., visits to hospitals for non-hospitalization reasons). The datasets contain 75,000-100,000 individuals for outpatient visits. This data includes exact date of birth and the survey date, which is equivalent to the exact date of visits and enables me to compute age in days at the time of outpatient visits. The sample size of the outpatient data is about 500,000-1,500,000.

A.2.1.2. Discharge data. The discharge data in the Patient Survey reports all the inpatients record discharged in the surveyed hospitals and clinics within September in the survey year. The datasets contain about 180,000-970,000 inpatients records per each survey year. The sample size gets larger in more recent years. The data includes the exact day of birth, admission, discharge, and surgery. It also contains information whether the patient needed surgery, and several types of main surgery (collected from 1999 on). Unlike the Comprehensive Survey of Living Conditions, the discharge data include patients who die in the hospital as well as clinics.

A.2.2. Survey of Medical Institutions

Detail: http://www.mhlw.go.jp/english/database/db-hss/dl/01_Outline_of_Survey.pdf

The Survey of Medical Institutions collects information on all medical institutions in Japan that are in practice at the time of survey. The survey was conducted every year until 1972 and every three years since then. The individual hospital/clinic level microdata files are available starting from 1972. The data collect information on the ownership of institutions, number of beds permitted, notification of emergency, teaching school status, number of physicians, clinical specialties, machinery and equipment, and their working conditions. I merge this hospital and clinic information to the Patient Survey based on institution ID.

A.2.3. Comprehensive Survey of Living Conditions (CSLC)

Detail: http://www.mhlw.go.jp/english/database/db-hss/cslc.html

The Comprehensive Survey of Living Conditions (CSLC) is a nationwide repeated cross-section survey of households that has gathered information on the health of the Japanese people since 1986. The CSLC collects information on socio-demographic characteristics, and health related topics. The long version of CSLC used in this study is conducted every three years for randomly sampled individuals based on the 3000-5000 districts from the National Census conducted every five years ending with last digit of zero or five.

The microdata files are available starting from fiscal year 1986. The survey reports births in months, so I use this information to compute the age in month combined with the information on month of the survey. The long version of CSLC consists of three questionnaires: Household, Health, and Income and Savings. A long-term care questionnaire was added in 2004. I mainly use the data on the health questionnaire that collects information on self-reported physical and mental health, and activity limitations.

I also use the insurance type information in the household questionnaire, to compute the average health insurance coverage of each health insurance type, which is mapped to the Survey of Medical Care Activities in Public Health Insurance to derive the amount of out-of-pocket medical expenditures. The household forms also include the basic individual-level socio-demographics such as gender, marital status, employment, and household size. The income and saving questionnaire asks the amount and source of income, and amount of saving and debt. Information on out-of-pocket medical expenditures at individual level is only collected in 2007. I use individual income and out-of-pocket medical expenditures to compute the welfare gains from risk reduction.

The survey covered 240,000-290,000 households and 740,000-800,000 household members in each survey round. The income and savings questionnaire is conducted for only around 15 percent of the whole sample.

A.2.4. Survey of Medical Care Activities in Public Health Insurance

Detail: http://www.mhlw.go.jp/english/database/db-hss/dl/shw-03.pdf

The Survey of Medical Care Activities in Public Health Insurance is a survey of health insurance claims data that gathers yearly information on detailed statements of medical fees and pharmacy dispensing fee. I use this information to derive the average monthly out-of-pocket medical expenditures for those who use medical institutions as described in Appendix A1.

Due to the monthly reimbursement to the medical institutions, the claim data is a summary of the medical expenditures per month per individual who uses medical institutions in June of the survey year. The data is collected from the prefectural branches of the Social Insurance Medical Fee Payment Fund for employment-based health insurance recipients and the Federation of National Health Insurance for National Health Insurance recipients. Health insurance claim data from the society-managed employment-based health insurance recipients is collected since 1999. Age is measured in year.

A.2.5. Vital Statistics: Mortality data

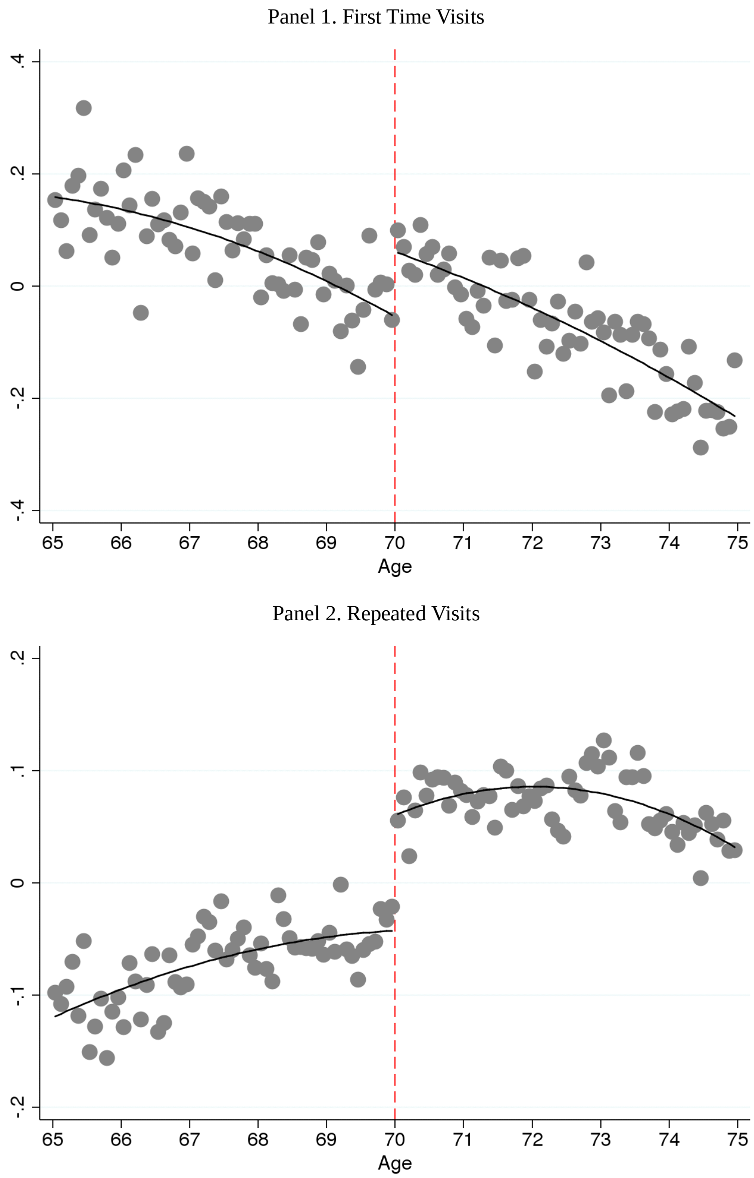
Detail: http://www.mhlw.go.jp/english/database/db-hw/outline/index.html

The 1984-2008 National Mortality Details Files is an annual census of deaths within Japan. The data contain the universe of deaths and information on the deceased's date of birth, and date of death, which enables me to compute age in

Cause of Death	1984 -1994 (ICD-9)	1995-2008 (ICD-10)
Main Cause		
Cancer	140-208	C00-C97
Heart Disease	390-398, 402, 404 410-429	I00-I09, I11, I13, I20-I51
Cerebrovascular Disease	430-434, 436-438	I60-I69
Respiratory Disease	460-519	J00-J99
Sub diagnosis		
Hypertensive Disease	401-405	I10-I15
Ischemic Heart Disease	410-414	I20- I25
Intracerebral Hemorrhage	431-432	I61, I69.1
Cerebral Infarction	433, 434, 437.7a, 433.7b	I63, I69.3

days at the time of death. The data also include gender, nationality, place of the death, and cause of deaths according to the International Classification of Disease (ICD). ICD9 was used till 1994, and ICD10 is used since 1995 in Japan. The ICD codes for each cause of death used in this paper are following;

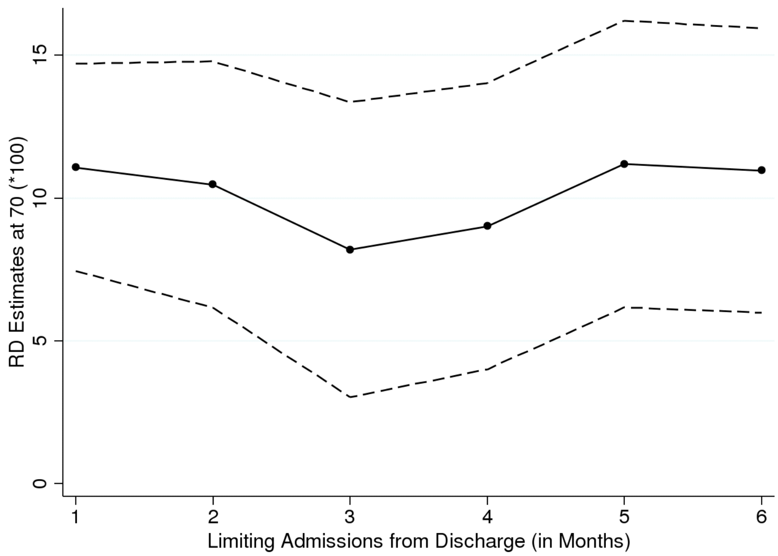
Figure A.1: Age Profiles for First Time and Repeated Outpatient Visits



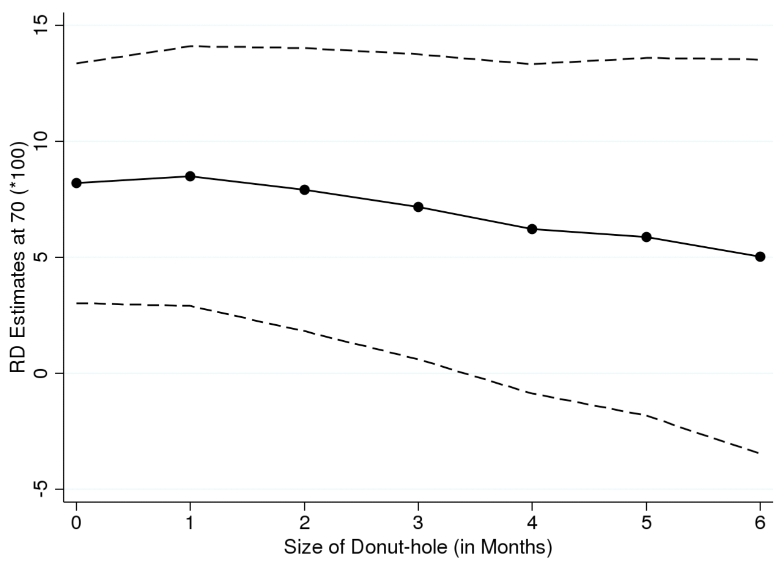
Note: The data come from pooled 1984-2007 outpatient data in the Patient Survey. The markers represent actual averages of residual of outcome that is regressed by birth month fixed effects and the survey year fixed effect to partial out the seasonality in birth and the underlying common shocks in the survey year. The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure A.2: Robustness of Results on Inpatient Admissions

Panel 1. Limiting the Sample by Different Windows from Discharge

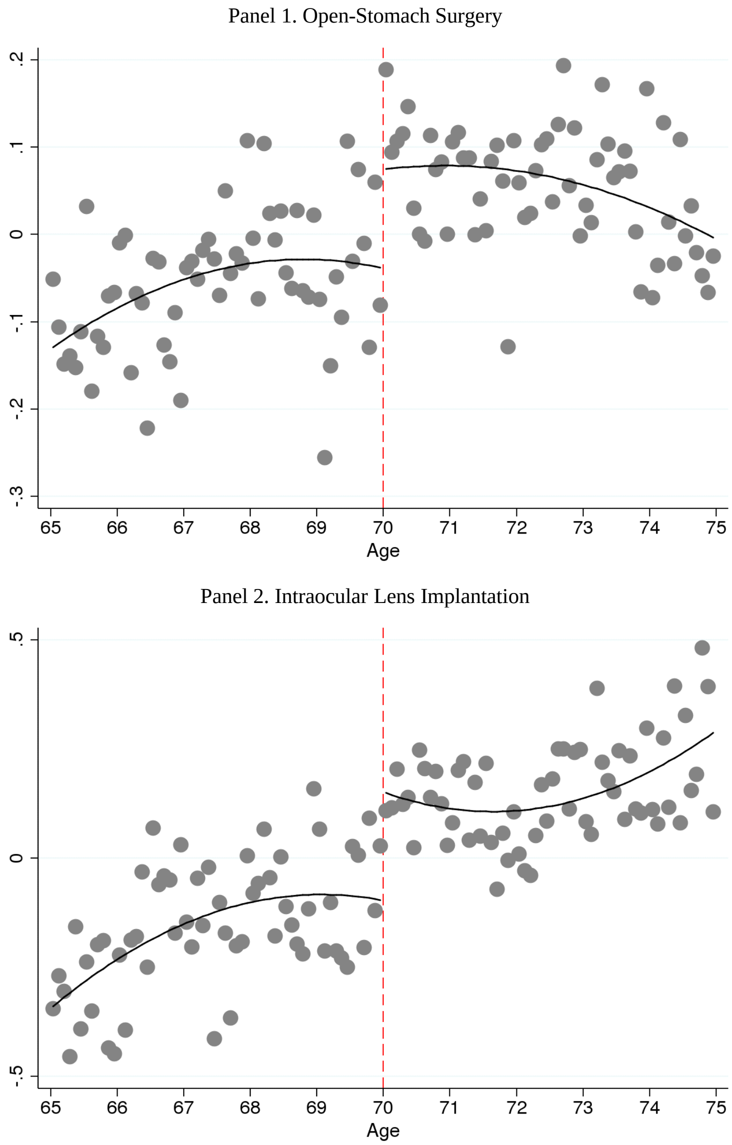


Panel 2. Estimates from "Donut-hole" RD



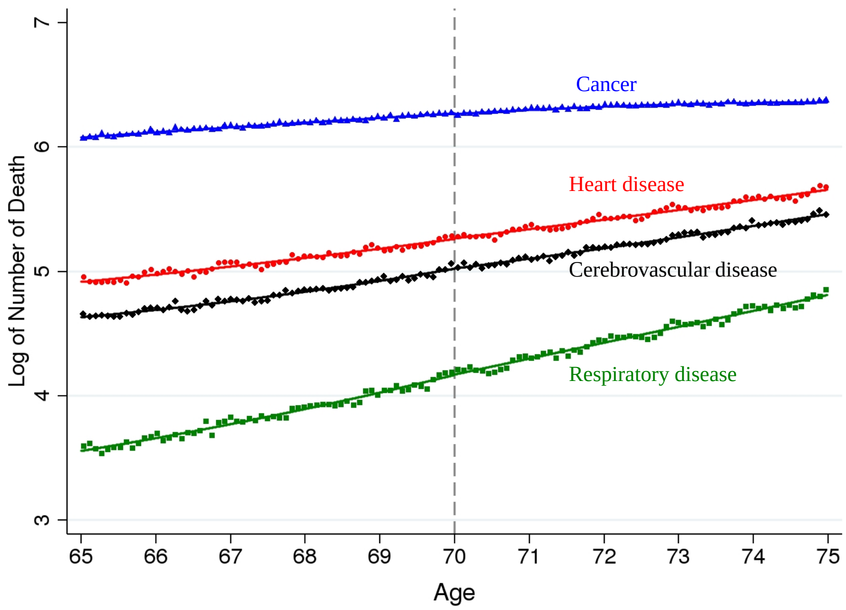
Note: The data come from pooled 1984-2008 discharge data in Patient Survey. The model here is quadratic age profile fully interacted with a dummy for age 70 or older. Dashed line is 95 percent confidence interval.

Figure A.3: Age Profile for Inpatient Admissions for Selected Surgery (log scale)



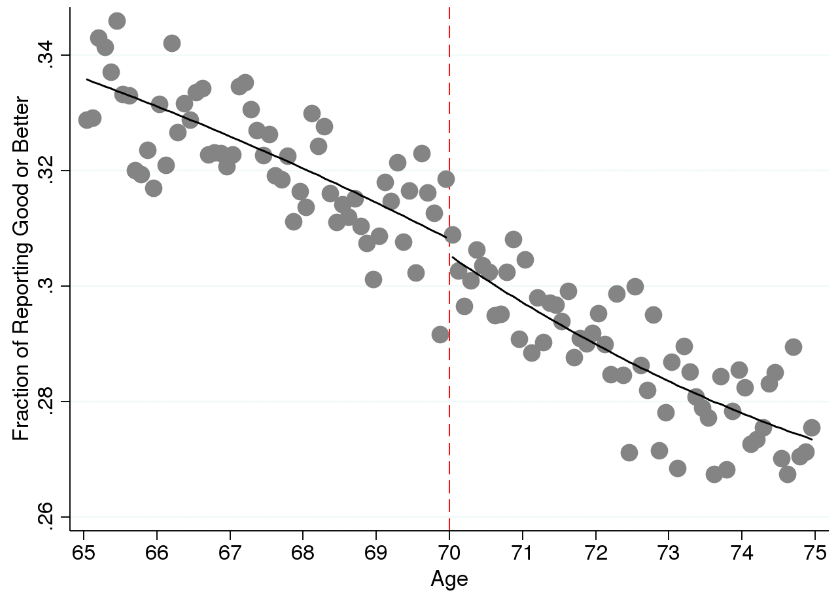
Note: The data come from pooled (1999, 2002, 2005, and 2008) discharge data in Patient Survey since specific surgery information is collected for only these four survey years. I use admissions within three months from discharge, and thus the sample size is 1,440. The markers represent actual averages of residual of log outcome that is regressed by birth month fixed effects, admission month fixed effects, and the survey year fixed effect to partial out the seasonality in birth and the underlying common shocks in the survey year. The lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure A.4: Age Profile for Cause-Specific Mortality



Note: The data come from pooled 1984-2008 mortality data. I use days to eligibility for the Elderly Health Insurance as a running variable. The cell is each 30 days interval from the day of eligibility at age 70. The markers represent the averages, and the lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Figure A.5: Age Profiles for Fraction in Good or Very Good Health



Note: The data come from pooled 1986-2007 Comprehensive Survey of Living Conditions. The markers represent actual averages (age in month), and the lines represent fitted regressions from models that assume a quadratic age profile fully interacted with a dummy for age 70 or older.

Table A.1: Top 10 Diagnosis for Outpatient Visits, and Inpatient Admission

Panel 1. Outpatient Visits

rank	Name of diagnosis	Percentage	ICD9 (3digit)
1	Essential hypertension	16.1%	401
2	Spondylosis and allied disorders	4.7%	721
3	Diabetes mellitus	4.7%	250
4	Osteoarthrosis and allied disorders	4.3%	715
5	Cataract	3.4%	366
6	Other and unspecified disorders of back	3.3%	724
7	Gastritis and duodenitis	2.3%	535
8	Occlusion of cerebral arteries	2.1%	434
9	Other disorders of bone and cartilage	1.9%	733
10	Disorders of lipoid metabolism	1.8%	272

Note: The data come from the pooled 1984-2008 outpatient visits data in the Patient Survey.

Panel 2. Inpatient Admissions

rank	Name of diagnosis	Percentage	ICD9 (3digit)
1	Cataract	4.4%	366
2	Angina pectoris	4.1%	413
3	Occlusion of cerebral arteries	3.8%	434
4	Diabetes mellitus	3.2%	250
5	Malignant neoplasm of stomach	3.1%	151
6	Benign neoplasm of other parts of digestive system	2.9%	211
7	Malignant neoplasm of liver and intrahepatic bile ducts	2.3%	155
8	Malignant neoplasm of colon	2.1%	153
9	Malignant neoplasm of trachea, bronchus and lung	1.8%	162
10	Cholelithiasis	1.5%	574

Note: The data comes from the pooled 1984-2008 discharge data in the Patient Survey.

Table A.2: Robustness of RD Estimates on Outpatient Visits for Selected Outcomes

Running Variable: Age in	Month			Day		
	Basic	Age 67-73	Cubic	Basic	Age 67-73	Cubic
	(1)	(2)	(3)	(4)	(5)	(6)
A. All	10.3***	11.3***	12.1***	11.4***	12.3***	12.7***
	(1.8)	(2.3)	(2.6)	(1.6)	(2.1)	(2.2)
B. By Visit Type						
Repeated visits	10.3***	11.2***	12.1***	11.4***	12.1***	12.5***
	(1.9)	(2.3)	(2.6)	(1.6)	(2.1)	(2.2)
C. Days from Last Outpatients Visits Among Repeated Visits						
1 day	16.4***	20.9***	21.6***	15.7***	17.1***	16.5***
	(4.4)	(6.1)	(6.5)	(2.1)	(2.7)	(2.9)
4-7 day	8.5***	6.6	8.7*	9.6***	11.7***	10.5***
	(3.0)	(4.1)	(4.6)	(2.3)	(3.1)	(3.2)
D. By Institution						
Clinic	13.8***	15.1***	16.0***	13.4***	14.2***	14.7***
	(1.8)	(2.3)	(2.6)	(1.1)	(1.5)	(1.5)
E. By Referral						
Without Referral	10.5***	11.6***	12.5***	11.5***	12.3***	12.8***
	(1.9)	(2.3)	(2.6)	(1.6)	(2.1)	(2.2)

Note: Each cell is the estimate from separate estimated regression discontinuities at age 70. “Basic” is the model that include quadratic of age, fully interacted with dummy for age 70 or older among people between ages 65-75. Controls are dummies for each survey year and each month of birth. I use pooled samples of the Patient Survey conducted every three year since 1984. Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table A.3: List of PQI (Ambulatory-Care-Sensitive Conditions)

Number	Name of Diagnosis
PQI 1	Diabetes, short-term complications
PQI 3	Diabetes, long-term complications
PQI 5	Chronic obstructive pulmonary disease
PQI 7	Hypertension
PQI 8	Congestive heart failure
PQI 10	Dehydration
PQI 11	Bacterial pneumonia
PQI 12	Urinary infections
PQI 13	Angina without procedure
PQI 14	Uncontrolled diabetes
PQI 15	Adult asthma
PQI 16	Lower extremity amputations among patients with diabetes

Note: I excluded PQ2 (Perforated appendicitis) from the analysis since this index is the number of admissions for perforated appendix as a share of admissions for appendicitis only. Also PQI 14 requires the fifth digit of the ICD9, which I don't have, since PQI 14 only include 25002 and 25003 (25000, 25001, and 25009 should not be included). To account for this, I only include diabetes (2500) which has secondary diagnosis.

Table A.4: Robustness of RD Estimates on Inpatient Admissions for Selected Outcomes

		Basic	Age 67-73	Cubic
		(1)	(2)	(3)
A	All	8.2***	10.0***	11.2***
		(2.6)	(3.4)	(3.6)
B	Surgery			
	With surgery	10.8***	17.4***	20.7***
		(3.8)	(5.0)	(5.2)
C	Type of Surgery			
	Open-stomach surgery	11.4**	17.4**	19.5***
		(5.6)	(7.0)	(7.4)
	Intraocular lens implantation	19.6***	18.9**	19.1*
		(6.2)	(8.0)	(9.8)
E	By Diagnosis			
	Cataract	22.6***	31.6***	46.4***
		(6.5)	(8.5)	(9.7)
	Occlusion of cerebral arteries	13.7***	16.3***	18.2***
		(4.6)	(5.9)	(6.3)
	Ischemic heart disease	14.5**	17.3*	16.4*
		(7.1)	(9.3)	(9.7)
	Cerebral infarction	12.8***	14.4**	14.5**
		(4.6)	(6.0)	(6.3)

Note: "Basic" is the model that include quadratic of age, fully interacted with dummy for age 70 or older among people between ages 65-75. Controls are dummies for each survey year, each month of birth, and each month of admission. I use pooled samples of Patient Survey conducted every three year since 1984. Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes..

Table A.5: RD Estimates of Inpatient Admissions by Characteristics of Hospital

		Basic	Age 67-73	Cubic
		(1)	(2)	(3)
A	Ownership			
	Governmental hospitals	7.0**	9.5**	11.9***
		(3.2)	(4.2)	(4.4)
	Public hospitals	10.1**	13.8***	17.1***
		(4.0)	(5.2)	(5.4)
	Not-for-profit hospitals	8.5***	9.7***	10.0***
		(2.8)	(3.6)	(3.8)
B	Teaching			
	Teaching hospital	6.3	5.9	10.1
		(5.0)	(6.4)	(6.5)
	Non Teaching hospital	8.4***	10.2***	11.3***
		(2.6)	(3.4)	(3.6)
C	Emergency Department			
	With	8.3***	10.3***	12.3***
		(2.8)	(3.7)	(3.8)
	Without	7.7***	9.6***	9.6**
		(2.8)	(3.6)	(3.8)
D	Size of hospital			
	1-99 beds	12.5***	14.3***	14.8***
		(3.4)	(4.3)	(4.5)
	100-299 beds	4.9	4.7	4.6
		(3.1)	(4.1)	(4.3)
	300-3000 beds	9.9***	12.7***	15.5***
		(3.3)	(4.3)	(4.6)

Note: Each cell is the estimate from separate estimated regression discontinuities at age 70. The specification is a quadratic of age, fully interacted with dummy for age 70 or older among people between ages 65-75. Controls are dummies for each survey year, each month of birth, and each month of admission. I use pooled samples of Patient Survey conducted every three year since 1984. Sample size is 3,240. Robust standard errors are in parenthesis. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. All coefficients on Post70 and their standard errors have been multiplied by 100, so they can be interpreted as percentage changes.

Table A.6: RD Estimates at Age 70 on Morbidity

	Self-reported Health				Stress-related			
	Good or Better Health		Linear Regression (1=poor 5=excellent)		Stress Dummy		Stressed due to own health and care	
	Age 68-9	RD at 70	Age 68-9	RD at 70	Age 68-9	RD at 70	Age 68-9	RD at 70
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. All	31.4	-0.3	2.8	1.1	41.1	0.4	25.3	0.2
		(0.6)		(1.3)		0.4		(0.7)
B By HH Income								
Above median	32.1	-0.1	2.7	2.3	39.2	-0.7	22.9	1.0
		(1.9)		(4.3)		(2.4)		(2.0)
Below median	30.1	1.4	2.8	-5.1	44.8	-3.2	29.2	-0.5
		(2.0)		(4.7)		(2.5)		(2.3)
years available	1986-2007				1995-2001			

Note: Entries in odd-numbered columns are the mean of age 68-69 years-olds of the outcome variables shown in column heading. Entries in even-numbered columns are estimated regression discontinuities at age 70, from models that include quadratic control for age, fully interacted with dummy for age 70 or older among people between age 65 to age 70. Other controls include indicators for gender, region, marital status, birth month, and survey year. Except column 4, estimates are based on linear probability model fit to pooled samples of CSLS conducted every three year since 1986. Standard errors (in parenthesis) are clustered at the age in month level as this is the most refined version of the age variable available. All regressions are weighted to take into account the stratified sampling frame in the data. ***, **, * denote significance at the 1%, 5% and 10% levels respectively. Available years for each outcome are described in the last row. Income is collected for roughly 15 % of all samples, and thus the sample size of Panel B is smaller than the full sample. All coefficients in even-numbered columns on Post70 and its standard error have been multiplied by 100 in order to interpret them as percentage changes.

Table A.7: Estimated Out-of-Pocket Medical Expenditure per Month across Survey Years

Panel A. Outpatient Visits

	Cost-Sharing			% reached stop-loss	
	Below 70	Above70	% reduction	Below 70	Above70
year	(1)	(2)	((1)-(2))/(3)	(4)	(5)
All	**3.99**	**1.02**	**74%**	**0.1%**	**0.6%**
1987	3.96	0.80	80%	0.1%	-
1990	4.26	0.80	81%	0.1%	-
1993	4.48	1.00	78%	0.1%	-
1996	4.23	1.02	76%	0.1%	-
1999	3.91	1.00	74%	0.2%	-
2002	3.61	1.30	64%	0.1%	0.5%
2005	3.97	1.28	68%	0.2%	0.7%
2008	3.69	1.20	68%	0.1%	0.5%

Panel B. Inpatient Admissions

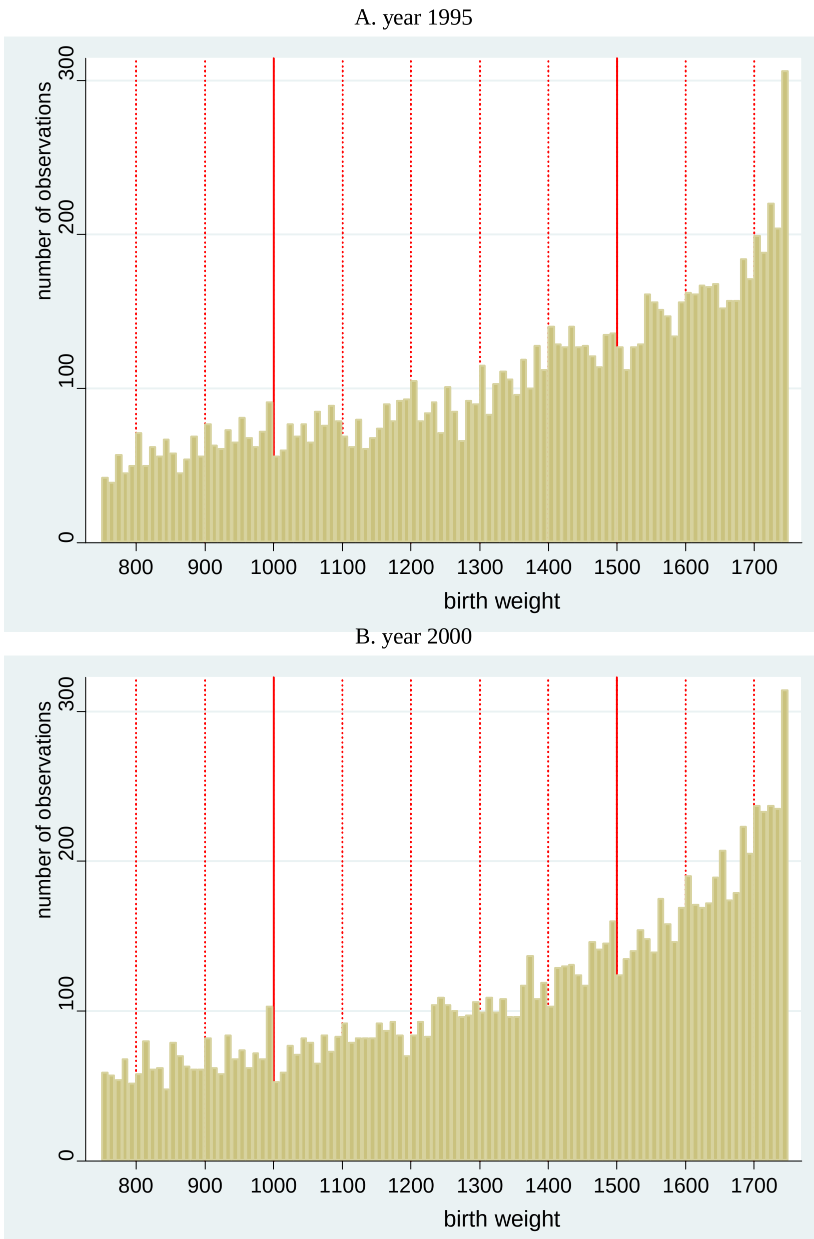
	Cost-Sharing			% reached stop-loss	
	Below 70	Above70	% reduction	Below 70	Above70
year	(1)	(2)	((1)-(2))/(3)	(4)	(5)
All	**37.95**	**12.44**	**67%**	**14.6%**	**0.0%**
1987	44.52	7.86	82%	26.6%	0.0%
1990	42.21	7.42	82%	21.6%	0.0%
1993	40.78	11.91	71%	11.5%	0.0%
1996	39.70	10.65	73%	11.5%	0.0%
1999	38.65	15.09	61%	9.2%	0.0%
2002	35.86	15.54	57%	8.7%	0.0%
2005	46.39	15.73	66%	18.3%	0.0%
2008	45.64	15.63	66%	13.5%	0.0%

Note: All money values without percentage sign are in thousand Yen (roughly 10 US dollar).

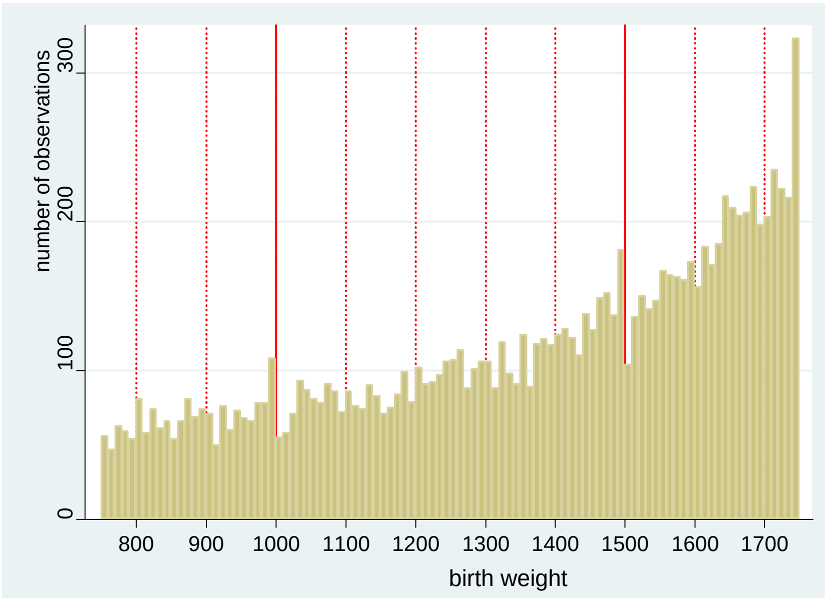
APPENDIX B

Supply Induced Demand in Newborn Treatment : Evidence from Japan

Figure B.1: The distribution of universe of birth in 1995, 2000 and 2005 (750-1750 grams)



C. year 2005



Note: The two solid vertical lines correspond to 1000 and 1500 grams, where the maximum number of the days those hospitals can claim reimbursement for NICU utilization differs. Other dotted line corresponds to every round value of 100 grams. The bin size is 10 grams.

Table B.1: Log difference in density for Figure B.1

	Cut-off	
year	1000 gram	1500 gram
1995	-0.33***	-0.095
	(0.119)	(0.088)
2000	-0.40***	-0.20***
	(0.118)	(0.083)
2005	-0.50***	-0.37***
	(0.115)	(0.084)

Note: To be consistent with Figure 4, I use the pilot bandwidth of 100 gram with the binsize of 10 gram. * p<0.10, ** p<0.05, *** p<0.01

Table B.2: Mother's delivery method

	All mothers	< 34 weeks of gestational length
Normal delivery (%)	0.077*	0.026
	(0.045)	(0.053)
All C-section (%)	-0.050	-0.058
	(0.048)	(0.061)
C-section: Emergency (%)	-0.049*	-0.074
	(0.029)	(0.066)
C-section: Elective (%)	-0.011	0.001
	(0.030)	(0.035)
Vacuum use (%)	-0.016	0.021*
	(0.013)	(0.012)
Forceps use (%)	-0.002	0.025*
	(0.009)	(0.013)
Sample size	53,094	5,080

Note: Each row corresponds to a separate OLS regression. The estimate on post is reported. Post is a dummy that equals one if hospital is under the new payment system and zero otherwise. All specifications include the year fixed effects and hospital fixed effects. Controls are age, age-squared, and multiple birth dummy. In addition to fixed effects and controls, we include 2002 hospital characteristics (number of beds, ownership of the hospital, a dummy for teaching hospital, level of hospital care (primary, secondary and tertiary), a dummy that takes the value of one if hospitals have an ER section, and a dummy that takes the value of one if hospitals have mandatory hospital within the same Health Service Area) each interacted with a linear time trend. Standard errors (in parentheses) are clustered at the hospital level. Significance level * p<0.10, ** p<0.05, *** p<0.01.

APPENDIX C

Effects of Universal Health Insurance on Health Care Utilization, Supply-Side Responses, and Mortality Rates: Evidence from Japan

C.1. Evidence against the Crowding-out of Employment-based Health Insurance by the NHI

As explained in Section 3.2, there are two potential channels through which the NHI expansion "crowded out" employment-based health insurance. First, the NHI could increase the number of self-employed workers by making ineligibility for employment-based health insurance a moot issue. Second, the introduction of the NHI could induce firms to reduce their size to fewer than five employees and receive an exemption from making financial contributions to employment-based health insurance.

To assess the first possibility, we calculate the ratio of self-employed workers to all individuals in the employed labor force, using data from the Population Censuses of 1950, 1955, and 1960. This self-employment ratio is the sum of the numbers of business owners without paid employees and family workers, divided by the number of all employed people 15 years old or over (14 for 1950). We exclude owners with paid employees, because they might be eligible for employment-based health insurance. Then, we regress the changes in this ratio from 1955 to 1960 on $impact_p$, the ratio of uninsured individuals in 1956. As shown in Table C.2, the ratio of uninsured individuals has no effect on the ratio of self-employed workers. Thus, we conclude that the first kind of crowding-out did not occur in the case of Japan in the 1950s.

Regarding the second possibility, we obtain data regarding the number of establishments, by size, from the Establishment Census. This survey has been conducted every three years; we use data from 1951, 1954, 1957, 1960, 1963, and 1966, and estimate equation (3.3) except that the base year (i.e., year with λ=0) is 1957. The estimated λ is shown in Table C.3.

If NHI expansion induced some firms to reduce their number of employees and thus receive an exemption from contributing to employment-based health insurance, the number of establishments with one to four employees should have increased during the 1956–60 period; the number of establishments with five to nine employees should also have decreased during the same period. Columns (1) and (2) of Table C.3A show that the number of establishments with one to four employees did not increase in response to NHI expansion, although the number of establishments with five to nine employees did decrease slightly. Columns (4) and (5) further show that, when looking at ratios rather than absolute numbers, establishments with one to four employees increased in the mid-1960s rather than in the late 1950s. Furthermore, these two estimates, λ_{63} and λ_{66}, seem to be driven solely by Tokyo and Osaka. As shown in Table C.3B, when we exclude Tokyo and Osaka, no λ remain statistically significant. Thus, column (4) of Table C.3A probably reflects the fact that Tokyo experienced a fall in the ratio of small establishments in the 1950s and had already reached by 1960 a much lower ratio than other prefectures, rather than a lagged response to the NHI expansion.

C.2. Impact on Household Out-of-Pocket Health Care Expenditures

Even with no improvement in health outcomes, health insurance may benefit insured individuals by reducing the risk of sudden out-of-pocket spending and thus by helping to smooth consumption (Finkelstein and McKnight 2008). To investigate whether, and to what extent, health insurance can reduce this risk, we need data regarding the distribution of out-of-pocket spending at the individual level. However, such data are not available. Thus, in this section, we instead explore the effect on average out-of-pocket medical expenditures.

Data pertaining to household medical out-of-pocket expenditures are taken from the National Survey of Family Income and Expenditures, which has been conducted every five years since 1959. This survey is nationally representative, in that both insured and noninsured individuals are included. Each surveyed household is asked to keep track of its household budget. Therefore, data on medical expenditures consist only of out-of-pocket medical expenditures by the household, and not payments made directly from the insurance system to medical providers. In addition, medical expenditures may include the purchase of nonprescription medication at drugstores. Medical spending by household in 1959, two years before the achievement of universal health insurance, was 2,206 yen (in 1980 prices) per month, or 1.8 percent of the total household income.

We examine the difference between 1959 and 1964 to estimate the impact of health insurance on out-of-pocket expenditures, as well as the difference between 1959 and 1969, to determine long-term effects. Specifically, we estimate the following first-difference regression:

$$dY = \beta_0 + \beta_1 impact_p + \beta_2' dX + \varepsilon_p \tag{C.1}$$

where X includes the same set of control variables added in rows (3) and (7) in Table 3.2.

As dependent variables, we use both the ratio of out-of-pocket medical expenditures to the total household expenditures and the log of out-of-pocket medical expenditures; Table C.4 presents the results thereof. The estimated coefficients are small and not statistically significant. These results suggest that the growth of household out-of-pocket medical expenditures did not vary with the proportion of people newly covered by health insurance owing to the introduction of universal health insurance.

The finding that health insurance had almost no impact on out-of-pocket medical expenditures is in stark contrast to those of studies of health insurance effects in the United States. For example, Finkelstein and McKnight (2008) found that the introduction of Medicare produced a 25-percent decline in out-of-pocket medical expenditures. This difference may be attributable to the difference in the coinsurance rate: in the case of Japan, newly covered NHI recipients still had to pay 50 percent of their own health care costs, whereas the introduction of Medicare reduced consumer costs to almost zero, save for a small deductible.

Table C.1: Variable Definitions and Data Sources

Variable name	Definition	Source
Admissions	Total number of new admissions in the calendar year. All hospitals, not including clinics.	(B)
Inpatient days	Total inpatient days (sum of days in the hospital of all patients) in the calendar year. All hospitals, not including clinics.	1950-51:(A) 1952-70:(B)
Outpatient visits	Total number of outpatient visits in the calendar year. All hospitals, not including clinics.	1950-51:(A) 1952-70:(B)
Expenditures by the NHI	Total healthcare expenditures paid through the NHI (i.e. total healthcare expenditures excluding out-of-pocket spending).	(I)
Number of medical claims	Number of claims made to the NHI by medical institutions.	(I)
Hospitals	Number of hospitals, all kinds, as of December 31	(D)
Clinics	Number of all clinics as of December 31.	(D)
Age specific mortality rates	Total number of deaths of people in the age group divided by population of the same age group interpolated from Census. Per thousand population.	(E) and (F)
Tooth cavities	Ratio of students who have tooth cavities. Based on mandatory medical examination of all students in elementary and junior high school students.	(J)
Physicians	Number of doctors who were working in hospitals as of December 31.	(D)
Nurses	Number of nurses (incl. practical nurses) who were working in hospitals as of December 31.	(D)
Beds	Total number of beds in hospitals and clinics, as of December 31.	(D)
Bed occ. rate	Bed occupancy rate, inpatient/365/number of beds as of July 1	(B)
Total population	Population as of October 1. For years 1950, 55, 60, 65 and 70, taken from Census. Data of inter Census years are interpolated by the Statistics Bureau.	(E) with interpolation
GDP deflator	Prefecture level GDP deflator in the 68SNA system with 1980 as the base year.	(G)
Real GNP per capita	Prefecture level GNP, deflated by prefecture GDP deflator.	(G)
Fiscal rev-exp ratio	Local government's revenue to expenditure ratio. Sum of prefecture and municipal governments. Revenue includes transfers from the national government but excludes transfers between prefecture and municipal governments.	(H)
Fiscal exp on health and sanitation	Local government's expenditure on health and sanitation. Sum of prefecture and municipal governments.	
Population by age group	Population by age group as of October 1. Interpolated from Census.	(E) with interpolation

Data sources:
(A) Japan Statistical Year Book, Bureau of Statistics
(B) Hospital Report, Ministry of Health and Welfare
(C) Annual Statistical Report of National Health Conditions, Health and Welfare Statistics Association
(D) Survey of Medical Institutions, Ministry of Health and Welfare

(E) Population Census, Bureau of Statistics
(F) Vital Statistics, Ministry of Health and Welfare
(G) Prefecture SNA in 68SNA format, available at http://www.esri.cao.go.jp/jp/sna/kenmin/68sna_s30/main.html
(H) Annual Report on Local Public Finance Statistics, Ministry of Home Affairs
(I) Annual Report on Social Security and Statistics, General Administrative Agency of the Cabinet
(J) School Health Survey, Ministry of Education, Science, Sports and Culture

Table C.2: The Effect of the NHI Expansion on the Changes in Self-employment Ratio 1955-1960

	All prefectures		Excl. Tokyo and Osaka	
	(1)	(2)	(3)	(4)
$Impact_p$ defined by equation (2)	-0.005	-0.010	0.001	-0.004
	[0.018]	[0.015]	[0.018]	[0.150]
Changes in Self-emp. ratio 1950-1955		0.389***		0.431***
		[0.104]		[0.097]
Observations	46	46	44	44
R2	0.00	0.19	0.00	0.26

Note: Robust standard errors are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table C.3: The Effect of the NHI Expansion on Establishment Size

A. All Prefectures

	log(number of establishments with 1-4 employees)	log(number of establishments with 5-9 employees)	log(number of all establishments)	% establishments with 1-4 employees	% establishments with 5-9 employees
λ_{51}	-0.064	0.001	0.069	-0.067	-0.012
	[0.087]	[0.134]	[0.228]	[0.118]	[0.018]
λ_{54}	0.034	0.046	0.029	0.005	-0.001
	[0.043]	[0.046]	[0.036]	[0.015]	[0.006]
λ_{60}	-0.059	-0.135*	-0.051	0.007	-0.006
	[0.047]	[0.069]	[0.046]	[0.015]	[0.005]
λ_{63}	-0.043	-0.115	-0.097*	0.040*	-0.010
	[0.048]	[0.095]	[0.053]	[0.020]	[0.008]
λ_{66}	-0.013	-0.224*	-0.094*	0.064**	-0.020*
	[0.052]	[0.119]	[0.056]	[0.027]	[0.010]
Observations	276	276	276	276	276
R-squared	0.999	0.999	0.997	0.91	0.988

B. Excluding Tokyo and Osaka

	log(number of establishments with 1-4 employees)	log(number of establishments with 5-9 employees)	log(number of all establishments)	% establishments with 1-4 employees	% establishments with 5-9 employees
λ_{51}	-0.062	0.063	0.129	-0.106	0.000
	[0.094]	[0.135]	[0.247]	[0.124]	[0.018]
λ_{54}	0.011	0.075	0.026	-0.009	0.005
	[0.046]	[0.051]	[0.039]	[0.013]	[0.006]
λ_{60}	-0.017	-0.049	-0.016	0.004	-0.003
	[0.037]	[0.052]	[0.040]	[0.014]	[0.004]
λ_{63}	-0.054	-0.05	-0.074	0.014	-0.001
	[0.058]	[0.101]	[0.063]	[0.013]	[0.007]
λ_{66}	-0.005	-0.081	-0.032	0.022	-0.008
	[0.064]	[0.109]	[0.066]	[0.019]	[0.009]
Observations	264	264	264	264	264
R-squared	0.997	0.998	0.992	0.823	0.975

Note: Standard errors, estimated with clustering by prefecture, are presented in the brackets. *, **, *** indicate statistical significance at the 10%, 5% and 1% levels, respectively.

Table C.4: The Effect of Universal health Insurance on Households' Out-of-pocket Medical Expenditure

	Ratio of medical expenditure in household expenditure		Log(medical expenditure)	
	1959-1964	1959-1969	1959-1964	1959-1969
$Impact_p$ defined by equation (2)	-0.002	-0.003	-0.037	-0.237
	[0.004]	[0.011]	[0.203]	[0.481]
Observations	46	46	46	46