Quasi-Experimental Approaches to Health and Utilization

Keio University
Graduate School of Business and Commerce
Ph.D. Thesis
Reo Takaku

Outline

Outline x

Chapter1 1
Introduction

Chapter2 10
Maternal Labor Supply, Childcare Provision and Child Health:
Regression Discontinuity Evidence from Japan

Chapter3 76
The Effect of Patient Cost Sharing on Health Care Utilization among Low-income Children

Chapter4 104
Does Reduced Patient Cost-sharing Improve Child Health ?

Chapter5 146
児童手当が両親の心理的健康に与える影響:
中低所得世帯における検証

Chapter6 (Reference Paper) 165
Hospital Response to Financial Incentives:
Evidence From Nighttime EMS in Japan

Acknowledgement 205

Chapter 1

Introduction

1 Credibility Revolution and Health Economics

Empirical strategy in micro econometrics has changed significantly over the past 25 years. Identification strategy has achieved a central role in empirical analysis, and strategy clarity is considered a core of identification. This movement arose when empirical economists started searching for credible experiments generated in real-world environments. Beforehand, many econometricians implicitly assumed that the best way to check the credibility of non-experimental results was to explore the fragilities of the non-experimental parameters by controlling many different sets of covariates. Leamer (1985) clearly expressed his optimistic views on sensitivity analysis, which can be used to control covariates systematically and not arbitrarily. After criticizing ad hoc sensitivity analysis of previous studies, he wrote "what we need instead are organized sensitivity analyses. We must insist that all empirical studies offer convincing evidence of inferential sturdiness," concluding

> An epidemic of overparameterization debilitates our data analyses. We need strong medicine to combat this disease. I know a global sensitivity analysis is a bitter pill to swallow. But try it, I think it's going to make us all feel much better. Maybe not entirely well, but better anyway.

The same optimistic sentiments on sensitivity checks were expressed in "I Just Ran Two Million Regressions" by Sala-i Martin (1997). The title of his AER paper eloquently shows how econometricians once believed the effectiveness of sensitivity checks for estimating causal parameters. In order to find factors associated with GDP growth, Sala-i Martin (1997) applied the extreme-bound tests proposed by Leamer (1985). His strategy amounts to finding covariates that show the estimated coefficients are stable even if other covariates are controlled. However, no empirical economist today is satisfied with such a remedy. Instead, we find that many different sets of covariates may not lead to identification of causal parameters. Such parameters may only be identified by finding a plausible source of exogenous variation from natural- and quasi-experiments in the absence of experimental randomization.

Angrist and Pischke (2010) call this change the "credibility revolution" , which breaks out a revolution in the manner of writing micro-econometrics papers in many fields. Today, many researchers spend an increased amount of time considering better research designs rather than making the econometric specification functional form sophisticated and implementing hundreds of sensitivity checks. Angrist and Pischke (2010) express their view briefly that " a clear-eyed focus on research design is at the heart of the credibility revolution in empirical economics. " Although the revolution has not yet swept macroeconomics and industrial organization, it may have already been completed in labor, public, and development economics.

In addition, statistical methodologies, which originate from epidemiological studies, have been extensively applied to economic empirical analysis. For instance, the difference-in-differences (DID) approach, which was incorporated into economic analysis in the mid- and late- 1980s, is an application of clinical outcome research for social science. As a result of the acceptance of epidemiological statistical tradition, complex empirical specifications have gradually become regarded as a sign of a desperate struggle to overcome problematic and demanding assumptions in identification. Instead, many empirical researchers have increasingly relied on naive OLS with clustered-robust standard errors[1]. Two major empirical strategies for identifying causal effects from observational data, namely DID and regression discontinuity design (RDD), have also been implemented using OLS without particular computational difficulties. Thus, researchers have focused their efforts the validity of the research design, because the interpretation of the results using DID and RDD are very sensitive to the identification assumptions made.

The credibility revolution started sweeping health economics after the turn of the 21st century pursuant to pioneering works in labor economics(Angrist and Krueger, 1991; Card and Krueger, 1994) and public economics (Gruber, 1994). This revolution seems to have completely changed the quality of empirical research in health economics research, as it has in the precedent fields. This is particularly noticeable in the Handbook of Health Economics Volume 1 published in 2000, as literature reviews in some of the book ' s chapters may have limited relevancy to recent academic arguments. Because of the rapid accumulation of theoretical and empirical literature, some chapters in the old handbook may have become obsolete within a decade of publication. Thus, the editors of the Handbook of Health Economics Volume 2 published in 2011 also acknowledged that "methods in health economics have evolved over the past ten years" and that "standards for identifying causal relationships have been elevated" (Mark V. Pauly and Barros, 2011). In addition to the incorporation of behavioral economic theories into healthcare analysis, the rapid evolution in health economic during the past 15 years seems to have been led by the credibility revolution in econometric analysis.

To explore the essence of the revolution, we can determine the main advantages of clear research design by considering randomization. In contrast to most micro-econometric literature without randomization, randomized controlled trials (RCTs) provide the causal effects of interest solely through a simple mean comparison of outcome variables between the treatment and control groups. Given that it is often difficult for empirical economists to choose which variables to control and that the empirical results often change based on the controlled variables[2], the advantage of RCTs seems obvious. As Griliches (1986) notes, econometrics

[1]A symbolic example that shows the nature of this change is the explanation on the two-stage least-square (2SLS) estimator in famous econometrics books. For instance, Hayashi (2000) notes that the 2SLS estimator is a special case of the GMM estimator, but Angrist and Pischke (2009) emphasize that the 2SLS estimator can be derived as a ratio of two OLS estimates: instrumental variable coefficients in reduced-form OLS and first-stage OLS.

[2]This is the reason why Leamer (1985) proposed his extreme-bound analysis.

would be unnecessary if perfect data could be obtained and randomization could be exploited for all questions we are trying to answer.

However, a fundamental limitation of the RCT is its huge cost and potential ethical problems for implementation. For instance, the total cost of the RAND Health Insurance Experiment, a notable experimental study implemented for health economic studies, amounted to 295 million USD in 2011 dollars (Greenberg and Shroder, 2004). Because of the cost concerns, implementing an RCT for every social scientific study is not feasible. Instead, a growing body of literature has exploited quasi-experiments. The advantage of quasi-experimental design is that it provides estimates similar to RCT with low implementation costs. In general, quasi-experimental studies utilize shocks and distortions in the real world generated by exogenous interventions. In some cases, the government seems to be a plausible outsider that exogenously intervenes for players. Although the first generation of the credibility revolution used macroeconomic regional conditions as exogenous shocks for individual behaviors, these shocks seldom contributed to valid causal identification because they hardly met exclusion criteria. Hence, recent papers have increasingly focused on the exogenous variation resulting from arbitrary intervention created by governments.

2 First Motivation: Application of the Credibility Revolution to Healthcare in Japan

The main contribution of my dissertation is to apply the credibility revolution to analysis on healthcare in Japan. Such an application is not new but still of particular importance because few quasi-experimental studies on the Japanese healthcare system have been conducted. Many studies on the Japanese healthcare system lack a clear identification strategy and simply compare the outcomes in different groups. Hence, policy implications from these studies are very limited (Ikegami, 2014). A low contribution of health economics to health policy in Japan is disappointing for our society, because economics should contribute significantly to population dynamics and the financial environment in Japan. For instance, Japan will face rapid ageing under tight fiscal conditions within 10 to 20 years. This suggests that there is no room for implementing inefficient policies, and thus, we must choose effective policies from many potential alternatives. Even today, these social pressures create demand for evaluating the costs and benefits of policies. The credibility revolution helps meet such demand, because it provides sophisticated statistical guidelines to evaluate the effectiveness of social programs (Angrist and Pischke, 2009). Through this dissertation, I try to follow these guidelines as faithfully as possible. In this sense, my dissertation is an application of existing methodologies, but it will contribute to the understanding of the policy effects with certain validity.

Here, I do not indicate the importance of “evidence-based policy” as my motivation. As explained in Kenjoh (2001) and Ikegami (2014), in Japan, policy is usually derived from power rather than from evidence. Hence, our empirical results likely do not directly affect the ultimate design of a given policy. Policymakers that quote our findings likely do so because our findings provide evidence favorable to them. Rather, I think that the main function of program evaluation by statistical specialists is to check the consequence of a policy introduced as a result of political power balances. By reviewing consequences, I hope that increasingly accurate knowledge on program effects will be shared among the population, resulting in higher accountability required from future politicians[3].

Following this motivation, I restrict my evaluation to the effects of various policies using standard program evaluation methodologies from the literature. My dissertation includes five chapters and attempts to uncover the effects of policies such as the "Medical Subsidy for Children and Infants" (MSCI) and child allowance using standard methodologies such as DID and RDD. In this sense, my dissertation is an orthodox attempt to show the possibility of applying program evaluation methodologies to healthcare in Japan.

3 Second Motivation: Child Health in Japan

The second motivation is related to the fields rather than to the methodologies. In four out of five chapters, I present an analysis on child health and healthcare utilization. This is because there is only a small volume of policies directed to households with children in Japan. As reported in the Ministry of Health and Welfare (2012), social expenditure for households as a percent of GDP is one-third that in France and Sweden. In addition, child poverty rates have been rising since the mid-1990s (OECD, 2012). In 2010, the poverty rate for children under 18 years in Japan was 15.7%, similar to those of Canada and Italy but almost double the rates of Germany and Sweden.

This is a problematic issue from a health perspective. First, increasing poverty and politically inactive attitudes for households with children may result in inequality in childhood health status. Hence, understanding the effectiveness of policies directed toward children with low socioeconomic status is critical. In addition, we should not underestimate the importance of changes in disease structures among children. During the past 50 years, infectious diseases have substantially diminished, and survival rates for children with cancer, congenital heart disease, leukemia, and other conditions have greatly improved. However, chronic conditions such as asthma, allergic dermatitis, and mental health problems have been increasingly prevalent among children. In Japan, the incidence rate of asthma was only about 1 percent in elementary school-aged children in 1990, but it increased to over 4 percent by 2012. Although there are no comprehensive statistics, anecdotal evidence suggests that mental health problems such as attention deficit hyperactivity disorder (ADHD) may be increasing, although such an increase may partly be the result of better awareness[4]. Given that these chronic conditions may be concentrated among children from low-income households (Perrin et al, 2014), recently increasing poverty among children may have accelerated health inequality across

[3] A similar view is expressed in Ikegami (2014), who presents a somewhat passive view on the role of intellectuals. However, Kenjoh (2005) expresses his view in a rather active style, hoping that intellectuals may affect the "high-spirited heresies" (in Japanese: Kigai no aru Itann Tachi) and thus help change policies. This is interesting, because both Kenjoh (2005) and Ikegami (2014) have the same view on the policy formation process. However, in either view, the possibility that policy analysis directly affects policy is clearly denied.

[4] These changes indicate that mortality rate, which is a standard health outcome, is no longer an adequate measure for population health.

socioeconomic status.

To improve the health of underprivileged children, we must first discern the effectiveness of alternative policies. For instance, the national government provides child allowance for households with children, and local governments subsidize out-of-pocket costs for child healthcare utilization. Next, we investigate whether an additional one billion JPY should be used for child allowances or medical subsidies. I did not find sufficient evidence in previous studies to answer this question, because few studies directly evaluate the effects of these policies. Hence, I construct my dissertation to fill the gap between high political needs and the existing literature and to find clear guidelines for future policy.

4 Outline and Summary

With the fulfillment of these needs as motivation, I address the following issues in this dissertation.

- Effects of Maternal Employment on Child Health (Chapter 2)
- Effects of Medical Subsidies for Children and Infants on Healthcare Utilization (Chapter 3)
- Effects of Medical Subsidies for Children and Infants on Child Health (Chapter 4)
- Effects of Child Allowances on the Psychological Health of Parents (Chapter 5)
- Manipulation of Hospital Arrival Time by Emergency Care Providers (Chapter 6)

In Chapter 2, I explore the association between child health and maternal employment to understand the social costs of increasing maternal employment. Over the past 3 decades in Japan, the number of women who participate in labor force and have children has gradually increased. According to the Labor Force Statistics, labor force participation rate for married women aged 30-34 was only 44.1% in 1985, but it increased to 55.6% by 2012 (Cabinet Office, 2013). However, existing knowledge on the cost and benefit of the increased labor force participation of married women is fairly limited, because a relatively small number of studies have tried to disentangle the "causal effect" of maternal employment on child health. Exploiting an exogenous shock to maternal employment as an instrumental variable, this chapter identifies the causal effect of maternal employment on child health in Japan. Specifically, I focus on the fact that many women exit the labor market when their firstborn child enrolls in elementary school because of shortage of afterschool childcare. This problem is notorious in Japan and is called the "wall for mothers with first graders." Facing such a "wall," mothers reduce their labor supply and younger siblings receive more parental care after the firstborn child enrolls in school. This institutional setting provides plausible variation to estimate the impact of reduced maternal labor supply assuming no direct effect of school enrollment for the firstborn child on the health of younger siblings. This empirical strategy provides a novel regression-discontinuity estimate based on unique Japanese institutional settings. In addition, this chapter may provide direct policy implications, because the national government in Japan is developing various new policies to encourage maternal employment.

In Chapters 3 and 4, I investigate the effects of reduced patient cost-sharing on healthcare utilization (Chapter 3) and health (Chapter 4), exploiting the expansion of the MSCI as a natural experiment. In Japan, the national co-insurance rate is 20 percent in preschool children and 30 percent in school-age children, but municipalities reduce these rates at their own financial expense. This subsidization program has been dramatically expanded in the last decade. In Chapter 3, I explore how the expansion of MSCI in Hokkaido prefecture affected the utilization of healthcare services among preschool children using insurance claims in a city. In addition, the effect on health status is examined in Chapter 4 using six waves of the Comprehensive Survey for Living Conditions, from 1995 to 2010.

Chapters 3 and 4 also provide important contributions on how patient cost-sharing is associated with healthcare utilization and patient health. Thus far, most existing knowledge on these issues results from the RAND Health Insurance Experiment (Manning et al, 1987), which is the most comprehensive randomized experiment on the effect of cost-sharing. However, the experimental results of 30 years ago would not be directly applicable today, especially to other countries. In addition, few studies have addressed the impacts of patient cost-sharing on child health. Because childhood health status is widely recognized as an important determinant for future achievement and health, creating new reliable estimates in the context of child health is of particular importance. The MSCI expansion in Japan provides a golden opportunity to investigate the real-world effects of patient cost-sharing among children.

In Chapter 5, the effects of the reform of child allowance led by the Democratic Party in 2010 is examined. The identification strategy is based on the fact that the increase in child allowance in 2010 differed across households according to the number of children and their age. For instance, the cash transfer from child allowance increased by 60,000 JPY for households with one school-age child, while it increased by 180,000 JPY for those with three school-age children. Hence, a situation occurs where one mother receives a large cash transfer but another mother receives less. Using these differential exogenous increases in child allowances, I estimate the effects of cash transfers on various mental health measures. Here, the investigated outcome is not that of children but of their mothers. For evaluating the effect of child allowance, child outcomes?such as test scores and health metrics?act as firsthand measures. However, maternal mental health is deeply associated with child outcomes, and children cannot be happy if they have an unhappy or discouraged mother. In this sense, the improvement of maternal psychological health and happiness may be an important channel through which child allowance improves child outcomes. This perspective provides useful insight for the evaluation of child allowance in Japan, because previous studies on the effects of child allowance in Japan consider only the effects on consumption, implicitly assuming that child allowance improves child outcomes through increased purchases of goods and services for children. This view seems to have very limited relevancy given the large overall effect of child allowance. Although this chapter is written in Japanese and forthcoming in Quarterly of Social Security Research (*Kikan Shakai Hosyo Kenkyu*), I add it to this dissertation because the scope and results of this study are relevant to other chapters.

The final chapter is separate from the others. In Chapter 6, coauthored with Atsushi Yamaoka, we attempt to uncover the association between hospital behavior and financial incentives. Although the theme is not related to child health, we adopt a careful identification strategy to uncover the causal effect of financial incentives. In this sense, this chapter is mainly based on the first motivation, namely the application of the “credibility revolution” to healthcare in Japan.

Concerning the identification issue, it is difficult to identify the causal effects of financial incentives on the behavior of healthcare providers because unobservable patient characteristics are associated with the choice of medical facilities, which are often subject to different payment systems such as fee-for-services (FFS) and capitation payment (CAP). Hence, a comparison of the treatment provided under FFS and CAP cannot provide a causal effect of payment systems on treatment choice, because hospitals with different payment systems receive different patients. In this chapter, we overcome this problem by focusing on a manipulative behavior among hospital care providers in Japan. Specifically, reimbursement for hospital care in Japan is linked to the number of “midnights” a patient stays in the hospital. Then, we argue that this “midnight-to-midnight” method may provide an incentive for healthcare providers to accept emergency patients before midnight, because they generate additional reimbursements through an extra night of hospitalization compared to those who arrive after midnight. We test this prediction using all 2.1 million administrative records of emergency medical transportation around midnight in Japan from 2008 to 2011. Given that the occurrence of emergency episodes is random, the detection of manipulating hospital arrival time may help clarify the effects of financial incentives.

References

Angrist JD, Krueger AB (1991) Does compulsory school attendance affect schooling and earnings? The Quarterly Journal of Economics 106(4):979–1014

Angrist JD, Pischke S (2009) Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press

Angrist JD, Pischke S (2010) The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of Economic Perspectives 24(2):3–30

Card D, Krueger AB (1994) Minimum wages and employment: A case study of the fast-food industry in new jersey and pennsylvania. American Economic Review 84(4):772–93

Greenberg DH, Shroder M (2004) The Digest of Social Experiments, third edition edn. Urban Institute Press., Washington, DC

Griliches Z (1986) Economic data issues. In: Griliches Z, Intriligator MD (eds) Handbook of Econometrics, Handbook of Econometrics, vol 3, Elsevier, chap 25, pp 1465–1514

Gruber J (1994) The incidence of mandated maternity benefits. American Economic Review 84(3):622–41

Hayashi F (2000) Econometrics. Princeton University Press

Ministry of Health L, Welfare (2012) White Paper

Ikegami N (2014) Rethinking health economics and policy researches [*aratamete iryou keizai seisaku kenkyuu wo kangaeru*]. Japanese Journal of Health Economic and Policy [*Iryou Keizai Kenkyu*] 26(1):1–2, URL http://www.ihep.jp/publications/study/search.php

Kenjoh Y (2001) Political Economy of Redistributive Policy (1)—Social Security and Healthcare in Japan, first edition edn. Keio University Press, Toyko, Japan

Kenjoh Y (2005) Korunakare gakumon 36: For high-spirited heresies being in any fields [in japanese; *dono sekai nimo iru kigai no areu itan tachi he*], URL http://news.fbc.keio.ac.jp/~kenjoh/work/korunakare36.pdf, (Accessed on March 5, 2014.)

Leamer EE (1985) Sensitivity analyses would help. American Economic Review 75(3):308–313

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A, Marquis S (1987) Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review 77(3):251–77

Mark V Pauly TGM, Barros PP (2011) Handbook of health economics. In: Mark V Pauly TGM, Barros PP (eds) Handbook of Health Economics, Handbook of Health Economics, vol 2, Elsevier, pp i –, DOI http://dx.doi.org/10.1016/B978-0-444-53592-4.00017-7

Sala-i Martin X (1997) I just ran two million regressions. American Economic Review 87(2):178–83

OECD (2012) Oecd family database: Child well-being module, URL http://www.oecd.org/els/family/oecdfamilydatabasechildwell-beingmodule.htm, (Accessed on March 5, 2014.)

Cabinet Office (2013) White paper on gender equality 2013, URL http://www.gender.go.jp/about_danjo/whitepaper/h25/zentai/html/zuhyo/zuhyo01-02-01.html, accessed on July 23, 2014

Perrin JM, Anderson LE, Van Cleave J (2014) The rise in chronic conditions among infants, children, and youth can be met with continued health system innovations 33(12):2099–2105

Chapter 2

Maternal Labor Supply, Childcare Provision and Child Health: Regression Discontinuity Evidence From Japan*

Reo TAKAKU†

Abstract

In Japan, mothers are likely to exit from labor market when their eldest child enrolls in elementary school because of many institutional barriers such as shortage of after school childcare. Using the eldest child's enrollment in elementary school as an exogenous shock to maternal labor supply, this paper explores how health of the younger preschool siblings responds to the decreased maternal labor supply. Using a regression discontinuity design, I marginally compare preschool children whose eldest sibling enrolls in elementary school or remains in preschool. The results show the maternal employment rate drops by 4-5 percentage points after the eldest child's school entry. In addition, reduction of maternal labor supply leads to an increase of parental care for the younger siblings. As a result of substantial decreases in maternal labor supply and increasing parental care, the probability of taking a "fever" decreases among the younger siblings, suggesting reduction of maternal labor supply improve child health. However, there seem to be no improvements on the other subjective and objective measures of child health such as the incidence of injuries and hospitalization. Taken together, this paper indicates that the reduction of maternal labor supply is associated with improvement of the health of preschool children, but the magnitude is not large at least in the short run.

Keywords : child health, maternal employment, regression discontinuity design, Japan

JEL classification : I0, J21,J13,C26

*This paper was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (Research Project Number:25780199) The views expressed in this paper do not necessarily reflect the views of the Ministry of Education, Culture, Sports, Science and Technology and Institute for Health Economics and Policy. I am grateful for the comments and suggestions from Michihito Ando, Yukiko Asai, Yoshikazu Kenjoh, Sayaka Nakamura, Kazunari Shinpo, Isamu Yamamoto and participants in the applied econometric conference in Osaka University and the annual meeting of Japanese Health Economic Association. Because of several revisions, this version is different from the final one which was submitted to academic journals. All errors are my own.

†Institute for Health Economics and Policy. Email:reo.takaku@ihep.jp

1 Introduction

Over the past three decades, the number of women with children who participate in labor forces has increased gradually in Japan, as well as in the other developed countries. According to the Labor Force Statistics, labor force participation rate for married women aged 30-34 was only 44.1% in 1985, but it had increased to 55.6% by 2012 (Cabinet Office, 2013). However, what we know on the cost and benefit of the increasing labor force participation in married women is fairly limited. In particular, relatively small number of studies has tried to disentangle the "causal effect" of maternal employment on child health, which would be potentially the major cost or benefit of increasing employment of mothers.

To uncover the causal effect of maternal employment, this paper employs a quasi-experimental strategy, exploiting the unique Japanese institutional settings on the childcare availability. My research design relies on the fact that mothers in Japan experience discontinuous reductions of childcare availability when their children enroll in elementary school on April at the age of 6[1]. Although school hours in the first and second grade are very short and there are also long seasonal vacations in elementary school, the availability of after school childcare for school-age children is quite limited. Such discontinuous reduction of childcare availability is notorious in Japan, and called "a Wall for the Mothers with First Grader" (in Japanese: *Shōgakkō Ichinensei no Kabe*). Facing the "wall", many mothers exit from labor market to provide after school childcare, even if they had worked when their children were preschool.

Exploiting this *Shōgakkō Ichinensei no Kabe*, the present paper establishes a novel Regression Discontinuity (RD) evidence on the impact of maternal employment on child health. Specifically, I explore how the health status of the *younger* preschool siblings changes before and after the eldest child's enrollment in elementary school. The intuition of this strategy is that the younger siblings are likely to receive maternal care because of mother's exit from labor market, when the eldest child has just enrolled in elementary school. If the other covariates are smooth around the cut-off month, the observed changes in health status at the same month should be attributed to the reduction of maternal labor supply and increasing parental care. This RD design is applied for the 6 waves of Comprehensive Survey of Living Condition from 1995 to 2010, which is the nationally representative sample of Japanese population.

Three findings are followed. First, RD estimates show that maternal employment rate significantly drops by 4-5 percentage points just after the eldest child's school entry. Given that the employment rate just before the cut-off month is 41%, the size of the employment loss due to *Shōgakkō Ichinensei no Kabe* is substantial. Second, as a result of reduced maternal employment, the younger siblings are more likely to receive parental care in weekday. If the quality of parental care may be higher than any other types of childcare, these findings suggests the quality of care received by the younger preschool siblings may be discontinuously improved. Third, reduction of maternal employment and increase of parental care improve health status of these siblings, with a significant reduction of children who take a "fever". However, I do not find significant

[1] Fortunately for my study, the school admission date is strictly enforced with almost complete compliance. Kawaguchi (2011) reports the percentage of children who cannot admit elementary school at April 2 is only 0.03 percent. As is mentioned in Shigeoka (2014b), this low exemption rate sharply contrasts with the situation in U.S.

improvements on the other subjective and objective measures of child health such as incidence of injuries and hospitalization. These results suggest that, in the short run, the decrease of maternal labor supply and increase of parental care are not associated with serious health conditions among preschool children. Finally, my main findings survive after the several robustness checks.

The renaming paper is organized as follows. Second section provides brief background of this paper such as prior literature reviews and institutional background on after school childcare in Japan. Section 3 gives explanations on the standard methodologies and RD design. Section 4 describes data and the definition of outcome variables. Section 5 shows the results from standard methodologies such as OLS and conventional IV. Next, section 6 summarizes results from RD design. In section 7, placebo tests which exploit potential timing of the treatment are implemented for a robustness check. Finally, concluding remarks are presented in section 8.

2 Background

2.1 Prior Literature

This subsection summarizes the review of prior literature which considers the identification of causal effect of maternal employment on child outcomes, mainly on child health. First studies which tried to explore the causality have employed maternal fixed effect to address omitted variable bias (Waldfogel et al, 2002; Anderson et al, 2003; Aizer, 2004; Ruhm, 2004; Aughinbaugh and Gittleman, 2004; Gordon et al, 2007). Using fixed effect model, unobservable factors which affect both mother's labor supply and child health (e.g. preference of mother) can be successfully eliminated if they were time-invariant. As Gordon et al (2007) mentioned, this assumption would be sufficiently plausible if additional covariates do not change the coefficient of maternal employment to a large extent. However, it is ad-hoc and far from completely plausible because we cannot include all potential covariates into our regression. In addition, it is impossible to address simultaneous bias or reverse causality with the inclusion of additional covariates.

Instead, some papers have employed instrumental variable technique. Nevertheless the conventional IVs such as female unemployment rate have failed to uncover the causal effect because of the low explanatory power (von Hinke Kessler Scholder, 2008; Cawley and Liu, 2012; Gwozdz et al, 2013; Datar et al, 2014; Wüst, 2014). The problem of weak instrument in this issue is widely recognized. For instance, Cawley and Liu (2012) wrote in their conclusion that "an important direction for future research is to find valid and powerful instruments for maternal employment, and investigate whether maternal employment has the causal effect of reducing mother's time spent in activities"[2].

Given the stream of existing literature, recent studies have exploited quasi-experimental changes in maternal employment in order to uncover causal effect of maternal employment on child health (Baker and Milligan, 2008; Gennetian et al, 2010; Morrill, 2011)[3]. These experimental studies have presented no or

[2]This sentence is also cited in Gwozdz et al (2013).

[3]Berger et al (2005) also try to uncover the causal effect using propensity score methods. However, as is pointed out in Ruhm

negative effect of maternal employment on child outcomes, while the results varies greatly. For instance, Baker and Milligan (2008) exploit the expansion of mandatory maternal leave, comparing the changes in outcome variables between pre- and post-expansion. Their paper reveal that the expansion of mandatory maternity leave in Canada reduced employment rate among mothers after birth and sharply increased the duration of breastfeeding, while they also show the impact on subjective health of mothers and children and children's cognitive ability would be generally weak[4]. The more strong negative effects are observed in Morrill (2011) who focuses on the fact that labor participation rate discontinuously increases when the mother's youngest sibling is eligible for kindergarten[5]. The IV estimates in Morrill (2011) indicate that maternal employment increases overnight hospitalizations by 4 percentage points, injuries and poisonings by 5 percentage points, and asthma episodes by 12 percentage points, each by around 200 percent. On the other hand, the study which exploits randomization shows more modest negative effect of maternal employment. Gennetian et al (2010) use the experimental data from a welfare-to-work program implemented in the early 1990s in the US, showing that a percentage point increase in employment induced by a welfare reform program decreases the probability of a child being in very good or excellent health by 0.6 percentage points.

Finally, as for the studies which use the sample of Japanese population, Tanaka (2008) investigates how maternal employment affects the educational attainment of children, using household data from the Japanese General Social Survey (JGSS). Although he finds maternal employment has negative effects on the children's educational attainment, but the empirical strategy is based on OLS, assuming maternal employment status is exogenously given.

2.2 After School Childcare and Maternal Labor Supply in Japan

Employed mothers arrange a variety of alternative cares to ensure safe after school environment. In general, Japanese children in lower grade levels appear to be in home care with parental supervision, compared with other developed countries, because use of out-of-school services are not common in Japan. According to an international comparative survey from OECD (2011), 80-90 % of children in the Nordic countries such as Denmark and Sweden use out-of-school-hours care services, while the rate is only 11.2 % in Japan. This rate is higher than Germany and Italy, but lower than France, UK and Canada[6]. The reason why is partly because of low labor participation rate in women in Japan, especially in the women with children. Historically, married women have been a main provider of after school care and supervision.

Recently, this tendency has been gradually changed as women with children have participated in labor

(2014), the results obtained from propensity score methods are likely to be sensitive to the choice of covariates and assumption to balance the characteristics between treatment and control.

[4]Based on the same research design, Baker and Milligan (2010) find weak impact of the increased maternal care on child's developmental outcomes, while the reform crowded out the home-based care by unlicensed non-relatives.

[5]There is another study which exploits variation of maternal labor supply which is caused by the exogenous changes in one of the siblings. Bettinger et al (2014) exploit the introduction of a program which was intended to give incentives for parents to stay home with children under 3 years in Norwegian, investigating the effect on the older siblings. They find a significant positive treatment effect on older siblings ' tenth-grade GPA, while their mothers reduced labor force participation.

[6]Recent international comparison study for children aged 8 to 13 years old(Akashi et al, 2014) also shows the likelihood of being alone or without adult supervision after school is higher in Japan than in Germany, UK, Franc and Korea.

market. Nevertheless, primal political interests have been on expansion of childcare services for preschool children, rather than school-age children. Indeed, the shortage of after school childcare has been regarded as a minor issue compared with that of preschool childcare. In the economic researches, this tendency is almost the same: Although there have been huge public debates on the considerable shortage of licensed childcare services for preschool children and how it constrains women's labor supply[7], there are no comprehensive studies on the shortage of after school childcare for first and second graders in elementary school. In the legalization process, provision of childcare for preschool children was legalized by the Child Welfare Act in 1947, while it was 1997 when provision of after school childcare was legalized. The lack of public attention to expand after school care inevitably results in the considerable shortage of care. As childcare for preschool children has been gradually expanded, the gap in childcare availability before and after school entry has increased rapidly and become like a "cliff". ASCLC (2013) estimates the utilization rate of after school childcare was only 67 % of the new students who had been left to daycare center when they were preschool, suggesting the large underling demand for after school care.

Inevitably, such a discontinuous reduction of childcare availability works as a strong barrier which keeps mothers away from labor market, although no research reveals the exact amount of employment loss due to this barrier [8]. On the other hand, there are also sufficient anecdotal evidences that suggest the discontinuous reduction of childcare availability, which is named "a Wall for the Mothers with First Grader" (in Japanese: *Shōgakkō Ichinensei no Kabe*), sharply decreases mother's participation for labor market.

In order to reduce the gap in the childcare availability before and after the school entry, prime minister Shinzō Abe recently has emphasized the importance of increasing after school childcare, including the policy to expand childcare for school-age children into his "growth strategy" published in June 2014 (Prime Minister of Japan and His Cabinet, 2014). Specifically, it intends that the supply of after school childcare will increase from current 0.9 million to 1.2 million by the end of 2019 fiscal year. According to the prime minister's explanation, this policy will effectively eliminate the current gap between availability of childcare before and after child's admission to elementary school and lead to increasing female labor supply.

3 Empirical Strategy

To clarify the nature of RD methodology, this paper compares the results from traditional methods with RDD. In this section, I explain my RD strategy after introducing standard methodologies.

[7]See Zhou and Oishi (2005), Lee and Lee (2014) and Kawabata (2014). Zhou and Oishi (2005) estimates underling demand for licensed childcare services and find that the size of underling demand was 111% of provided amount of child care services. Lee and Lee (2014) and Kawabata (2014) investigate the effect of childcare availability on mother's employment. In particular, Kawabata (2014) finds childcare provision for mothers with children under 3 years old helps to augment their participation for labor market.

[8]JILPT (2013) presents the average employment rate of mothers based on the age of the youngest child and find the reduction of employment rate when children admit to elementary school, although the number of observation used in the study is small.

3.1 Standard Methodology

Most analysis which investigate the effect of maternal employment on child outcome begin their studies with estimating simple OLS, based on following equation,

$$H_{it} = a_0 + a_1 MEMP_{it} + X_{it} a_2 + \epsilon_{it}, \tag{1}$$

where H_{it} is a child outcome such as health status, $MEMP_{it}$ is a variable which denotes maternal employment, X_{it} is a vector of covariates and ϵ_{it} is an error term. In this equation, it should be noted that there is no consensus which variable should be used for $MEMP_{it}$. Some studies utilize mother's working hour as $MEMP_{it}$ (Anderson et al, 2003; Gordon et al, 2007; Datar et al, 2014), while the other utilize a binary variable which captures whether a mother worked or not (Morrill, 2011; Cawley and Liu, 2012; Wüst, 2014) or a categorical variable which varies according to working status such as full-time or part-time (Bernal and Keane, 2011; Gwozdz et al, 2013). Among these variables, I use a binary variable which captures extensive margin of labor supply because of the limitation of the data. Even if which variables are used, we can obtain unbiased estimate of a_1 once we control all variables which potentially affect dependent variable.

However, the condition of such unbiased estimate is so demanding that many studies try to address the empirical inconsistency which accrues from the use of OLS[9]. Among them, the conventional method which is applied repeatedly is an instrumental variable technique which utilizes region-level female unemployment rate as an IV for maternal employment. Although some studies criticize the use of this IV because of several serious problems, it is conventionally applied in order to check the robustness of the the main results from OLS or maternal FE (Anderson et al, 2003), frequently failing to uncover causal effect of maternal employment due to weak instrument (von Hinke Kessler Scholder, 2008; Cawley and Liu, 2012; Gwozdz et al, 2013; Datar et al, 2014; Wüst, 2014).

To replicate their results, I estimates two-stage least square model with regional female unemployment rate as an IV, based on following first stage regression,

$$MEMP_{it} = b_0 + b_1 UNEMP_{rt} + X_{it} b_2 + \varepsilon_{it}, \tag{2}$$

where $UNEMP_{rt}$ is a local female unemployment rate in year t and region r, which is obtained from the Annual Labor Force Survey[10]. On this conventional IV specification, we should note that there are serious threats for the exclusion restriction even if the first stage F statistics would be enough high to reject the null hypothesis of weak instrument[11]. First, Cawley and Liu (2012) insist that local macroeconomic

[9] Mother fixed effects are widely used to address omitted variable biases if we have longitudinal data (Anderson et al, 2003; Ruhm, 2004; Aizer, 2004; Wüst, 2014), but my data are from a repeated cross sectional survey. Hence, the results from mother FE cannot be presented.

[10] Regional classification is not based on 47 prefectural border, but 10 conventional regions. The data are from http://www.stat.go.jp/data/roudou/longtime/03roudou.htm (Accessed on July 10, 2014.).

[11] Following a conventional wisdom, the instrument is not weak if the F statistics is over 10.

condition would not only affect endogenous maternal labor supply but also affect child health directly. This is a main reason why female unemployment rate is not valid instrument. Second, the exclusion restriction would be violated because we can measure the amount of maternal employment only partially because of the limitation of data. For instance, some papers measures maternal labor supply by the extensive margin (i.e. whether a mother works or not), assuming intensive margin is not affected, but local macroeconomic environment would affect working hours (intensive margin) as well as employment itself. Since mother's working hours would be relevant to child outcomes, the violation of exclusion restriction may exaggerate the impact of maternal employment[12].

3.2 Regression Discontinuity Design

3.2.1 Identification Assumption

Since low availability of after school childcare makes mothers exit from labor market in Japan, the younger siblings are likely to receive more parental care when the eldest child enrolls in elementary school. Exploiting the discontinuous reduction of the mothers' participation in labor market, this paper implements a RDD analysis to estimate the effect of maternal employment on child health. The RDD in this paper is implemented with the elder sibling's age in month as an assignment variable. In addition, since the treatment does not necessary affect maternal employment with full compliance, "fuzzy" RD, rather than "sharp" RD is employed (Lee and Lemieux, 2010). Before introducing the empirical specification, I should clearly state two identifying assumptions. First, my RDD requires that underlying characteristics of the younger preschool distributes continuously around the threshold, namely the month of the the eldest child's enrollment in elementary school (*continuity assumption*). Second, the eldest child's enrollment in elementary school must affects child health status only through maternal employment (*exclusion restriction*).

On the first assumption, we should note that the characteristics of children may not be continuous around a school-entry-age cutoff date since parents may manipulate birth timing. This potential problem is intrinsic to the any school-entry-age cutoff RD (Dobkin and Ferreira, 2010; McCrary and Royer, 2011; Shigeoka, 2014b). In Japan, parents may want to deliver after April 2 because of the potential advantage of cognitive and non-cognitive achievement over the those born in March (Kawaguchi, 2011; Shigeoka, 2014b)[13]. In addition, birth timing may exhibit natural seasonality[14] and reflect socio economic backgrounds of the parents, which are not observable. Then, marginal comparison between children born around school-entry cutoff date necessary reflects these unobservable characteristics. Although this problem is clearly minor if

[12] The latter threat would be crucial even if we exploit quasi-experiments. Off course, suitable quasi-experiment provides an identification which assures that the IV affects outcome variables only through maternal labor supply, but it is very difficult to find a quasi-experiment which affects only the extensive margin of maternal labor supply. For instance, in Morrill (2011)'s study, the endogenous variable is a dummy which takes one if the mother participates in labor market, implicitly assuming the treatment does not affect working hours and shift time.

[13] Kawaguchi (2011) shows test score is higher among older children than younger counterparts in 4th to 8th grade. Shigeoka (2014b) finds that there is a considerable manipulation of birth timing around April 2 using the universe of births during 1974 to 2010 and there effects are heterogeneous, showing births by younger mothers, 2nd-born births, and male births are more shifted than births by older mothers, 1st-born births, and female births.

[14] Kawaguchi (2011) suggests that farmers are likely to have birth in winter because of low work load.

birth month of the younger siblings is independent from that of the eldest child, we should again note that birth month of the eldest child may reflect household-level characteristics which must affect the younger siblings.

Whatever the reason, heaping and seasonality in birth months is a challenge for RDD analysis, since the Comprehensive Survey for Living Conditions is only conducted in one day in June. To address this threat, I control birth quarter fixed effects. Although it is possible to include birth month fixed effects, but due to limited range of the bandwidth, quarter rather than month effects seem to absorb seasonality of birth timing accurately even under the relatively small bandwidth[15]. It should be noted that, due to the consideration of seasonality in birth timing, my analysis deviates from a conventional wisdom that recommends to choose smaller bandwidth (Hahn et al, 2001). For instance, we cannot shorten the bandwidth within 6 months around the threshold since the coefficient of quarter birth dummies cannot be estimated. When heaping and seasonality bias would be relevant, as in the school-entry cutoff RD, large bandwidth combined with seasonal dummies would be the second best strategy since small bandwidth would make non-random heaping bias more severe (Barreca et al, 2011).

Second assumption is exclusion restriction that the reduced form effects of the eldest child's enrollment in elementary school on health of the younger siblings would come only from the changes in maternal employment. Given many other potential paths through which mothers take to adjust they labor supply than the adjustment in the extensive margin, this assumption may be somewhat demanding. For instance, mothers may reduce working hours for their part-time jobs and the others may shift the working time to mid-night or early morning to provide after school childcare when their child admits to elementary school. In addition to the adjustment in the decision whether or not to exit from labor market, these adjustments may also affect child health. This potentially violates the exclusion restriction of my strategy. Even if it is the case, however, we should note that the reduced form estimate reveal the meaningful total effect of maternal labor supply, which is caused by the eldest child's enrollment in elementary school. Although I am aware of the difficulties to detect the entire response in mother's labor supply for the treatment, it seems to be a reasonable assumption that the treatment does not affect other mediators than mother's response in labor market. For instance, the treatment may not affect housing choice and father's contribution for childcare. In addition, if health condition of the eldest sibling is highly responsive to the environmental changes due to the enrollment in elementary school, health of the younger siblings are possibly affected by inter-household infection. However, we do not have enough reasons to believe that it is the case. At least, I believe that mothers' decision on how much they work is a major and primal mediator which accounts for the effect of the eldest child's school entry on the health status of the younger preschool siblings. Thus, throughout this paper, I present reduced form estimates and IV estimates together and try to provide careful interpretations.

[15] From the same reason, Shigeoka (2014a) controls birth month fixed effects in order to rule out seasonality of birth. Note that his age-based RDD utilize relatively wide bandwidth (10 years from 65 years to 75 years), but this paper uses much smaller bandwidth (6 years before and after school-entry cutoff month).

3.2.2 Econometric Specification

Following the discussion above, I begin with estimating following equation to examine how the eldest child's school entry constrain maternal employment, controlling seasonality and heaping in birth timing of the eldest sibling by birth quarter fixed effects,

$$MEMP_{it} = \alpha_0 + \alpha_1 D_{it} + f(Z_{it}) + X_{it}\theta + \tau + year + pref + \zeta_{it}, \quad (3)$$

where D_{it} is a binary variable which takes one if the eldest child is school-age and otherwise zero, X_{it} is a vector of covariates and ζ_{it} is an error term. $f(Z_{it})$ is a polynomial function of age in months of the eldest child (Z_{it}), which is specified as,

$$f(Z_{it}) = \sum_{k=1}^{n} \left((Z_{it} - C)^k + D_{it}(Z_{it} - C)^k \right), \quad (4)$$

where, n is an order of polynomial. In this specification, C is the cut-off value which standardizes the term $Z_{it} - C$ as zero when the eldest child admits to elementary school on April at the age 6. Specifically, C is set as 75 since the survey I use in this study is held in June every 3 years[16]. X_{it} is a vector of covariates which includes age of the children and ζ_{it} is an error term. $year$ and $pref$ are year effects and prefecture fixed effects, respectively.

In addition, τ is a birth quarter fixed effects which absorb seasonality and heaping in the birth timing. If the season of birth of the eldest child reflects unobservable characteristics of the household, the seasonal pattern is supposed to be stable across cohorts. Hence, it may be absorbed into τ with a sufficiently long bandwidth. In the baseline specification, I estimate this equation with a bandwidth of ± 36 months from a cut-off month and first order polynomial. And then, several specification checks are implemented. Following a conventional wisdom, two checks are implemented (Lee and Lemieux, 2010). First the results from narrower windows such as ± 24 months and ± 12 months are presented. Second, the results from quadratic and cubic polynomial are presented with the baseline bandwidth fixed at ± 36 months. Compared with standard approaches, the setting of bandwidth here may be wide because of the necessity of controlling birth quarter fixed effect correctly.

Some other remarks are stated. First, since this equation is applied for preschool-aged children who have at least one elder sibling, a child exits from my analysis when he newly enrolls in elementary school. For instance, with a bandwidth of ± 24 months, the younger siblings born a year after the eldest child drop from analysis when $Z_{it} - C = 12$ since they would enroll in elementary school. Second, for the calculation of standard errors, I calculate the standard errors that are clustered at the eldest child's age in months since

[16]Because child admits to elementary school on April at the age 6 (6 years old = 72 months), the youngest age of the children who are allowed to admit is those who are 72 months on April 1 every year. These children will be almost 75 months when the survey is held at the end of June.

conventional standard error which does not take into account discreteness of assignment variable tends to overestimate the precision of the estimated effects (Lee and Lemieux, 2010).

Following the estimation of the first stage effect of the eldest child's admission to elementary school on maternal employment, the same equation is applied for the child health outcomes (H_{it}). This reduced form equation, which captures an intention-to-treat (ITT) effect (Lee and Lemieux, 2010), directly measures the impact of the eldest child's school entry on child health. Formally, following equation is used,

$$H_{it} = \beta_0 + \beta_1 D_{it} + g(Z_{it}) + X_{it}\delta + \tau + year + pref + +\eta_{it}, \tag{5}$$

where $g(Z_{it})$ is a polynomial function of age in months of the eldest child and η_{it} is an error term.

Since the model is exactly identified, 2SLS estimates, which capture the causal effect of maternal employment on child health, are numerically identical to the ratio of two coefficients ($\frac{\beta_1}{\alpha_1}$) if the same bandwidth is chosen for equation (3) and (5), and the same order of polynomial is used for $g(\cdot)$ and $f(\cdot)$. Throughout this paper, I do not use optimal bandwidth calculation proposed by Imbens and Kalyanaraman (2012) because of discreteness and limited range of our assignment variable. Since our assignment variable is age in month and this variable consists of only 72 discrete values (36 months before and after enrollment in elementary school) at the maximum, we prefer to provide RD estimates with varying bandwidths rather than to present a single RD estimate with an "optimal" bandwidth based on additional assumptions.

Finally, the ITT effect in equation (5) captures the total effect of the treatment on health outcomes. It can be reasonably assumed that this total effect is generated through the changes in maternal labor supply and the other paths may play minor role. For instance, emigration to seek better public school is possible at the time when the eldest child enrolls in elementary school, but certainly such behavior is not common.

4 Data

4.1 Comprehensive Survey of Living Condition

I use one of the most comprehensive data of children's health status in Japan. The Comprehensive Survey of Living Condition is a nationally representative survey of stratified random sample of Japanese population. This survey has been conducted every three years since 1986 and there are 11 rounds available under the permission of Ministry of Health, Labour and Welfare. From all rounds, I construct a set of repeated cross section data, polling the data of preschool children from 6 rounds from 1995 to 2010.

Child health variables are obtained from the health questionnaire. In addition, I also use the household questionnaire which contains broad household characteristics such as composition of household, working status of parents and in which public health insurance the child is covered. These household characteristics are combined to child health data with 100 % matching.

From complete data set, I exclude children who receive public welfare program and those with lone parents. Furthermore, children without siblings are dropped because my identification strategy is based on the change in maternal labor supply due to the eldest child's school entry. Without siblings, this strategy cannot be applied. Off course, such a restriction of sample would inevitably narrow external validity of the analysis. In addition, children whose parents are over 70 years and below 20 years, and children under 6 months are also excluded. It should be noted that OLS and conventional IV are also applied for the sample without excluding firstborn children. For the RD design, after excluding them, I choose the observations within each bandwidth.

4.2 Outcome Variables

4.2.1 Maternal Labor Supply

In the first stage RD estimates, the maternal labor supply is measured through its extensive margin. Using the question "Do you work currently with remuneration?", I create a binary variable which takes one if a child's mother answers yes for this question and otherwise zero. Since the CSLC does not ask working hours, effect on the intensive margin for labor supply is not investigated. Instead, the CSLC contains the question on the types of employment contract only for the respondents who work with remuneration. Using this question, the types of employment contract are classified into 4 groups; self-employed and workers for family business, general employee[17], employee with short-term employment contract and the people with the other job contract such as an executive of firm. Among them, employee with short-term employment contract includes those who work with the contract less than 1 year. If these women are likely to exit from labor market, the share of them may decrease when the eldest child enrolls in elementary school.

4.2.2 Childcare Provision

Note that mother's exit from labor market does not necessary lead to increase of parental care for the younger siblings because the new first grader in elementary school may require more care from parents than in preschool-age. Hence, it is possible that the amount of parental care for the younger preschool siblings does not increase largely even if their mother quit job. To answer this question, I focus on the choice of childcare provider. In Japan, childcare for preschool children is provided by parents, daycare center, relatives and kindergarten. Among them, enrollment in kindergarten is not allowed for child under 3 years old. In addition, note that a child cannot go to kindergarten without extensive care from parents in Japan because of short stay hours. As a result, children are left daycare center if their parents have full-time job. In short, childcare system in Japan implicitly supposes that a child grows up with stay-at-home parenting by 3 years old and goes to kindergarten after April at the age of 3, if one of their parents, actually their mother, does not have full-time job. If their mother have full-time job or need to work for long hours, their children are allowed to be left on daycare center.

Based on this childcare system, I construct two variables to measure the amount of parental care. First,

[17] Part-time employee is included in general employee if their contract is over 1 year or not based on effective contract.

children are defined as receiving parental care if they go to kindergarten or their parents can provide care in daytime. Note that this variable can take one even if the parents use daycare center and own parental care together. In this sense, the first variable is a generous measure of parental care because a child is regarded as receiving parental care even if their parents take care of him only one day of a weak. Second variable tries to capture the intensive utilization of daycare center. I classify a child receives care only from daycare center if he goes to daycare center and the parents do not provide any care for him in daytime. If this variable takes one, the child is regarded as heavily relying on daycare center. Through investigating the changes in these variables, I evaluate the effect of the eldest child's school entry on childcare provision for the younger preschool siblings.

4.2.3 Child Health

Subsequently, child health status is measured through three outcomes. The first is subjective symptom which is measured through a question that "In the last few days, have you experienced any symptoms of illness or injury?". On behalf of preschool children, parents answer this question[18]. Second outcome is a binary variable which captures current outpatient visits due to injury such as fracture. This outcome is divided into fracture and other injuries such as cut and skin burn. In the absence of supervision from parents, children may more and more experience injuries as a result of dangerous activities without supervision, especially in the case of infants. Empirically, Currie and Hotz (2004) show the quality of supervision is related to the incidence of unintended injuries among children under age 5. Morrill (2011) also reports maternal employment is related to the incidence of injury and poisoning. Provided that quality of maternal care is higher than that of childcare facilities, the increasing maternal care would decrease the incidence of these accidents. Third outcome is the probability of hospitalization which is a major and one of the most standard outcome which measures child health. This variable takes one if parents answer that their child admitted to hospital at the time the survey was held, and otherwise zero. Since admission to a hospital requires the judgment of a medical professional, hospitalization is regarded as an objective measure of health status[19].

4.3 Descriptive Statistics

The summary statistics is presented in Table 1. In this table, I report the summary statistics in all preschool children and the RD sample, separately. The full sample is used for OLS and conventional IV analysis. For the RD sample, the sample includes preschool children whose eldest sibling is aged 36 months before and after the school entry. In addition, firstborn child is excluded in the RD sample. The number of observation is 148,699 in full sample and 57,211 in the RD sample, respectively. The number of observation for the probability of symptom and visit are less than the maximum because these variables are observed for the children who did not admit to hospital. The number of observation in Panel B, where the summary statistics of the employment contract is presented, is small. This is because these variables are observed when the

[18] Although subjective health is commonly used for empirical studies, CSLC asks it only for school-age children.
[19] Since CSLC does not survey the reason of hospitalization, our measure includes all admissions to hospital.

mother has worked. In Panel C, the probability of receiving any parental care is about 65 percent and the utilization rate of daycare center is 25 percent. Since the data for childcare are available since 1998, the sample size for these variable is smaller than other variable in Panel A. Finally, it should be noted that the means in the two samples are almost identical in all variables, regardless of the small sample size in RD analysis.

On the covariates, child's age in month, gender, age of head and spouse, working status of household head, number of siblings and total household members and insurance plans are controlled. In Japan, children are covered under the same health insurance plan as their designated household head. Broadly, there are three types of health insurance plans for working-age adults in Japan: Society-Managed Health Insurance (SMHI); a health insurance plan managed by the National Health Insurance Association (NHIA); and Citizens' Health Insurance (CHI), which is a residence-based health insurance plan. These three plans account for almost 90%[20] of health insurance for those under 75 years of age. Adults who work for large firms participate in SMHI, whereas those who work for small and medium enterprises are included in the NHIA. Other adults must obtain coverage through the CHI in their residential area[21]. Based on these institutional settings, 4 binary variables are created to control the types of health insurance the children are enrolled[22]. The summary statistics for the covariates are reported in Table A1 in the Appendix.

5 Results from Standard Methodologies

5.1 OLS

I begin my analysis with presenting the results from OLS in Table 2. Column (1) reports the coefficient of maternal employment in the full sample of preschool children. These results correspond to the results from routinely-used approaches which try to uncover correlation between maternal employment and child health among all preschool children, without restricting the analysis to local subpopulation. And then, for the convenience of the comparison of results from RDD, results from restricted samples are reported. From Column (2) to (5), the eldest children in each household are excluded and bandwidth is also changed from 36 months in Column (3) to 12 months in Column (5). In all equations, we include full covariates described in Appendix Table A1, as well as year effects and prefecture fixed effects, and standard errors are clustered at prefecture-level.

The coefficients in Column (1) reveal striking negative correlation between maternal employment and child health, with the coefficient of 0.015 for the probability of having any symptoms being significant at 99 percent confidence interval. This suggests that children whose mother participates in labor market are more likely to feel symptoms of illness or injuries, compared with the counterparts whose mother does not work.

[20]The remaining 10% is included in Mutual Aid Associations that cover those employed in the public sector.
[21]A comprehensive and historical review of the Japanese health care system is provided in Ikegami et al (2011) .
[22]Since benefit package such as co-insurance rate is unified across all insurance plans, the difference in child health across plans cannot be attributed to the insurance policies. Rather, these variable implicitly control the occupational choice of household head.

Given the mean of the probability, the coefficient suggests maternal employment increases it by 5.6 percent (0.015/0.265=0.056). In addition, I find significant negative effects in major 4 items, namely fever, cough, wheezing and stuff nose. On the fever and wheezing, the coefficients are robustly significant once the eldest children are excluded and the bandwidth becomes narrower. On the other hand, there may be no significant effect on the probability of being injured and hospitalized. All the coefficients of maternal employment on these outcomes are not significant.

5.2 Conventional Instrumental Variable

Next, the results from conventional IV, which utilize region-level female unemployment rate as an IV, are presented in Table 3. As is in the results from OLS, I present several results based on alternative inclusion criteria from Column (1) to (5). In the IV analysis, we should note that the regional unemployment rate is sufficiently relevant to maternal employment, at least in the full sample results in Column (1). The first stage F statistics are over 10, suggesting the IV is not weak. Given that many studies try but do not report IV results since region-level female unemployment rate is likely to be weak in the first stage (von Hinke Kessler Scholder, 2008; Cawley and Liu, 2012; Gwozdz et al, 2013; Datar et al, 2014; Wüst, 2014), it is of importance to note that region-level female unemployment rate meets the standard requirement in the first stage. Hence, the results in Column (1) are of importance in the sense that this column reports the results which these previous studies would have obtained. In the subsample analysis from Column (2) to Column (1), however, the instrument seems to be weak, with the F statistics below 10. This is probably because mothers with one child, who are excluded in the subpopulations, are more responsive to macroeconomic conditions than those with two or more children. Then, I cannot compare the results from conventional IV with those from RDD.

In Column (1), the results from the conventional IV seem to be consistent with those from OLS. The conventional IV estimates present significant effects of maternal employment in all selected symptoms. However, the coefficients are too large and less precise. For instance, the point estimate on the probability of having any symptoms suggests that the probability would increase by 97.7 percent with mother being employed. This magnitude is clearly unrealistic, suggesting the violation of exclusion restriction. Major concern here is that region-level female unemployment rate affects various aspect of female labor market such as working hour, as well as mother's participation in labor market, and all these aspects affect child health. This implies that region-level female unemployment rate is correlated severely with the error term in the structural equation of interest[23].

Regardless of these limitations, the results from OLS and conventional IV are fairly similar in full sample analysis. In both methodologies, maternal employment is associated with the increasing probability of having symptom but not with the probability of visits due to injuries and hospitalization.

[23] If high unemployment rate improve child health directly (Ruhm, 2000) (e.g. through reduction of traffic accident and air pollution), the exclusion restriction is also violated.

6 Results from RDD

6.1 Identification Checks

Before moving to the RD results, I present two standard validity checks (Lee and Lemieux, 2010). First, I examine whether the density of the assignment variable, age in month of the eldest children, is continuous at the discontinuity. Since age in month is not continuous but discrete, I implement parametric version of McCrary (2008)'s density test[24]. Second, I examine the discontinuities in all covariates by using the same parametric regression. In order to implement these tests, the data are collapsed into survey year and age in month. With the bandwidth of 36 months, the sample size is 432 (6 years * 36 months * 2). In addition, my tests address the potential seasonality and heaping in birth timing by controlling birth quarter fixed effects. This is a point which is different from standard application of parametric test of discontinuity, but necessary for causal identification since parents may choose the timing of birth according to their socio economic characteristics, as is explained in previous section. Far from completely plausible, we can assume that these sorting of birth timing are controlled by birth quarter fixed effect if they are stable across cohorts. In this sense, the test here examines the discontinuity conditional on seasonality of birth timing of the eldest child. Finally, the following equation are applied,

$$y_{it} \quad = \quad c_0 + c_1 D_i + h(Z_i) + \tau + year + \kappa_{it}, \tag{6}$$

where y_{jt} represents a variable of interest aggregated in age group j by year t, $h(Z_{jt})$ is a polynomial function of age profile and κ_{jt} is an error term. If the distribution of the y_{it} is smooth around the threshold, we can expect c_1 is not different from zero. For the density test, I count the number of children included in the analysis by the age in month of the eldest child and then this number is regressed with the discontinuity term and the polynomial function.

I show the bin-mean plot of the number of children who are included analysis in Figure 1-(a). The x-axis of the figure represents age of the firstborn children which is standardized 0 at the month when they enroll in elementary school and y-axis represents the count of the younger siblings in each bin. The lines are the quadratic fit for the number of observations. For the standardization, I distract 75 from age in months of each child. Since school admission is April at the age of 6 (72 months) in Japan, the age of the youngest first graders is 75 months at the end of June, when the CSLC is in field. Hence, "0" in x-axis means that the firstborn children are the youngest in all first graders in elementary school and "-1" means that they are the eldest among preschool children. The parametric test here is based on the marginal comparison between these age groups.

Next, when value of x-axis is -36, which means age of the firstborn children is 36 months before their school entry, the count is only around 400, but it gradually increases up to about 1,100. The upward slop

[24] The original McCrary's density test is for the RD design with continuous assignment variable, rather than discrete variable.

in the count is because the younger siblings had not born yet when the firstborn child was in very small age. On the contrary, the count decreases to 600 when the value of x-axis reaches to 36 because the younger siblings enroll in elementary school and drops from the analysis. Although the figure shows a non-linear pattern of the count, it also suggests that the count may be smooth at the cut-off month (X-axis = 0).

Other figures in Figure 1 present the same plot of main covariates such as child age, sex and age of household head. Since the running variable is age of the firstborn children, the mean age of the younger siblings also increase as the running variable increases (1-(b)). However, there is no systematic jump around the threshold in this figure. We also find that the share of girls and age of household head are completely smooth. On the other hand, we may find discontinuity in the threshold in Figure 1 -(e). Importantly, the number of household member exhibits small jump, suggesting the eldest child's school entry may increase the number of household. If some household decided to live together with grandparents of the children at the timing of school entry in order to provide after school childcare, this irregular jump may be plausible, threatening one of the important assumptions in this study. However, we should also note that the number of household member show some regular patterns (waves) across the assignment variable. By looking at the short period from "0" to "18", in particular, I find that the waves in Figure 1 -(e) seem to be consistent with those in the age of household head in Figure 1 -(d). This finding reflects the fact that elder parents are likely to have elder grandparent and live with them together. Hence, it is reasonable to interpret the jump as a result of household-level sorting across birth timing. Again, we can check this point by incorporating birth quarter fixed effect.

The results of parametric tests are presented In Table 4. estimates in Column (1) to (3) include a linear trend of running variable and it's interaction term with the discontinuity term (a dummy variable which takes 1 for the eldest child's school entry). For the bandwidth selection, Column (1) includes 36 months before and after the cut-off month. Subsequently, the bandwidths are narrower in Column (2) and (3). Estimates in Column (4) to (6) test the robustness of the results with higher order of polynomial such as quadratic (Column 4), cubic (Column 5) age profiles and their interaction with the discontinuity term.

In the Panel A, where the results of parametric density test are reported, the coefficients are not statistically significant in all columns, suggesting there is no bunching in the sample size, conditional on birth quarter fixed effects. This supports the validity of my RDD. In the Panel B, I examine the discontinuities in covariates which are included in the main analysis. In the regression on head's working status, the coefficient of the discontinuity term is estimated significantly, but most of them are less precise and are not robust for alternative specification. In addition, I find that other covariates are smoothly distributed around the threshold. Although 1 -(e) suggests number of household members bunches around the threshold, the parametric test for discontinuity does not reject the hypothesis of no discontinuity. This result largely comes from the fact that I control unobservable characteristics of household which is correlated with the birth timing of the eldest child by controlling birth quarter fixed effects[25].

[25] Without controlling these effects, the parametric test find significant discontinuities in the number of household members.

Although the results from these specification checks secure the smoothness of covariates conditional of birth quarter fixed effects, it is somewhat ad hoc and far from completely plausible. Then, after introducing main results, I provide an additional robustness checks by applying "donut-hole' RD" (Barreca et al, 2011) in Appendix B, excluding the observations near the threshold. There are two reasons why "donut-hole' RD" provides meaningful robustness checks for my analysis. First, it provides a well-established test to check heaping induced bias in RD estimates, although there is no consensus on the optimal size of the donut. Second, as is discussed in Shigeoka (2014a), it addresses inter-temporal substitution of labor supply around the threshold. For instance, mothers are likely to work harder in the previous year of the eldest child's school enrollment because they have to reduce working hours after the enrollment. If it is the case, RD estimates for maternal employment may capture the magnitude of such substitution, rather than local randomization around the threshold. To address potential bias from inter-temporal behavior, as well as heaping bias, robustness of the main results are checked with "donut-hole' RD".

6.2 Effect on Maternal Employment

Once the RD design passes the identification checks, variation in the treatment near the threshold is regarded as if they were randomized. Based on such a local randomization, causal effects of treatment on outcome variable would be revealed.

First stage effect on maternal employment is graphically presented in Figure 2 with the corresponding estimates summarized in Table 5. In Figure 2, the share of the younger siblings whose mother worked with rewards is plotted by the standardized age in month of the firstborn children. We show, in this Figure, a clear evidence on the discontinuous reduction of maternal employment when the eldest child enrolls in elementary school. First, maternal employment rate increases from about 30 % to 40 % when the age of the firstborn children increases from 36 months before school entry to the cut-off month. Nevertheless, just after the cut-off month, maternal employment rate drops by around 5 percentage points, and again begins to increase. These findings are plausible as the effect of *Shōgakkō Ichinensei no Kabe* and directly show that mothers are likely to exit from labor market at the timing of the eldest child's school entry.

Next, Table 5 reports the corresponding RD estimates, with and without covariates. First, 5 naive RD estimates without covariate are thoroughly significant, with the coefficients ranging from -0.025 in Column (1) to -0.081 in Column (3). These estimates suggest that maternal employment rate dropped by 2 to 8 percentage points when the eldest child enrolls in elementary school. In addition, the results are robust for the inclusion of covariates. The point estimate ranges from -0.032 to -0.078. This is a substantial decrease in the employment rate; from the average in the last 6 months before the cut-off month (41%), the probability of being employed decreases by 8% to 17%. In addition, the size of the employment change is as large as that found in Morrill (2011)[26]. However, the magnitude of the reduction depends on the choice

[26] As in this paper, Morrill (2011) utilizes a discontinuous increase of maternal employment rate when the youngest siblings are eligible for kindergarten. First stage coefficients in Morrill (2011) range from 4-8 percentage points, which are the same absolute values with my paper.

of bandwidth. To check the robustness for alternative bandwidth more carefully, Appendix Figure C1 plot the RD estimates from linear age profile by the length of bandwidth. The figure suggests that RD estimates from the bandwidth of 12 months may be irregular. On the other hand, in many bandwidth from 13 months to 36 months, point estimates are stable around 4-5 percentage point (Figure C1-(a), (b) and (c)).

Again, shortage of after school childcare in Japan may explain the observed reduction of maternal employment. In the absence of proper after school childcare, mothers with lower graders in elementary school seem to exit from labor market even if they had participated when their children were preschool.

6.3 Effect on Types of Employment Contract

In addition to the reduction of mother's participation in labor market, types of the employment contract can change before and after the cut-off month. If part-time workers and workers with short-term contract are more likely to exit, we would find the share of these workers discontinuously decreased after the eldest child's school entry. The results are summarized in Figure 4 which plots the age profile of the share of 4 alternative employment contracts and corresponding table is presented in Table 6. The figure and the RD estimates do not exhibit no sign of systematic changes in the type of employment contract, but point estimates suggests rough pattern on the characteristics of workers who exit. The sign of RD estimates are positive in self-employed and general employee, but negative in employee with short-term contract and other workers. Given that self-employed workers may not quit job, positive signs on this group are consistent with prediction. In addition, the point estimates on short-term employee are negative, suggesting this group may have high tendency to quit job after the eldest child's school entry, while the estimates are far from precise. The low precision of the estimates are partly explained from the rough classification of types of workers in CSLC. Importantly, by definition, part-time and full-time workers are included in general employee together. Then, we cannot observe labor supply response in part-time workers separately from full-time workers. In addition, some part-time workers would have misreported themselves as "other employee" because "general employee" implicates full-time workers. Probably consistent with this prediction, the RD estimates also suggest the reduction of employment rate in "other employee", indicating part-time workers are likely to exit from labor market around the threshold.

6.4 Effect on Childcare Provision

Does the reduction of maternal labor supply due to *Shōgakkō Ichinensei no Kabe* lead to increase in parental care for the younger preschool siblings ? To answer this question, I turn to investigate the parental response on the childcare provision. As is explained above, two dependent variables are constructed. The results are summarized in Figure 5 and Table 7. First, Figure 5 -(a) shows a discontinuous increase in the probability of receiving parental care around the threshold, although there seem to be substantial time-series patterns. On the contrary, the share of children who are left in daycare center and do not receive parental care at all in daytime slightly decreases after the eldest sibling's school entry. However, as in the figure on the probability of receiving parental care, the age profile may exhibit some cyclical and seasonal fluctuations. These fluctuations are because of the fact that the enrollment in daycare center and kindergarten is allowed in April and the opportunity to enroll in the other month is very restricted.

Parametric RD estimates with the eldest child's birth timing controlled by 4 seasonal dummies, which are summarized in Table 7, show the discontinuity estimates after controlling fluctuated age-profile on the childcare provision. Results from the regressions without and with covariates are reported in Panel A and B, respectively. First, I find the results are not robust for alternative bandwidth selection from Column (1) to (3). This suggests that cyclical fluctuation according to age-profile, observed in Figure 5 -(a), heavily affects the results from RD. The results from Column (4) and (5) also suggest cyclical fluctuation matters. In Column (4), with bandwidth of 36 months, the RD estimates on the probability of receiving parental care show no significant jump. Once the fluctuation is more precisely controlled by cubic polynomials in Column (5), however, the RD estimates turn out to be significant.

Second, while the results on parental care are sensitive for the choice of bandwidth, I find the estimates become more precise once covariates are controlled. In Column (2), the RD estimate without covariates is positive but insignificant but that with covariates is 0.022 and statistically significant, suggesting that we would have significant treatment effects once potential covariates are appropriately controlled. As a result, the RD estimates are robustly significant for alternative bandwidths conditional on covariates and linear age-profile[27].

On the results on utilization of daycare center, the results are much ambiguous. The point estimates are consistently negative, but less precise even if covariates are controlled. Given that mothers who do not work outside are not allowed to use public daycare center in principle[28], a reasonable consequence of the shrinking maternal labor force participation is a significant decrease of utilization of daycare center. Again, these imprecise estimates may be attributed to the difficulties to separate the treatment effect from underlying cyclical pattern, even after controlling the eldest child's birth quarter. Despite these difficulties, however, I show in Appendix C that the point estimates exhibit certain robustness for alternative choice of bandwidth[29]. These results give me some confidences that the results here are totally spurious and driven by "waves" in the utilization rate of daycare center.

[27] See Appendix Figure C2-(a), (b) and (c). In these figures, I present RD estimates from various bandwidth from 12 months to 36 months, one by one. Although some estimates are not significant at 95 percent interval, lower bound of the confidence interval is not below zero so largely.

[28] In Japan, parents who want to use daycare center need to be judged whether they meet official criteria set by the government. This official criteria consists of various requirements such as household income, working status and health of parents and help from relatives. If mother works full-time, they are likely to be allowed to use public daycare center.

[29] Appendix Figure C2-(a) plots the RD estimates and the 95 percent confidence intervals, changing the bandwidth from 12 months to 36 months. Although the upper bounds of the 95 percent confidence interval seem to be slightly over zero in the bandwidth less than 26 months, all the point estimates are stable and almost significant, ranging from -0.02 to -0.04.

6.5 Effect on Child Health

6.5.1 Subjective Symptom

The bin-mean plot of the probability of having any symptoms is presented in Figure 5. The quadratic fit in the left side of vertical line exhibits a downward slope since children feel less symptom as they grow up and the probability becomes stable around 25% in the right side. On the other hand, there seems to be no significant jump at the threshold. Indeed, at the margin of the threshold, the probability of having symptom seems to be smooth regardless of sudden decrease of maternal employment rate and increasing parental care at the same timing. Table 8 summarizes the results of the reduced form RD specification and the corresponding IV estimation on this outcome. Here, I find the results depend on choice of polynomials and bandwidth; In Column (1) and (2), RD estimates are highly significant, but in the other column they turn to be insignificant.

Although the results on the symptom vary across specifications, we should note the IV estimates in Column (1) and (2) suggest very large impact of maternal employment. For instance, the coefficient in Column (2) is 0.808 which suggest maternal employment induces a 81 percentage point increase of the probability of having a symptom. This magnitude seems to be unrealistically large. As is explained in the results from conventional IV, this overestimation indicates the violation of exclusion restriction. Since the treatment affects various aspect of labor force adjustment over working hours and intensity of the work allocated to the mothers, IV estimates are severely overestimated. However, reduced form effects still provide meaningful information on the effect of such an extensive adjustment of maternal labor force on child health. Hence, from here, I highlight my results based on reduced form results, rather than IV results.

Since the results on all kinds of symptom depend on the specification, I focus on the selected symptoms. Table 9 reports the reduced form estimates on the 10 main symptoms which are consistently asked in the CSLC from 1995 to 2010. In addition, the bin-mean plots are presented for 4 major symptoms in Figure 6. Although the interpretation on the results of each items is too specific and beyond the scope of this paper, the reduced form estimates are generally insignificant. Only in "fever" which may be related to infectious diseases (Column 1), I find negative effects in many specifications. Given increasing parental care and decreasing utilization of daycare center, reduction of "fever" may be plausible since infectious diseases such as common cold are prevalent in daycare center(Silverstein et al, 2003). In the graphical representation, Figure 6 suggests that there seems to be a slight discontinuity around the threshold only in "fever", although this discontinuity may be driven by the irregular reduction at the cut-off month[30].

On the other hand, maternal employment is irrelevant to the major chronic conditions among children such as asthma and allergic dermatitis. In Column (5) which shows the effect on the probability of being "wheezing", IV estimates are not significant in most specifications. In addition, I find no effect on the "rush"

[30]To check the robustness, it is useful to implement "donut-hole RD", excluding the observations near the threshold. The "donut-hole RD" on fever is presented in the Appendix Figure B1. Although the RD estimates are slightly insignificant in two estimations with small bandwidth, the other "donut-hole RD" estimates are significant. Then, I think the reduction of the probability of taking a fever is not spurious.

in Column (9) which may be related to chronic skin problems.

It should be noted that OLS and conventional IV find a significant effect on "wheezing", suggesting maternal employment increases childhood asthma, but my RD estimates do not support it. The difference comes from the different nature of estimates. Since RD estimate captures local average treatment effect (LATE) (Imbens and Angrist, 1994), it can be different from global estimates captured in OLS. Major difference is that OLS estimates may include long-run cumulative effect of maternal employment[31], but RD estimates extract marginal changes in outcome variables which is caused by treatment. In addition, since conventional IV and my fuzzy RD exploit different local shocks on the quantity of maternal labor supply, the results can also be different. However, it is conclusive that reduced maternal labor supply and increasing parental care are not associated with chronic condition at least on the margin, because chronic conditions may not be responsive to short-run environmental changes.

6.5.2 Injury

With parental cares and supervision, children may be kept away from dangerous activities which would result in serious injuries such as fractures. Hence, we can predict that reduced maternal employment and increasing parental care, which are caused by the treatment, would have reduced the incidence of injuries[32]. To examine this possibility, I run the same parametric RD model for the probability of outpatient visits due to injuries as an outcome. This outcome covers outpatient visits for fractures, skin burn and the other injuries. First, the graphical representation offers an intuitive understanding of the results. Figure 7-(a) shows the bin-mean plot of the probability of physicians visit due to all injuries, and then Figures 7-(b) and (c) provide results based on the types of injuries; Figure 7-(b) focuses on the incidence of fracture, which should be regarded as a serious injury, and Figure 7-(c) shows the results on other injuries. On the incidence of all causes of injuries in Figure 7-(a), the profile of the probability seems to be sufficiently smooth around the threshold, while the dispersion is fairly large. In addition incidence of fractures is also smooth without any discontinuous reduction around the threshold (Figure 7-(b)), suggesting that maternal employment is irrelevant to the incidence of serious injuries in Japan, at least on the margin.

The RD results are presented in Table 10. In all specifications, I find no significant effect in reduced form estimates (Panel A). Hence, IV estimates on the effect of maternal employment are also insignificant. Given that OLS and conventional IV suggest no association between maternal employment and these outcomes, my results are robust for alternative specifications. Finally, it should be noted that these results are different from previous results in the US, presented by Currie and Hotz (2004) and Morrill (2011). For instance, Currie and Hotz (2004) argue that unintentional injuries in daycare center is associated with mother's working status and find a regulation on daycare center in the US reduced the incidence. In addition, they find the policy impact is observed only in the children of working mothers because they are likely to use daycare

[31] However, OLS estimates may be biased by omitted variable and reverse causality.

[32] Fujiwara et al (2010) find paternal involvement on childcare reduce the risk of injuries among the children at 18 months in Japan. For instance, they find that taking a child for a walk by the father prevents all cause injuries. As a potential mechanism, they point out an increase of the quality and quantity of childcare by reducing maternal stress.

center more than the children whose mother is not employed[33]. Morrill (2011) estimates more directly the impact of maternal employment on hospitalization due to injuries and poisoning among school-age children and find significant association, while her estimates on this outcomes do not seem to be robust.

The difference in the effect of maternal employment on incidence of injuries between Japan and US are attributed to the quality of daycare center. In the context of the effect of maternal employment on childhood overweight, Greve (2011) has already provided an useful argument. As in my study, Greve (2011) points out that there is no statistical association between maternal employment on child overweight status in Denmark, while existing studies in North America point at rising maternal employment as an explanation for the increasing trend in child weight. According to Greve (2011), these difference may be attributed to the difference in the quality of childcare and father's contribution to children's health. Although father's contribution to childcare is much lower in Japan than in the US, I claim the quality of daycare may partly explain the differences because public daycare center in Japan is stringently regulated, while such a regulation results in the considerable shortage of child care facilities and many mothers do not participate in labor market because they cannot find any vacancy in publicly-licensed daycare center (Zhou and Oishi, 2005; Unayama, 2012)[34].

6.5.3 Hospitalization

Finally, the effects of hospitalization are examined with the same analytical framework as the other outcomes. In the many previous studies, hospitalization is one of the most important outcomes which measure child objective health, although it can be affected by a variety of socio economic environment such as the amount of patient cost sharing and access to medical facilities. However, since these factors would be stable during the short bandwidth before and after the threshold, hospitalization is regarded as a proper measure for objective health status. The results are presented in Figure 8 and Table 11. Contrary to Morrill (2011) who finds large negative effect of maternal employment on child's hospitalization, the effects on hospitalization are not significant in the reduced form estimates, except in Column (5). Only in this column, the estimate show significant and strong effect of the eldest child's school entry on the younger siblings' hospitalization. However, the estimate in Column (5) is not robust for a robustness check based on "donut-hole" RD. Appendix Figure B1 shows that, without donut-hole, the RD estimate are significant at 95 percent confidence interval, but all the donut-hole RD estimates are not significant. This strongly suggests RD estimate in Column (5) may be spurious. Hence, I conclude there is no association between maternal labor supply and child hospitalization. The corresponding bin-mean plot in Figure 8 also shows that the probability of hospitalization is completely smooth around the cut-off month, while the dispersion of the data is somewhat large.

[33] They also note the importance of investigating the relationship between increasing maternal employment and childhood injuries because, compared with test scores, accident rates have direct relevancy with maternal employment.

[34] In Japan, there is considerable shortage in the supply of public daycare center, despite low labor force participation rate among women with children. Although employment rate in the women with youngest child aged 4-6 is about 80% in Denmark(Greve, 2011), but the employment rate in women with youngest child under 6 was only around 40% in Japan (Ministry of Health and Welfare, 2010)

7 Placebo Test

Since RD estimates can varies across alternative choice of bandwidths and polynomial functions, as is shown previously, it is requisite to implement various robustness checks. One popular robustness check for the instability of RD estimates is to implement placebo test by changing the timing of the treatment to an arbitrary point where no environmental changes occur in our experimental settings. If my treatment confounds with unobservable determinants which are also correlated to the assignment variable, we are likely to find significant association between the placebo treatment and the outcomes. On the contrary, if the placebo test finds no significant effect, validity of our results may be enhanced. In particular, the advantage of implementing placebo test is to check the robustness for the seasonal trends of outcome variables which are associated with birth month of the eldest child. For instance, as is mentioned previously, birth timing of the eldest child potentially reflects household characteristics which are related to the outcomes in the younger siblings. Although we roughly control them by incorporating birth quarter fixed effects in the regression analysis, it is far from completely plausible. Here, the placebo test may complement my baseline results since we would find strong significant association between placebo treatments and the outcomes if the seasonality were serious threat for my RD analysis. On the contrary, the seasonal heaping may not give a serious bias for the RD estimate if placebo treatments have no significant association with any key variables.

Based on this idea, I compare the results between "real" and placebo treatment. As a placebo test, I deliberately change the timing of treatment from 2 years (24 months) before the eldest child's school entry to 5 years (60 months) after the"real" cut-off point. The results based on other potential treatments are not reported due to small sample size. For each placebo test, bandwidth is also changed from 12 months to 36 months and linear and cubic polynomials are controlled. Then, I obtain 4,250 RD estimates (25 months * 2 polynomials * 85 treatments) for one outcome variable. To show the results graphically, the average value of t statistics of placebo RD estimates at various potential cutoffs are plotted according to the timing of potential treatment. Here, I report the results on maternal employment, probability of receiving parental care, taking a "fever" and hospitalization, results on other outcomes are presented in Appendix C to save the space.

The results are presented in Figure 9. In figure (a) I find strongest negative impact of the treatment at the "real" threshold, while placebo treatments are generally insignificant. Although the treatment effect is negative and significant at 20 months before the "real" threshold, I do not know plausible reason: This may be an irregular exception which is associated with the eldest child's growth. Rather, it should be noted that there are no large negative impact in the right of the threshold, suggesting that the promotion of the eldest child in elementary school does not induce any shrinking in maternal employment, but the school entry does. Given that seasonality of the eldest child's birth timing is assumed to be stable across the every cohorts, the strong negative impact only found at the threshold indicate that the eldest child's school entry causes the reduction of maternal employment.

In figure (b), the placebo test provides rather ambiguous results on the probability of receiving parental care since some potential treatment effects exhibit relatively high value of t statistics exceeding 1.5. However, I also find the largest impact in the RD estimates with "real" treatment (average t statistics = 1.90) compared to other potential treatments. This again suggests that the treatment effect is likely to be valid. Next, I find significant and negative impact on the probability of taking a "'fever" with the proper treatment, but rarely find with potential treatments. All in all, figure (a), (b) and (c) seem to show that the eldest child's school entry induces reduction of maternal employment and increase in parental care for the younger siblings, and then as a result, the younger siblings become less taking infectious diseases. In addition, this effect is not observed with the other potential treatments. Finally, in figure (d), I find slightly larger negative t statistics with the "real" treatment, compared with the other placebos. However, this t statistics does not suggest significant association between the treatment and hospitalization.

On the further discussion on the placebo test, see Figure C1 to C9 in Appendix C.

8 Discussion

Regardless of increasing labor force participation among women with children in the past 30 years, there are limited studies which disentangle causal effect of maternal employment on child health. Exploiting unique institutional settings in the childcare availability in Japan, this paper shows how reduction of maternal employment affects the health among preschool children. Specifically, Identification strategy in this study is based on the fact that mothers in Japan are likely to exit from labor market to provide after school childcare for their school-age children when they newly enter elementary school because the childcare availability for school-age children is quite limited compared with that for preschool children. Such discontinuous reduction of childcare availability is notorious in Japan, and called "a Wall for the Mothers with the First Grader" (in Japanese: *Shōgakkō Ichinensei no Kabe*). Exploiting the *Shōgakkō Ichinensei no Kabe*, this paper establishes a novel RD evidence on the impact of maternal employment on child health. Indeed, using firstborn child's admission to elementary school as an exogenous shock to maternal employment, I explore how health of the younger preschool siblings responds to the decreased maternal employment rate. The data of child health are from Comprehensive Survey of Living Conditions from 1995 to 2010 which is the stratified random sample of Japanese population, and then fuzzy RD design is applied for them, focusing on the short windows from the eldest child's school entry.

The results show that, in the first stage estimate, the maternal employment rate drops by 4-5 percentage points just after the eldest child's school entry. Given that the employment rate just before the cut-off month is 41%, the size of the employment loss due to *Shōgakkō Ichinensei no Kabe* is substantial. In addition, I find significant increase in parental care for the younger siblings. These findings suggest the quality of care received by the younger preschool siblings may be improved. As a result, heath status of these siblings is improved with a significant reduction of children who take a "fever". However, I do not find significant improvements on the other subjective and objective measures of child health such as incidence of injuries and hospitalization. These results suggest that, in the short run, the decrease of maternal labor force participation and increase of parental care are not associated with serious health conditions among preschool children.

On the other hand, there remain several limitations. First, we should also note that the RD results here do not capture the long-run effect of maternal labor supply. Given that health is capital (Grossman, 1972), slight short-term effect on child health, observed in this study, may be accumulated in the long-run and result in huge deterioration in health status in the later stage of life. Unfortunately this study does not have any clear answers on this very important issue. However, given that there are persistent gradient between household income and health and maternal labor force participation necessary increase household income, long-run detrimental effect of maternal employment on child health may be rather weaker than short-term effect. If it the case, results in this paper lead to an opportunistic view on the extensive increase in maternal employment rate in Japan. Second, this paper focuses on the younger siblings, excluding the eldest. Since many studies shows children of different birth order are raised differently and have different cognitive ability[35], exclusion of the eldest child may limit the external validity of my analysis.

References

Aizer A (2004) Home alone: supervision after school and child behavior. Journal of Public Economics 88(9-10):1835–1848

Akashi Y, Iwasaki K, Kanefuji F, Kobayashi J, Tsuchiya T, Yuki M (2014) International Comparison of After School Activities among Children. Fukumura Publishing Company

Anderson PM, Butcher KF, Levine PB (2003) Maternal employment and overweight children. Journal of Health Economics 22(3):477 – 504, DOI http://dx.doi.org/10.1016/S0167-6296(03)00022-5, URL http://www.sciencedirect.com/science/article/pii/S0167629603000225

Aughinbaugh A, Gittleman M (2004) Maternal employment and adolescent risky behavior. Journal of Health Economics 23(4):815–838, URL http://ideas.repec.org/a/eee/jhecon/v23y2004i4p815-838.html

Baker M, Milligan K (2008) Maternal employment, breastfeeding, and health: Evidence from maternity leave mandates. Journal of Health Economics 27(4):871–887, URL http://ideas.repec.org/a/eee/jhecon/v27y2008i4p871-887.html

Baker M, Milligan K (2010) Evidence from maternity leave expansions of the impact of maternal care on early child development. Journal of Human Resources 45(1)

Barreca AI, Lindo JM, Waddell GR (2011) Heaping-induced bias in regression-discontinuity designs. NBER Working Papers 17408, National Bureau of Economic Research, Inc, URL http://ideas.repec.org/p/nbr/nberwo/17408.html

Berger LM, Hill J, Waldfogel J (2005) Maternity leave, early maternal employment and child health and development in the us. The Economic Journal 115(501):F29–F47, DOI 10.1111/j.0013-0133.2005.00971.x, URL http://dx.doi.org/10.1111/j.0013-0133.2005.00971.x

Bernal R, Keane MP (2011) Child care choices and children ’ s cognitive achievement: The case of single mothers. Journal of Labor Economics 29(3):pp. 459–512

[35]See literature review in Haan et al (2014)

Bettinger E, Haegeland T, Rege M (2014) Home with mom: The effects of stay-at-home parents on children's long-run educational outcomes. Journal of Labor Economics 32(3):443–467

Cawley J, Liu F (2012) Maternal employment and childhood obesity: A search for mechanisms in time use data. Economics & Human Biology 10(4):352–364, DOI http://dx.doi.org/10.1016/j.ehb.2012.04.009, URL http://www.sciencedirect.com/science/article/pii/S1570677X12000652, obesity and the Family across the Life Course

After School Childcare Liaison Council (2013) Report on the national survery on after school childcare, http://www2s.biglobe.ne.jp/ Gakudou/2013kasyosuu.pdf (Accessed on July 3, 2014)

Currie J, Hotz VJ (2004) Accidents will happen?: Unintentional childhood injuries and the effects of child care regulations. Journal of Health Economics 23(1):25–59, URL http://ideas.repec.org/a/eee/jhecon/v23y2004i1p25-59.html

Datar A, Nicosia N, Shier V (2014) Maternal work and children's diet, activity, and obesity. Social Science and Medicine 107(0):196 – 204, DOI http://dx.doi.org/10.1016/j.socscimed.2013.12.022, URL http://www.sciencedirect.com/science/article/pii/S0277953613007144

Dobkin C, Ferreira F (2010) Do school entry laws affect educational attainment and labor market outcomes? Economics of Education Review 29(1):40 – 54, DOI http://dx.doi.org/10.1016/j.econedurev.2009.04.003, URL http://www.sciencedirect.com/science/article/pii/S0272775709000685

Fujiwara T, Okuyama M, Takahashi K (2010) Paternal involvement in childcare and unintentional injury of young children: a population-based cohort study in japan. International Journal of Epidemiology 39(2):588–597

Gennetian LA, Hill HD, London AS, Lopoo LM (2010) Maternal employment and the health of low-income young children. Journal of Health Economics 29(3):353–363, URL http://ideas.repec.org/a/eee/jhecon/v29y2010i3p353-363.html

Gordon R, Kaestner R, Korenman S (2007) The effects of maternal employment on child injuries and infectious disease. Demography 44(2):307–333, URL http://ideas.repec.org/a/spr/demogr/v44y2007i2p307-333.html

Greve J (2011) New results on the effect of maternal work hours on children's overweight status: Does the quality of child care matter? Labour Economics, Elsevier 18(5):579–590, URL http://ideas.repec.org/a/eee/labeco/v18y2011i5p579-590.html

Grossman M (1972) On the Concept of Health Capital and the Demand for Health. Journal of Political Economy 80(2):223–55

Gwozdz W, Sousa-Poza A, Reisch LA, Ahrens W, Eiben G, Fernand-Alvira JM, Hadjigeorgiou C, Henauw SD, Kovcs E, Lauria F, Veidebaum T, Williams G, Bammann K (2013) Maternal employment and childhood obesity: A european perspective. Journal of Health Economics 32(4):728 – 742, DOI http://dx.doi.org/10.1016/j.jhealeco.2013.04.003, URL http://www.sciencedirect.com/science/article/pii/S0167629613000520

Haan MD, Plug E, Rosero J (2014) Birth Order and Human Capital Development: Evidence from Ecuador. Journal of Human Resources 49(2):359–392

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

Ministry of Health L, Welfare (2010) Actual situation of working women [in japanese: Hataraku jyosei no jitujyou], URL http://www.mhlw.go.jp/bunya/koyoukintou/josei-jitsujo/10.html, accessed on January 30, 2015

von Hinke Kessler Scholder S (2008) Maternal employment and overweight children: does timing matter? Health Economics 17(8):889–906, URL http://ideas.repec.org/a/wly/hlthec/v17y2008i8p889-906.html

Ikegami N, Yoo BK, Hashimoto H, Matsumoto M, Ogata H, Babazono A, Watanabe R, Shibuya K, Yang BM, Reich MR, Kobayashi Y (2011) Japanese universal health coverage: evolution, achievements, and challenges. The Lancet 378(9796):1106 – 1115

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

Imbens GW, Angrist JD (1994) Identification and estimation of local average treatment effects. Econometrica 62(2):467–75

Prime Minister of Japan and His Cabinet (2014) Rivision of the japan rivival strategy: 10 challenges for reform, http://www.kantei.go.jp/jp/singi/keizaisaisei/pdf/10challenge02shousaiJP.pdf (Accessed on July 3, 2014)

JILPT(2013) (2014) Momen Between Childcare and Work: Reanalysis of JILPT the National Survey for the Household with Children 2011. The Japan Institute for Labour Policy and Training

Kawabata M (2014) Childcare access and employment: The case of women wih preschool-aged children intokyo. Review of Urban and Regional Development Studies 26(1):40–56, DOI 10.1111/rurd.12018, URL http://dx.doi.org/10.1111/rurd.12018

Kawaguchi D (2011) Actual age at school entry, educational outcomes, and earnings. Journal of the Japanese and International Economies 25(2):64–80

Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. Journal of Economic Literature 48(2):281–355, DOI 10.1257/jel.48.2.281, URL http://www.aeaweb.org/articles.php?doi=10.1257/jel.48.2.281

Lee GH, Lee SP (2014) Childcare availability, fertility and female labor force participation in japan. Journal of the Japanese and International Economies 32(C):71–85, URL http://ideas.repec.org/a/eee/jjieco/v32y2014icp71-85.html

McCrary J (2008) Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142(2):698–714

McCrary J, Royer H (2011) The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. American Economic Review 101(1):158–95

Morrill MS (2011) The effects of maternal employment on the health of school-age children. Journal of Health Economics 30(2):240–257

OECD (2011) Oecd family database, URL http://www.oecd.org/els/family/42063425.pdf, accessed on July 4, 2014

Cabinet Office (2013) White paper on gender equality 2013, URL http://www.gender.go.jp/about_danjo/whitepaper/h25/zentai/html/zuhyo/zuhyo01-02-01.html, accessed on July 23, 2014

Ruhm CJ (2000) Are recessions good for your health? The Quarterly Journal of Economics 115(2):617–650

Ruhm CJ (2004) Parental employment and child cognitive development. Journal of Human Resources 39(1), URL http://ideas.repec.org/a/uwp/jhriss/v39y2004i1p155-192.html

Ruhm CJ (2014) Maternal employment and child development

Shigeoka H (2014a) The effect of patient cost sharing on utilization, health, and risk protection. American Economic Review 104(7):2152–84

Shigeoka H (2014b) School entry cutoff date and the timing of births. mimeo

Silverstein M, Sales AE, Koepsell TD (2003) Health care utilization and expenditures associated with child care attendance: A nationally representative sample. Pediatrics 111(4):e371–e375

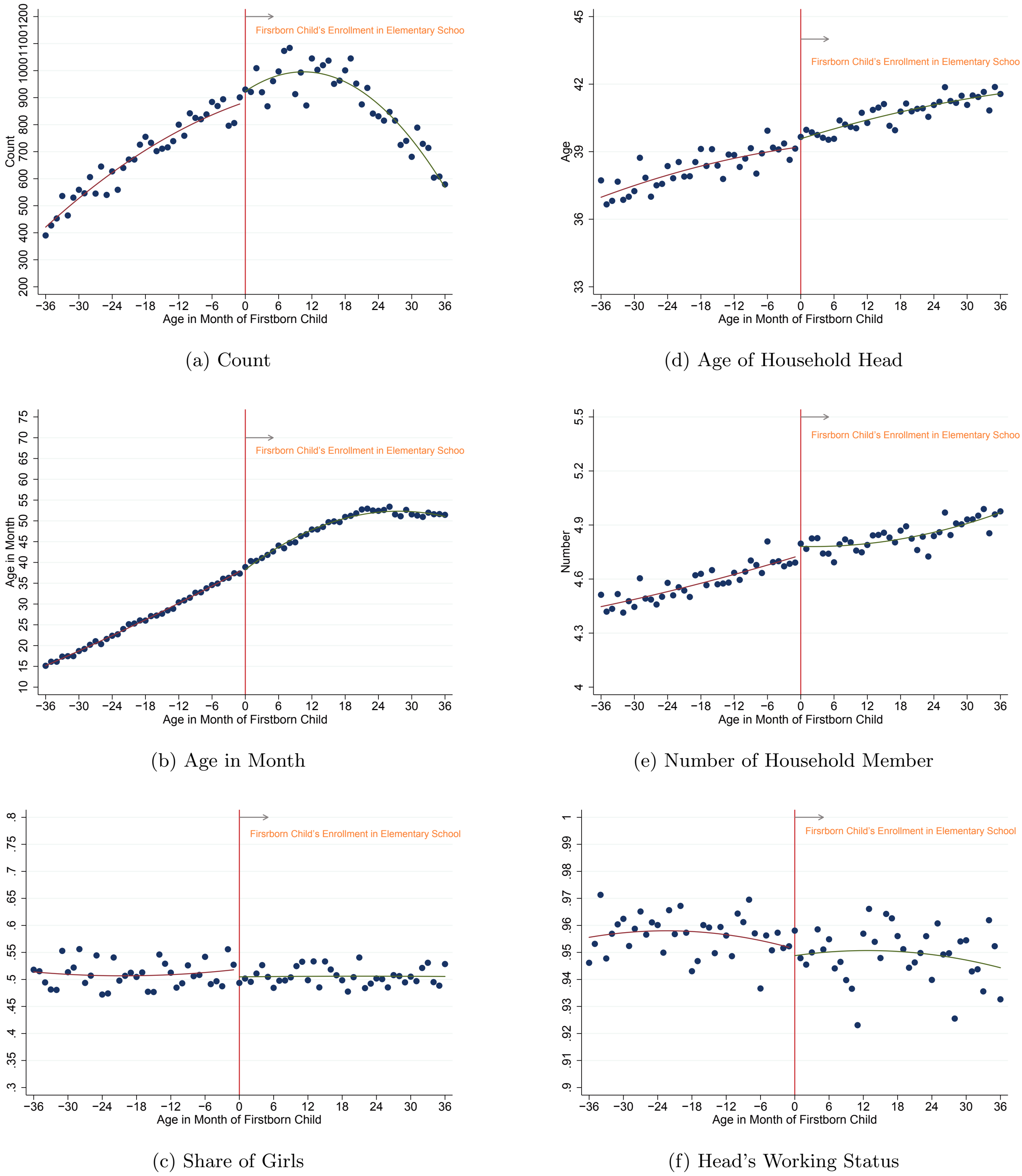
Tanaka R (2008) The gender-asymmetric effect of working mothers on children's education: Evidence from japan. Journal of the Japanese and International Economies 22(4):586–604

Unayama T (2012) The Possibility of Pursuing Both Marriage/Childbirth and Employment, and the Development of Nursery Schools. Japanese Economy, ME Sharpe, Inc 39(1):48–71, URL http://ideas.repec.org/a/mes/jpneco/v39y2012i1p48-71.html

Waldfogel J, Han WJ, Brooks-Gunn J (2002) The effects of early maternal employment on child cognitive development. Demography 39(2):369–392, URL http://ideas.repec.org/a/spr/demogr/v39y2002i2p369-392.html

Wüst M (2014) Maternal employment during pregnancy and birth outcomes: Evidence from danish siblings. Health Economics Forthcoming

Zhou Y, Oishi A (2005) Underlying demand for licensed childcare services in urban japan. Asian Economic Journal 19(1):103–119, DOI 10.1111/j.1467-8381.2005.00206.x, URL http://dx.doi.org/10.1111/j.1467-8381.2005.00206.x



(a) Count

(d) Age of Household Head

(b) Age in Month

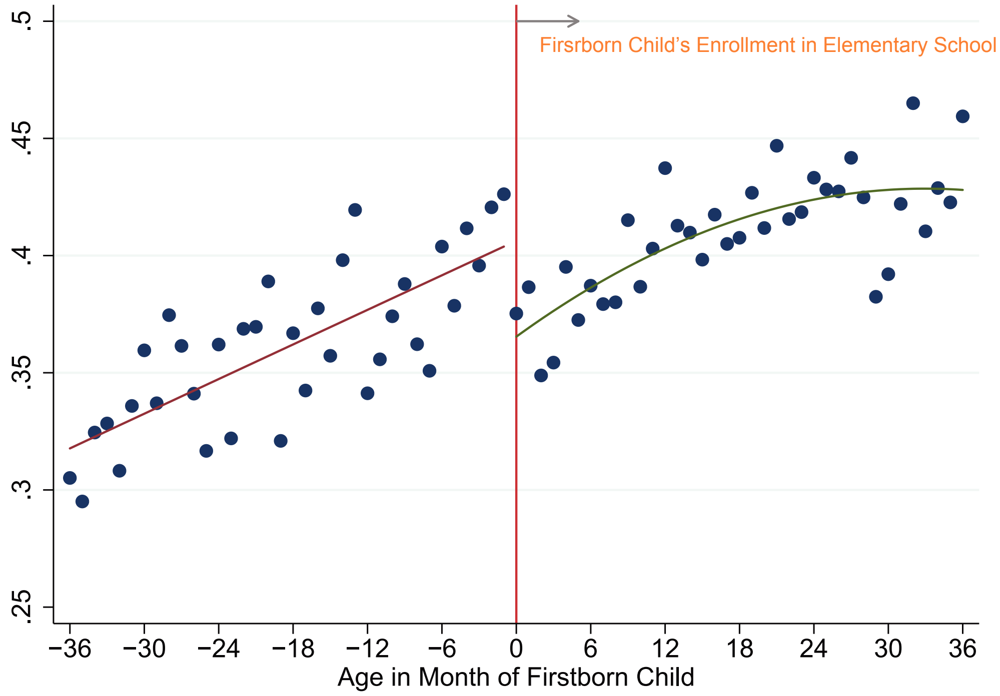
(e) Number of Household Member

(c) Share of Girls

(f) Head's Working Status

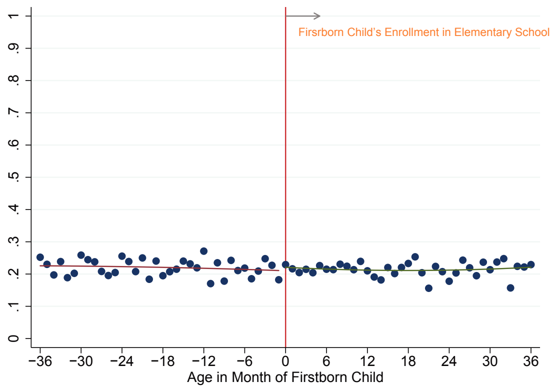
Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. Count is the number of observation in each bin. The sample includes preschool children except the firstborns. The lines are the quadratic fit.

Figure 1: Identification Checks

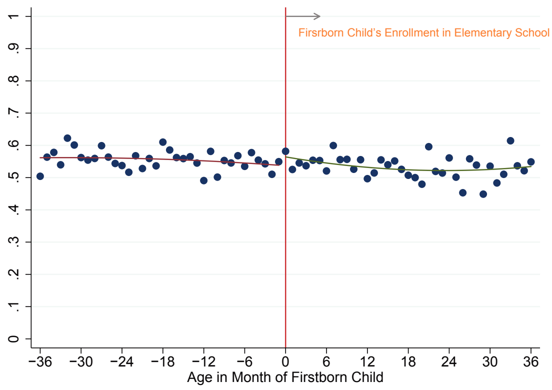


Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. The lines are the quadratic fit.

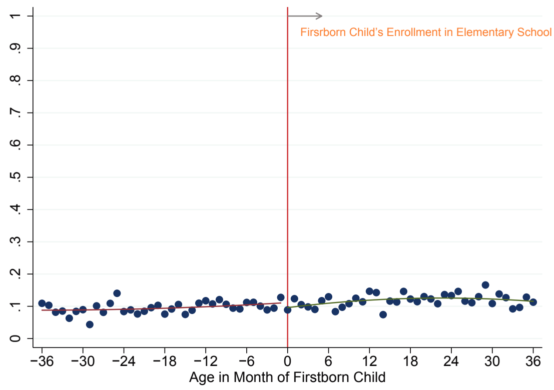
Figure 2: Share of Working Mothers



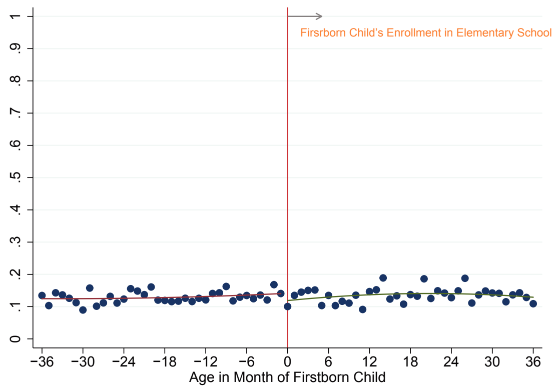
(a) Self-employed



(b) General Employee



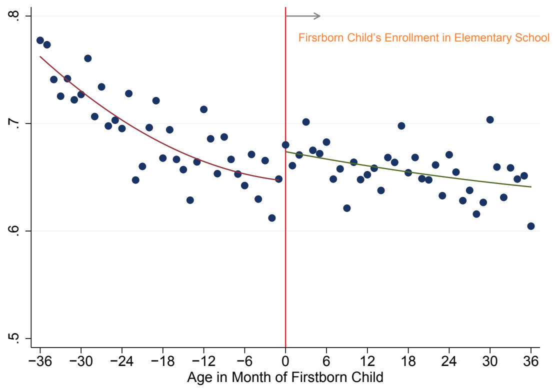
(c) Employee with Short Contract



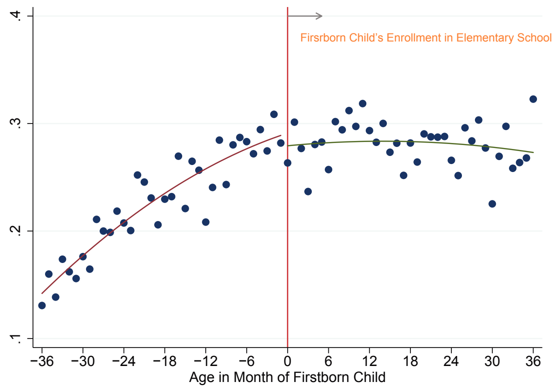
(d) Others

Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. Count is the number of observation in each bin. The sample includes preschool children except firstborns. The children whose mother is not working are excluded. "Self-employed" includes the wives whose husband is self-employed and works for family business. "Employee with short contract" includes the workers whose employment contract is less than 1 years and "general employee" is the employed whose employment contract lasts for 1 year or the employed without term of the contract. "Others" includes other workers. The lines are the quadratic fit.

Figure 3: Types of Mother's Employment Contract



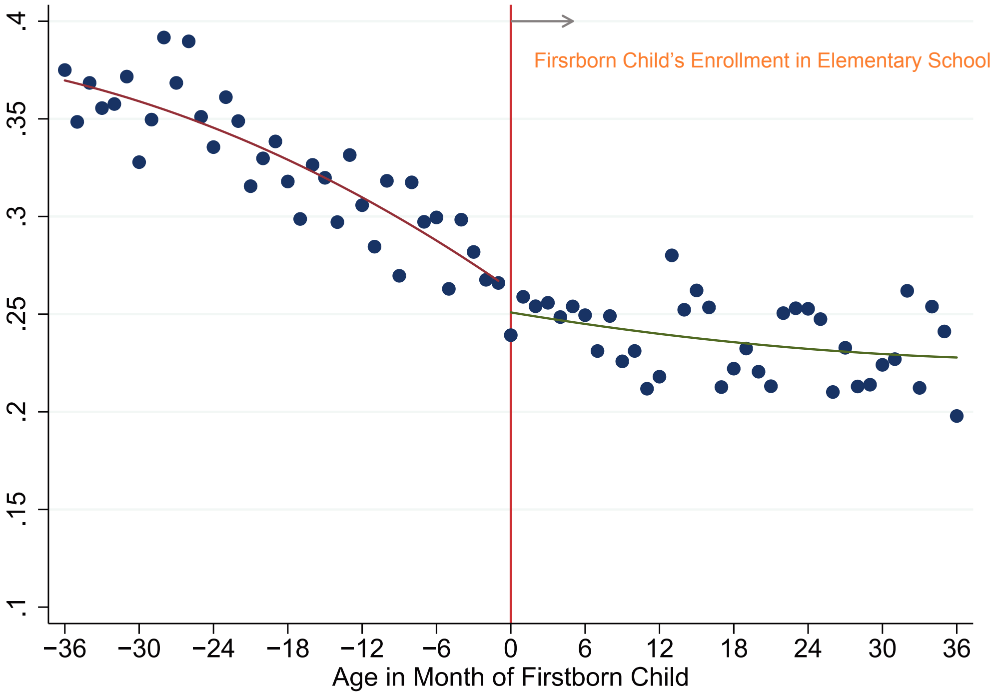
(a) Parental Care



(b) Daycare Center

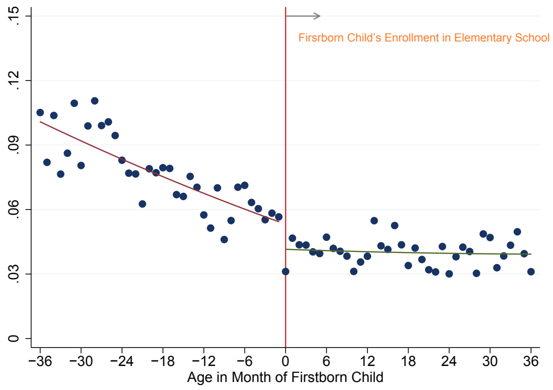
Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. Children defined as receiving parental care if they enroll in kindergarten or the parents take care of them in day time. If children are enrolled in daycare center and usual care providers for them are not their parents, they are regarded as receiving care only from daycare center. Enrollment for kindergarten is prohibited for children below 36 months. The lines are the quadratic fit.

Figure 4: Childcare Providers in the Day Time

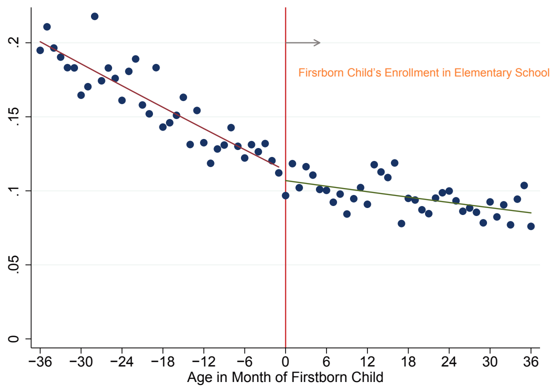


Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. The lines are the quadratic fit.

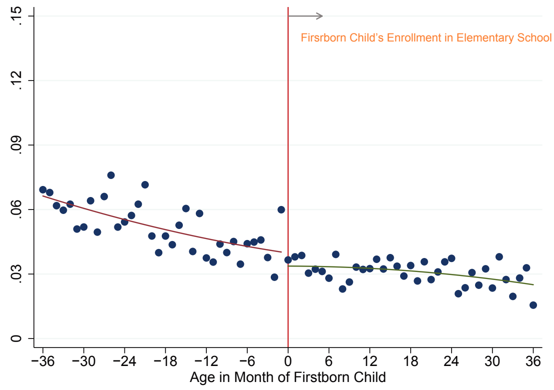
Figure 5: Probability of Having Any Symptoms



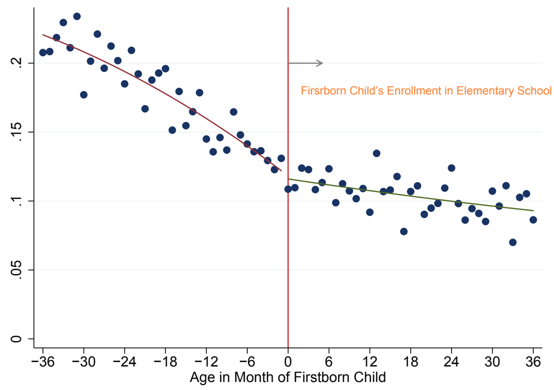
(a) Fever



(b) Cough



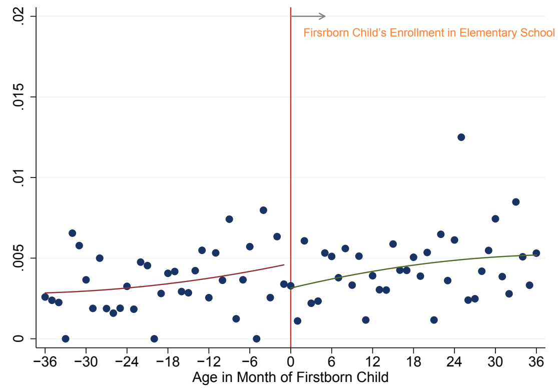
(c) Wheezing



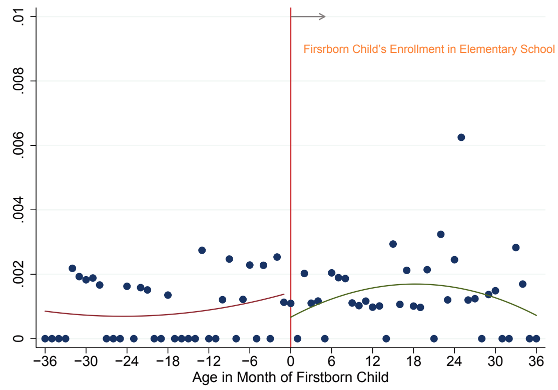
(d) Stuff Nose

Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. The lines are the quadratic fit.

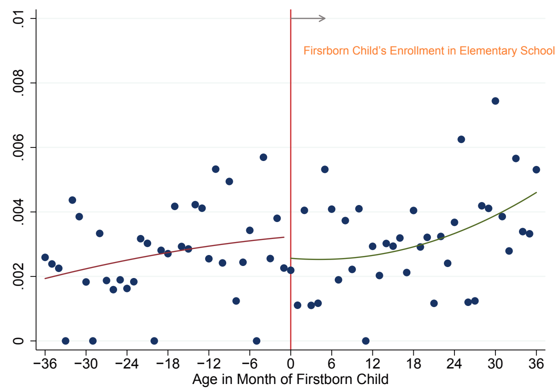
Figure 6: Probability of Having Any Symptoms : Selected Items



(a) All Injuries



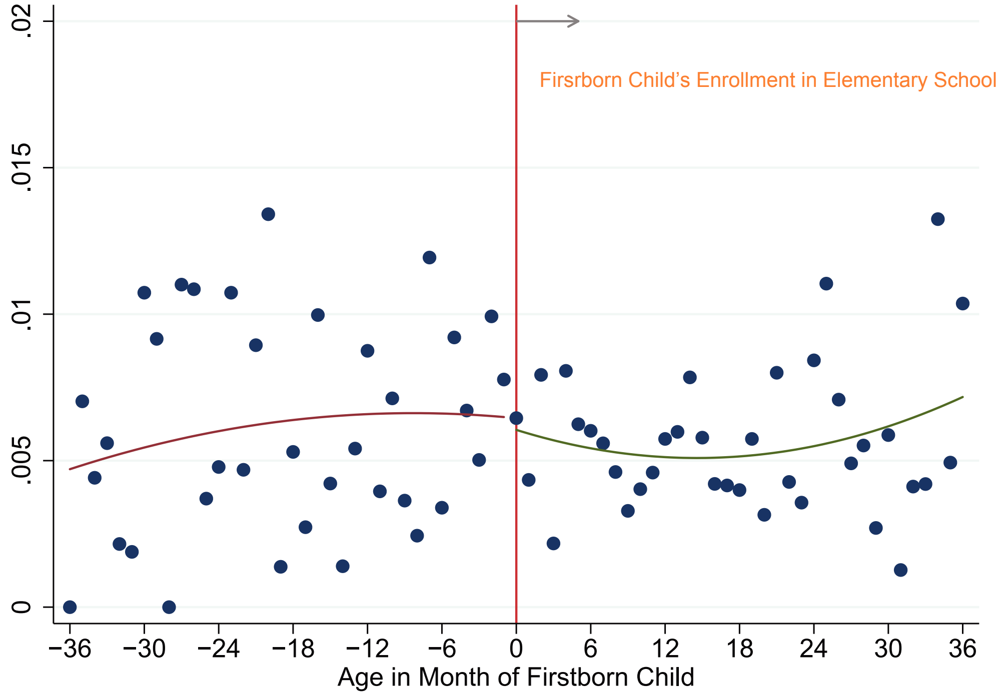
(b) Fractures



(c) Other Injuries and Skin Burns

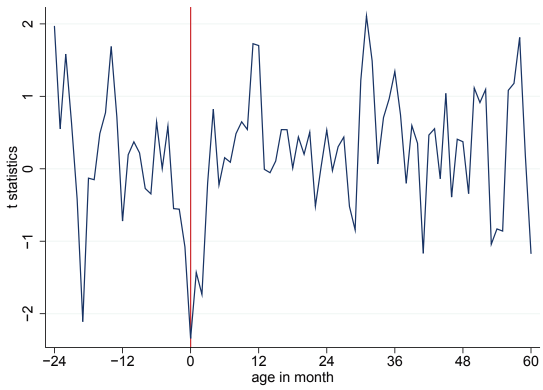
Note: "All injuries" include the outpatient utilization for fractures, and the other injuries and skin burn. Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. The lines are the quadratic fit.

Figure 7: Probability of Current Outpatient Visits due to Injury

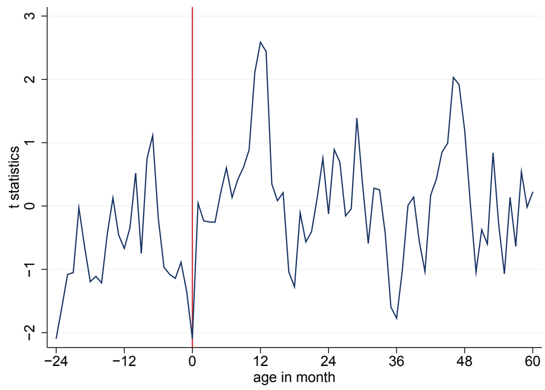


Note: Horizontal axis represents the age in month of the eldest child standardized by the month when they enroll in elementary school. The sample includes preschool children except firstborns. The lines are the quadratic fit.

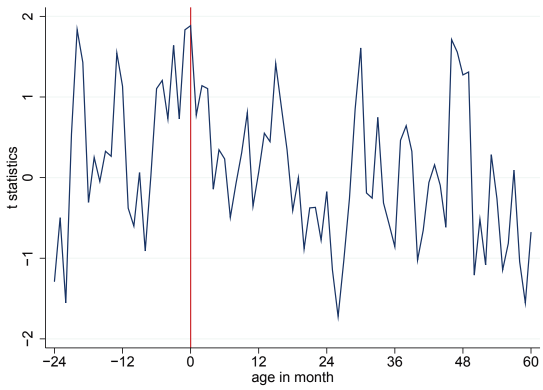
Figure 8: Probability of Hospitalization



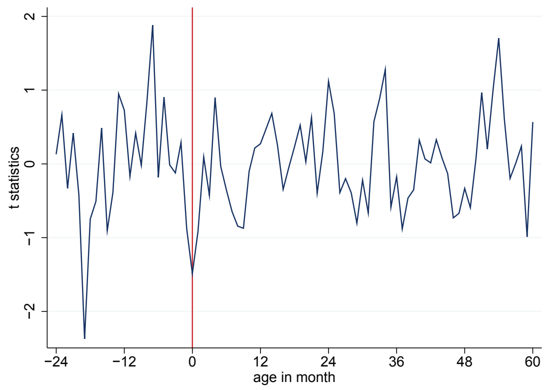
(a) Maternal Employment



(c) Fever



(b) Parental Care



(d) Hospitalization

Note: Horizontal axis represents the timing of the placebo treatment which is measured by the standardized age in month of the eldest child. Vertical axis represents the average value of t statistics. These t statistics are averages of different control function specifications; specifically, a linear and cubic control function estimated on each side of the cutoff, with the bandwidth ranged from 12 months to 36 months. For each placebo treatment, 50 t statistics (2 polynomials * 25 bandwidths) are derived. All specifications include controls for the fixed effects of the eldest child's birth quarter as well as the other covariates summarized in Appendix Table A1. Vertical line represents the age of 75 months, standardized at zero, over which the eldest child enrolls in elementary school.

Figure 9: Placebo Test

Table 1: Means of the Dependent Variables by Samples

	All Preschool Children			RD Sample		
	Obs.	Mean	S.D	Obs.	Mean	S.D
Panel A. Health Outcomes						
Probability of Symptom	145,306	0.271	0.444	55,788	0.273	0.446
Probability of Visit due to Injuries	146,178	0.004	0.065	56,238	0.004	0.064
Fracture	146,178	0.001	0.033	56,238	0.001	0.034
Other Injuries	146,178	0.003	0.056	56,238	0.003	0.054
Probability of Hospitalization	148,699	0.006	0.076	57,211	0.006	0.076
Panel B .Maternal Employment						
Mother Employed	148,699	0.383	0.486	57,211	0.389	0.488
Self-Employed, etc.	56,988	0.207	0.405	22,283	0.216	0.412
General Employees	56,988	0.560	0.496	22,283	0.541	0.498
Employee with Short-term Contract	56,988	0.107	0.309	22,283	0.110	0.312
Other Employee	56,988	0.126	0.332	22,283	0.133	0.340
Panel C .Childcare Provision						
Parental Care	111,420	0.67	0.466	42,216	0.66	0.50
Daycare Center	111,420	0.24	0.4357	42,216	0.26	0.43

Note: "All children" includes all preschool children aged 6 months and over. "RD sample" summarizes the means in the sample which is used for RD analysis with bandwidth of 36 months.

Table 2: Results From OLS

		All Children	Exclude Eldest Children			
		No	No	36 months	24 months	12 months
		(1)	(2)	(3)	(4)	(5)
Symptom	Any	0.015***	0.010**	0.010*	0.010	0.016**
		(0.003)	(0.004)	(0.005)	(0.006)	(0.007)
	Fever	0.008***	0.007***	0.008***	0.009***	0.009***
		(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
	Cough	0.004*	0.000	-0.003	-0.003	0.001
		(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
	Wheezing	0.009***	0.007***	0.008***	0.007***	0.006***
		(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
	Stuff Nose	0.009***	0.003	0.002	-0.001	-0.000
		(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
	Obs.	145,310	78,070	55,788	41,264	22,053
Injury	Any	0.000	-0.000	0.000	0.000	-0.000
		(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
	Fracture	0.000	0.000	0.000	0.000	0.001
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
	Other Injuries	-0.000	-0.000	-0.000	-0.000	-0.001
		(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
	Obs.	146,182	78,696	56,238	41,601	22,233
Hospitalization		0.001	-0.000	-0.000	0.000	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	Obs.	148,703	80,033	57,211	42,324	22,619
Covariates		X	X	X	X	X
Prefecture Fixed Effects		X	X	X	X	X
Year Effects		X	X	X	X	X

Note: The table reports the coefficients of maternal employment on main outcomes, using OLS. Standard errors are clustered at prefecture level. Outcome variables are present in the left column. In Column (1), the result form all preschool children aged 6 months and over are included. From Column (2) to Column (3), the eldest child is excluded. To compare the results with those from RD analysis, the sample is restricted in several intervalsbefore and after the threshold month when the eldest child enrolls in elementary school. The bandwidth is reported in the upper raw of the table. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 3: Results From Conventional IV

		All Children	Exclude Eldest Children			
Bandwidth		No	No	36 months	24 months	12 months
		(1)	(2)	(3)	(4)	(5)
Symptom	Any	0.977**	0.865*	0.748	0.854	1.003
		(0.471)	(0.520)	(1.122)	(0.690)	(0.694)
	Fever	0.521**	0.717*	1.318	0.756	0.722
		(0.253)	(0.402)	(1.399)	(0.511)	(0.490)
	Cough	0.721**	0.619	0.761	0.638	0.855
		(0.321)	(0.410)	(0.982)	(0.524)	(0.698)
	Wheezing	0.382**	0.339	0.496	0.212	0.199
		(0.188)	(0.227)	(0.537)	(0.237)	(0.275)
	Stuff Nose	0.587*	0.646	0.788	0.889	0.877
		(0.340)	(0.472)	(1.109)	(0.708)	(0.631)
	First Stage F stat	15.04	5.58	1.38	3.63	4.25
	Obs.	145,310	78,070	55,788	41,264	22,053
Injury	Any	-0.002	-0.030	-0.065	-0.051	-0.044
		(0.032)	(0.051)	(0.132)	(0.064)	(0.061)
	Fracture	-0.003	0.003	-0.053	-0.035	-0.044
		(0.019)	(0.025)	(0.072)	(0.036)	(0.033)
	Other Injuries	-0.001	-0.032	-0.013	-0.012	0.005
		(0.028)	(0.052)	(0.105)	(0.048)	(0.053)
	First Stage F stat	17.88	6.36	2.11	4.53	4.67
	Obs.	146,182	78,696	56,238	41,601	22,233
Hospitalization	Any	0.038	-0.002	-0.045	0.074	-0.030
		(0.055)	(0.069)	(0.151)	(0.104)	(0.083)
	First Stage F stat	14.31	4.32	5.71	6.22	5.39
	Obs.	148,703	80,033	57,211	42,324	22,619
Covariates		X	X	X	X	X
Prefecture Fixed Effects		X	X	X	X	X
Year Effects		X	X	X	X	X

Note: The table reports the coefficients of maternal employment on main outcomes, using region-level female unemployment variable as an instrumental variable for maternal employment. Standard errors are clustered at prefecture level. Outcome variables are present in the left column. In Column (1), the result form all preschool children aged 6 months and over are included. From Column (2) to Column (3), the eldest child is excluded. To compare the results with those from RD analysis, the sample is restricted in several intervals before and after the threshold month when the eldest child enrolls in elementary school. The bandwidth is reported in the upper raw of the table. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 4: Identification Checks

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Panel A. Parametric Density Test					
Ln Count	-0.037	-0.007	-0.232	0.022	-0.037
	(0.024)	(0.029)	(0.146)	(0.039)	(0.065)
Panel B. Discontinuity in Covariates					
Age in Month	0.231	0.249	1.013	0.525	0.366
	(0.229)	(0.289)	(1.395)	(0.411)	(0.708)
Share of Girl	0.002	-0.006	-0.002	-0.010	-0.020
	(0.008)	(0.011)	(0.052)	(0.018)	(0.033)
Age of Heads	0.150	0.249	0.702	0.223	0.752
	(0.212)	(0.278)	(1.291)	(0.371)	(0.614)
Age of Spouses	0.135	0.205	0.392	-0.082	0.672
	(0.202)	(0.263)	(1.194)	(0.346)	(0.589)
Number of Children	0.004	-0.009	-0.037	-0.013	0.010
	(0.011)	(0.015)	(0.076)	(0.020)	(0.033)
Number of Household Members	0.024	0.014	0.122	0.010	0.105
	(0.020)	(0.026)	(0.123)	(0.036)	(0.064)
Head's Working Status	-0.008*	-0.011**	0.014	-0.011	-0.019
	(0.004)	(0.005)	(0.026)	(0.007)	(0.013)
Municipality-based Insurance	-0.004	0.001	0.002	-0.016	-0.020
	(0.008)	(0.011)	(0.043)	(0.016)	(0.028)
Employment-based Insurance	-0.000	-0.007	-0.015	0.010	0.011
	(0.009)	(0.011)	(0.047)	(0.017)	(0.029)
Number of Observations	432	288	144	432	432
Polynomial Order	One	One	One	Two	Three
Year Effect	X	X	X	X	X
Birth Quarter Dummies	X	X	X	X	X

Note: Panel A summarizes the results of density test. Panel B examines the discontinuity of the covariates at the cut-off month. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 5: Effect on Maternal Employment

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Without Covariates	-0.025***	-0.031***	-0.081***	-0.041***	-0.064***
	(0.009)	(0.011)	(0.024)	(0.013)	(0.019)
With Covariates	-0.032***	-0.033***	-0.078***	-0.038***	-0.062***
	(0.009)	(0.011)	(0.025)	(0.013)	(0.019)
Number of Observations	57,211	42,324	22,619	57,211	57,211
Polynomial Order	One	One	One	Two	Three
Year Effect	X	X	X	X	X
Birth Quarter Fixed Effect	X	X	X	X	X

Note: This table summarizes the RD estimates based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 6: Effect on the Type of Employment Contract

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Self-Employed , etc.	0.010	0.025*	0.019	0.018	0.006
	(0.011)	(0.013)	(0.025)	(0.016)	(0.023)
General Employee	0.011	0.014	0.041	0.037*	0.039
	(0.014)	(0.018)	(0.034)	(0.021)	(0.030)
Employee with Short-term Contract	-0.007	-0.017	-0.024	-0.021	-0.011
	(0.009)	(0.011)	(0.025)	(0.013)	(0.019)
Other Employee	-0.014	-0.022*	-0.035	-0.033**	-0.034
	(0.010)	(0.012)	(0.023)	(0.015)	(0.021)
Observations	22,283	16,518	8,721	22,283	22,283
Covariates	X	X	X	X	X
Year Effects	X	X	X	X	X
Birth Quarter Fixed Effect	X	X	X	X	X
Polynomial Order	One	One	One	Two	Three

Note: This table summarizes the RD estimates based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 7: Effect on the Choice of Childcare Provider

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Panel A. Without Covariates					
Any Parental Care	0.037***	0.017	0.054**	0.021	0.040*
	(0.010)	(0.012)	(0.025)	(0.014)	(0.021)
Daycare Center Only	-0.016*	-0.007	-0.042*	-0.009	-0.032
	(0.009)	(0.012)	(0.024)	(0.014)	(0.020)
Panel B. With Covariates					
Any Parental Care	0.046***	0.022*	0.055**	0.021	0.040*
	(0.010)	(0.012)	(0.024)	(0.014)	(0.021)
Daycare Center Only	-0.026***	-0.011	-0.042*	-0.006	-0.032
	(0.009)	(0.012)	(0.023)	(0.014)	(0.020)
Number of Observations	42,656	31,560	16,852	31,560	16,852
Polynomial Order	One	One	One	Two	Three
Year Effect	X	X	X	X	X
Birth Quarter Fixed Effect	X	X	X	X	X

Note: This table summarizes the RD estimates based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 8: Effect on the Probability of Having Any Subjective Symptoms

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Reduced Form Estimates	-0.016**	-0.026***	-0.016	-0.015	-0.024
	(0.008)	(0.010)	(0.022)	(0.012)	(0.018)
IV Estimates	0.542*	0.808***	0.225	0.407	0.382
	(0.283)	(0.311)	(0.141)	(0.265)	(0.288)
First Stage F stat	17.84	13.59	14.32	12.69	12.94
Number of Observations	55,788	41,264	22,053	55,788	55,788
Polynomial Order	One	One	One	Two	Three
Covariates	X	X	X	X	X
Year Effect	X	X	X	X	X
Birth Quarter Fixed Effect	X	X	X	X	X

Note: This table summarizes the reduced form RD estimates and IV estimates, based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 9: Effect on Individual Symptoms: Reduced Form Estimates

		Fever	Cough	Headache	Wheezing	Toothache	Stuff nose	Diarrhea	Stomachache	Rash	Cut
Polynomial Order	Bandwidth Choice	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
One	36 months	-0.006	-0.003	-0.000	-0.000	0.001	-0.001	0.003	-0.002	-0.002	0.000
		(0.004)	(0.006)	(0.001)	(0.003)	(0.001)	(0.006)	(0.002)	(0.003)	(0.003)	(0.002)
Observations		57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211
One	24 months	-0.011**	-0.008	-0.000	-0.004	0.000	-0.004	0.002	-0.002	-0.002	0.002
		(0.005)	(0.007)	(0.001)	(0.004)	(0.002)	(0.007)	(0.003)	(0.004)	(0.004)	(0.002)
Observations		42,324	42,324	42,324	42,324	42,324	42,324	42,324	42,324	42,324	42,324
One	12 months	-0.022**	-0.014	0.002	-0.016**	0.002	-0.003	-0.003	0.006	0.006	0.003
		(0.009)	(0.014)	(0.002)	(0.008)	(0.004)	(0.013)	(0.005)	(0.007)	(0.007)	(0.004)
Observations		22,619	22,619	22,619	22,619	22,619	22,619	22,619	22,619	22,619	22,619
Two	36 months	-0.012**	-0.009	-0.000	-0.006	-0.000	-0.003	0.003	-0.001	-0.001	0.003
		(0.006)	(0.008)	(0.001)	(0.005)	(0.002)	(0.009)	(0.003)	(0.005)	(0.005)	(0.003)
Observations		57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211
Three	36 months	-0.022***	-0.018	0.001	-0.009	0.002	-0.012	0.002	0.000	0.000	0.003
		(0.008)	(0.012)	(0.002)	(0.007)	(0.002)	(0.013)	(0.005)	(0.007)	(0.007)	(0.004)
Observations		57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211	57,211
Mean of Dep.		0.036	0.096	0.004	0.029	0.010	0.117	0.012	0.007	0.026	0.013

Note: This table summarizes the reduced-form estimates based on alternative specifications. Each column corresponds to the symptoms and means of the variable are reported in the bottom row. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$..

Table 10: Effect on the Probability of Current Outpatient Visits due to Injury

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Panel A. Reduced Form Estimates					
All Injuries	-0.001	-0.001	-0.001	-0.002	-0.002
	(0.001)	(0.001)	(0.003)	(0.002)	(0.002)
Fracture	-0.000	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Other Injuries	-0.001	-0.001	-0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Panel B. IV Estimates					
All Injuries	0.043	0.042	0.019	0.045	0.030
	(0.032)	(0.036)	(0.024)	(0.040)	(0.030)
Fracture	0.001	0.012	0.005	0.022	0.016
	(0.014)	(0.014)	(0.007)	(0.018)	(0.016)
Other Injuries	0.041	0.025	0.011	0.018	0.007
	(0.025)	(0.028)	(0.019)	(0.027)	(0.019)
Number of Observations	56,238	41,601	22,233	56,238	41,601
Polynomial Order	One	One	One	Two	Three

Note: "Visits due to injury" includes the outpatient utilization for fractures, the other injuries and skin burn. This table summarizes the reduced form RD estimates and IV estimates, based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 36 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$

Table 11: Effect on the Probability of Current Hospital Admission

	36 months (1)	24 months (2)	12 months (3)	36 months (4)	36 months (5)
Reduced Form Estimates	-0.002	-0.002	-0.016*	-0.002	-0.006**
	(0.001)	(0.002)	(0.008)	(0.002)	(0.003)
IV Estimates	0.060	0.065	0.063*	0.043	0.108**
	(0.043)	(0.049)	(0.033)	(0.048)	(0.046)
Number of Observations	57,211	42,324	22,619	57,211	57,211
Polynomial Order	One	One	One	Two	Three
Covariates	X	X	X	X	X
Year Effect	X	X	X	X	X
Birth Quarter Fixed Effect	X	X	X	X	X

Note: This table summarizes the reduced form RD estimates and IV estimates, based on alternative specifications. Column (1), (2) and (3) show the results of RD regression with a linear polynomial, based on alternative bandwidth selection from 24 months to 6 months. Column (4) and (5) show the results of RD regression with alternative polynomial orders, while bandwidth is fixed at 36 months. All equations control sex, age in month, age of household head, number of children under 15 years old, number of household member, working status of household head, insurance plans, survey year effects and prefecture fixed effects. Standard error is clustered at the age in month of the firstborn child. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$

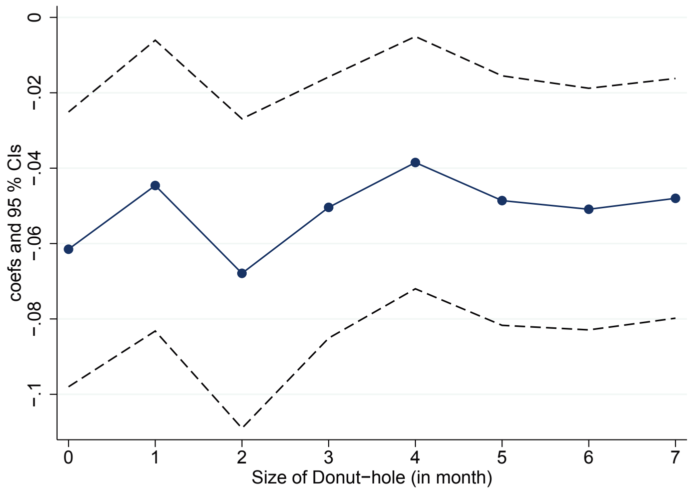
A Descriptive Statistics

Table A1: Means by Samples

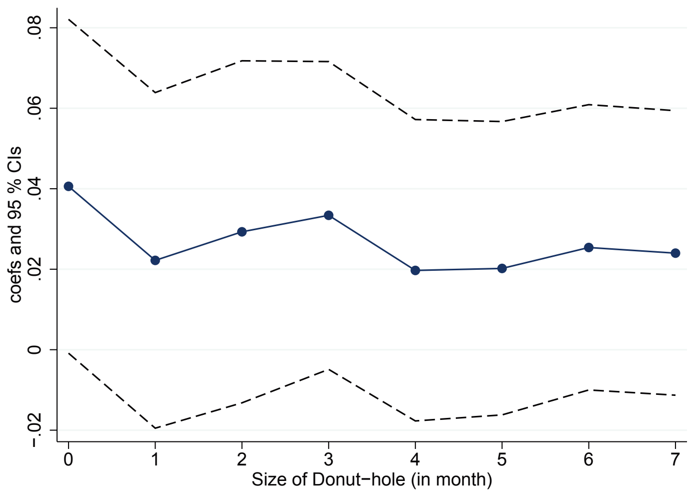
	All Preschool Child	RD Sample
	(1)	(2)
Mother Employed	0.38	0.39
	(0.49)	(0.49)
Age in Month	41.40	39.19
	(19.88)	(18.74)
Female	0.51	0.51
	(0.50)	(0.50)
Firstborn	0.46	n.a.
	(0.50)	n.a.
N. of Children	2.05	2.36
	(0.80)	(0.54)
N. of Household Members	4.40	4.73
	(1.18)	(1.03)
Age of Household Head	38.99	39.65
	(10.97)	(10.88)
Age of Spouse	36.63	37.24
	(10.45)	(10.42)
Head's Working Status	0.95	0.95
	(0.21)	(0.21)
Insurance Status: CHI	0.22	0.23
	(0.42)	(0.42)
Insurance Status: NHIA	0.04	0.04
	(0.19)	(0.19)
Insurance Status: SMHI	0.72	0.72
	(0.45)	(0.45)
Insurance Status: Other	0.02	0.02
	(0.13)	(0.13)
Observations	148,699	56,632

Note: Column 1 includes all preschool children aged 6 months and over. Column 2 summarizes the means in the sample which is used for RD analysis with bandwidth of 36 months. Firstborn children are excluded in Column (2).

B Donut-hole RD



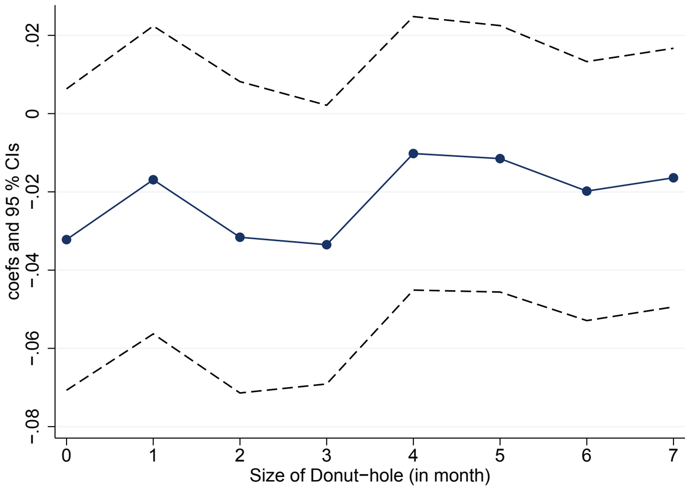
(a) Maternal Employment



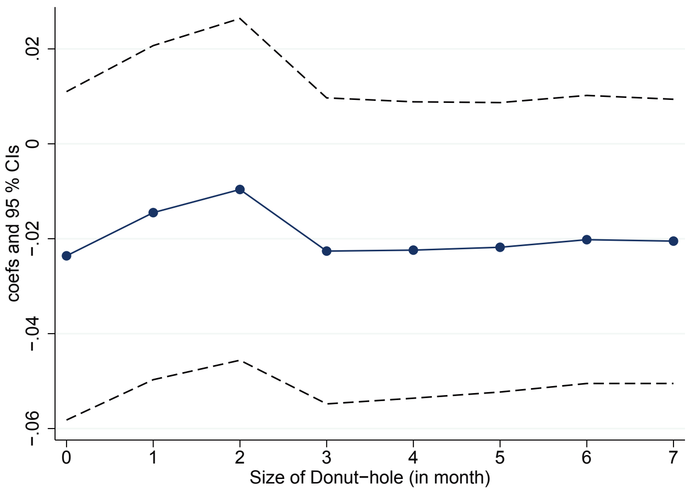
(b) Parental Care

Note: Horizontal axis represents the size of donut-hole which is the number of month excluded from RD estimation. "0" represents a baseline specification where no observations are excluded. The model here is based on cubic age profile fully interacted with a dummy for school-entry age or order. In all estimations, the bandwidth is fixed at 36 months. Dashed lines are 95 percent confidence interval.

Figure B1: Donut-hole RD Estimates



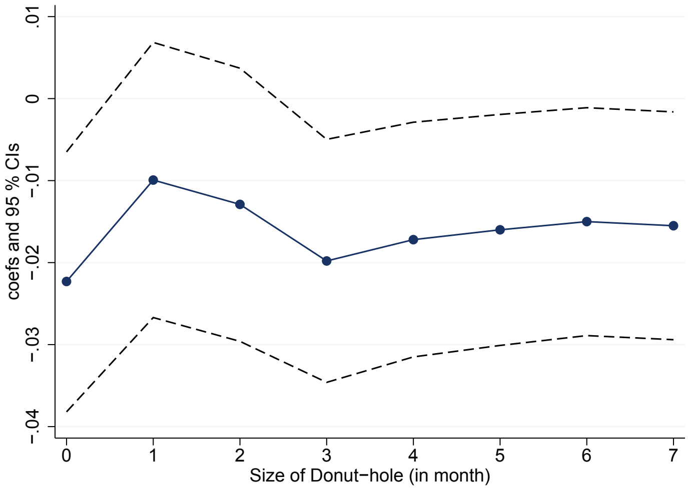
(c) Daycare Center Only



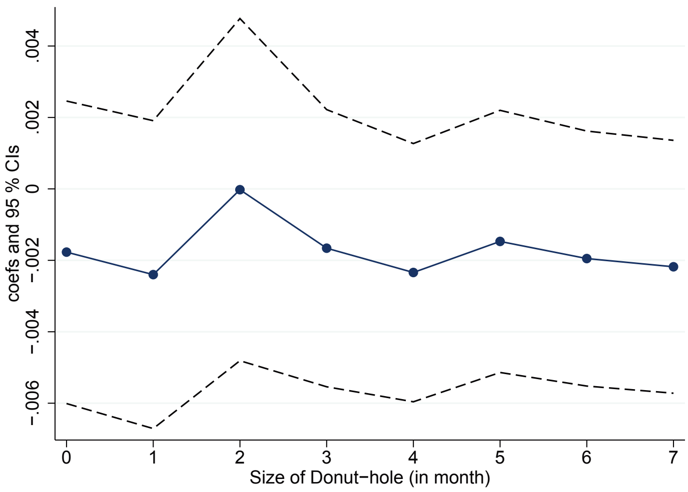
(d) Any Symptom

Note: Horizontal axis represents the size of donut-hole which is the number of month excluded from RD estimation. "0" represents a baseline specification where no observations are excluded. The model here is based on cubic age profile fully interacted with a dummy for school-entry age or order. In all estimations, the bandwidth is fixed at 36 months. Dashed lines are 95 percent confidence interval.

Figure B1: Donut-hole RD Estimates



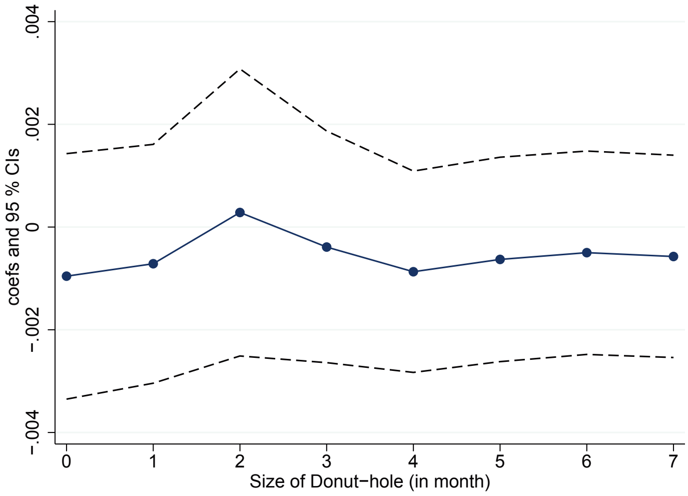
(e) Fever



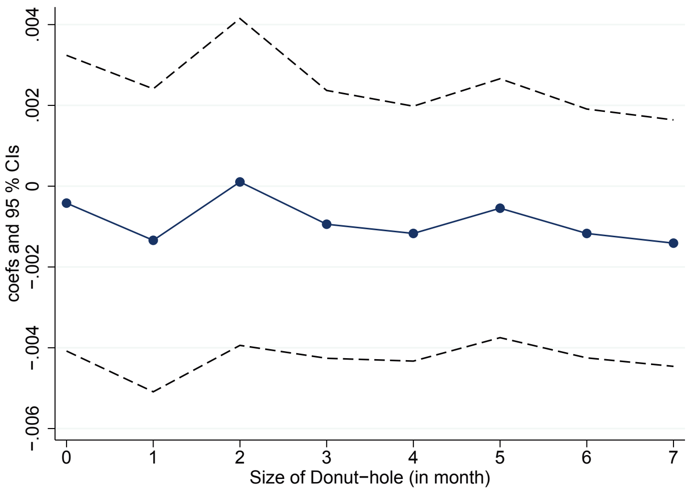
(f) All Cause Injuries

Note: Horizontal axis represents the size of donut-hole which is the number of month excluded from RD estimation. "0" represents a baseline specification where no observations are excluded. The model here is based on cubic age profile fully interacted with a dummy for school-entry age or order. In all estimations, the bandwidth is fixed at 36 months. Dashed lines are 95 percent confidence interval.

Figure B1: Donut-hole RD Estimates



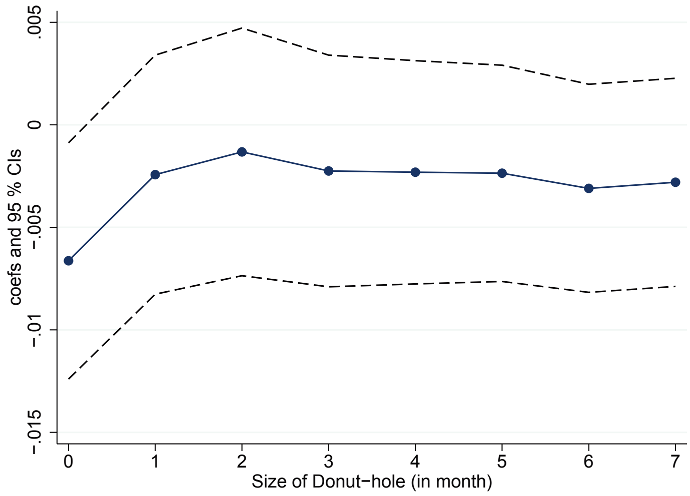
(g) Fractures



(h) Other Injuries

Note: Horizontal axis represents the size of donut-hole which is the number of month excluded from RD estimation. "0" represents a baseline specification where no observations are excluded. The model here is based on cubic age profile fully interacted with a dummy for school-entry age or order. In all estimations, the bandwidth is fixed at 36 months. Dashed lines are 95 percent confidence interval.

Figure B1: Donut-hole RD Estimates



(g) Hospitalization

Note: Horizontal axis represents the size of donut-hole which is the number of month excluded from RD estimation. "0" represents a baseline specification where no observations are excluded. The model here is based on cubic age profile fully interacted with a dummy for school-entry age or order. In all estimations, the bandwidth is fixed at 36 months. Dashed lines are 95 percent confidence interval.

Figure B1: Donut-hole RD Estimates

C Additional Placebo Test

In Appendix C, I focus on two potential treatments and present whole results without taking the average value of them. First, I choose the month when the eldest child becomes 4th grade in elementary school as a timing of the placebo treatment. Specifically, a dummy variable which take a value of one if the eldest child is over 111 months. This age is 36 months (3 years) after the month of school entry. In addition, another placebo test which exploits the timing of 123 months, when the eldest child becomes 5th grader in elementary school, is also implemented.

The RD estimation is based on the reduced form regression which directly examines the association between the treatment (the timing of becoming 1st, 4th or 5th grader) and outcome variables. For each outcome, linear, quadratic and cubic polynomials and their interactions with the cut-off dummy are fitted in order to control underlying trends which are associated with eldest child's age in month. The results are presented one by one from Figure C1 to Figure C9. While the main findings are not so different from those written in the main text, I briefly summarize them below.

C.1 Maternal Employment

Figure C1 presents the results on maternal employment. In this figure, figure (a) to (c) present the results by "real" treatment which exploit the eldest child's school entry. On the other hand, the middle column shows the results by the placebo treatment which exploits the eldest child's promotion to 4th grade. And then, results from another placebo test which exploit the promotion to 5th grade are presented in right column. In each test, I show the RD estimate with alternative bandwidth changed from 12 months to 36 months. In addition, linear polynomial is controlled in the upper row, and quadratic and cubic one are controlled in the middle and bottom row, respectively.

The figure (a) to (c) clearly show robustly that the eldest child's school entry reduces maternal employment ,while some estimates from the narrow bandwidth are not significant because of larger standard error. These results exhibit a clear contrast with those from two placebo test. For instance, as in figure (b), we see that standard error becomes larger as bandwidth goes narrower in the figure (e). However, treatment effects are significant only in figure (b) which is from the "real" treatment. This suggests that, while RD estimates are not significant with narrow bandwidth in figure (b), it does not suggests there is no treatment effect. All in all, two placebo tests show no significant effect under various assumption of underlying trends and RD estimates are significant only with "real" treatment, strongly suggesting the treatment effect is causal.

C.2 Childcare Provision

Results on childcare provision in Figure C2 and C3 are somewhat difficult to interpret, since they vary across the assumption of polynomials. The most RD estimates from two placebo tests are not significantly different from zero. Rather, placebo test with the promotion to 4th grade as a treatment seems to suggest that the eldest child's promotion to 4th grade makes the younger siblings' probability of receiving parental care lower, while the coefficients are less precise. I find significant positive effect only in some estimates in Figure (a) and (c), suggesting that reduction of maternal employment due to the treatment really increases parental care for the younger siblings. However, the results in these figures only suggest weak significance and in figure (b), with quadratic polynomial, the "real" treatment effects are no longer significant. These results suggest we can find significant increase in parental care under some assumptions but not under other assumptions. Given that odd-degree polynomials are preferred since these perform better at boundary points (Fan and Gijbels, 1996), however, we may conclude that the treatment effect is not spurious.

On the use of daycare center in Figure C3, both "real" and placebo tests show no significant effect, although some RD estimates are negative and significant in the "real" treatment, suggesting the eldest child's school entry reduces the possibility of using daycare center in daytime.

C.3 Symptom

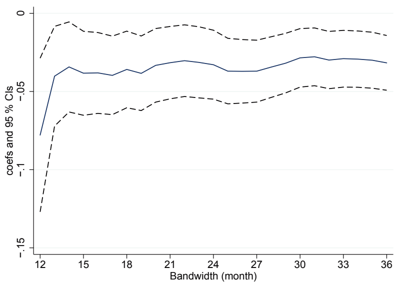
Results on the probability of having any symptoms are summarized in C4. On this outcome, I should note a place test with the eldest child's promotion to 4th grade suggests negative and significant effect in figure (d) to (f). This suggests the negative but imprecise treatment effects observed in left column are likely to be spurious. On the other hand, I seem to find negative and causal effect on the probability of taking a "fever". Some estimates in the left column in Figure C5 shows strong negative treatment effect. While those in the middle column also suggest significant negative effect, the estimates in the left column are more highly significant than those in the middle column.

C.4 Injury

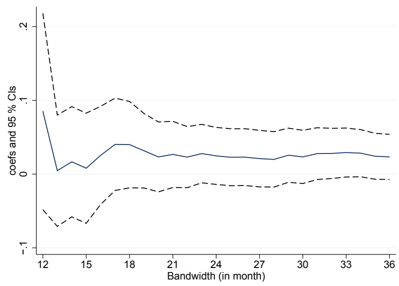
From Figure C6 to C8, the same placebo tests are implemented for the incidence of injuries. Supporting the validity of my placebo tests, all of the placebo treatment effect are not significant, as well as the "real" treatment effect.

C.5 Hospitalization

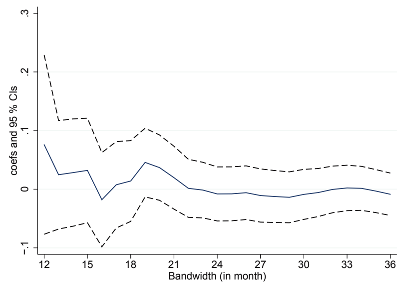
The two placebo treatments also are not significant for hospitalization in Figure C9. Although the upper limits of 95 percent interval is very close to zero in the many estimates in the left column, more than half are still insignificant. Given that the results in hospitalization cannot survive for the robustness checks by "donut-hole" RD, I can conclude there is no significant reduction of hospitalization caused by the eldest child's school entry.



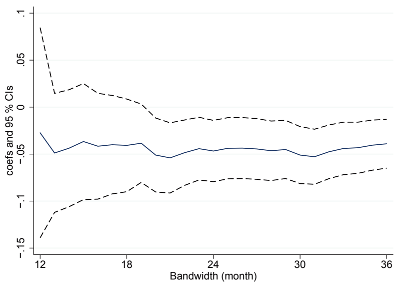
(a) Linear: 1st Grade



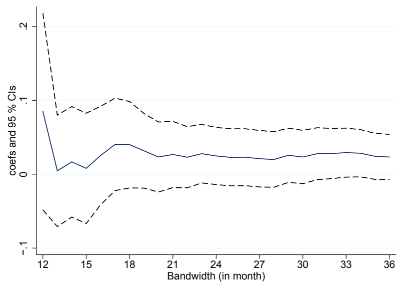
(d) Linear: 4th Grade



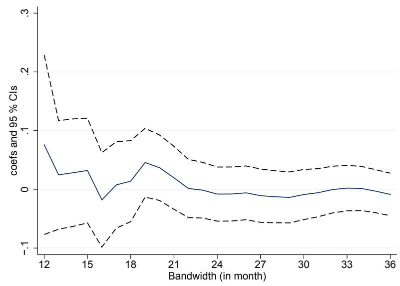
(g) Linear: 5th Grade



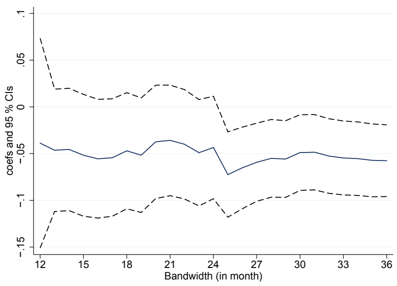
(b) Quadratic: 1st Grade



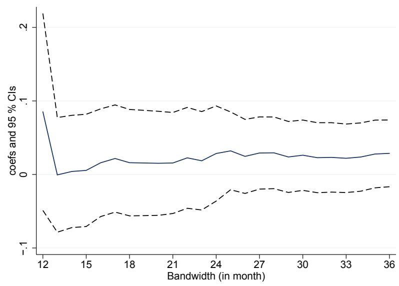
(e) Quadratic: 4th Grade



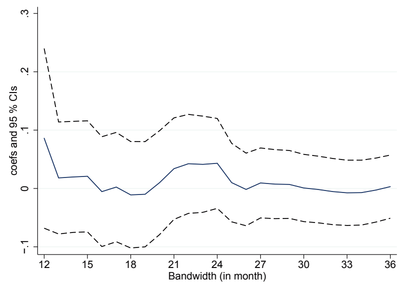
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



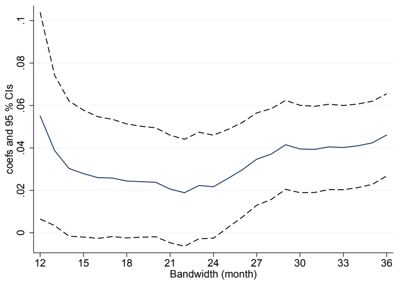
(f) Cubic: 4th Grade



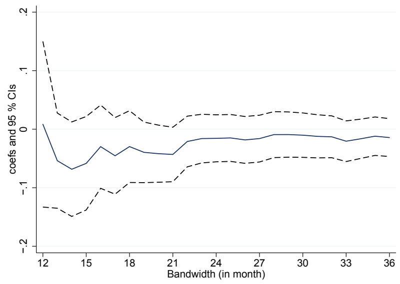
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

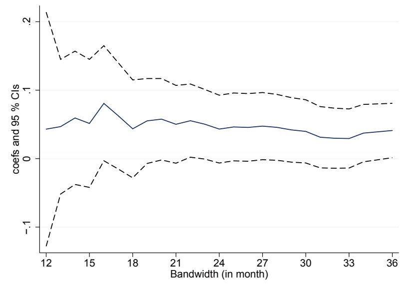
Figure C1: Mother's Working Status



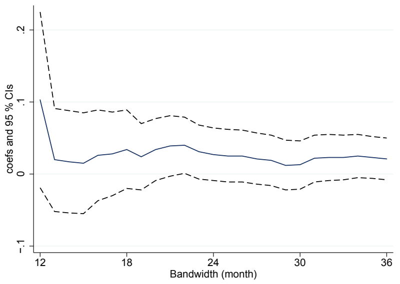
(a) Linear: 1st Grade



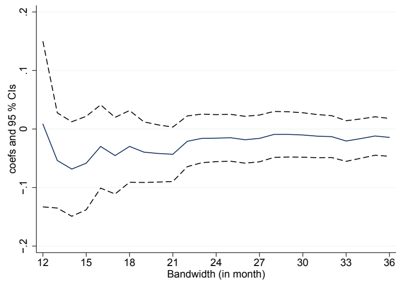
(d) Linear: 4th Grade



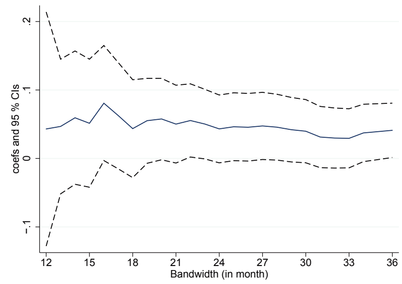
(g) Linear: 5th Grade



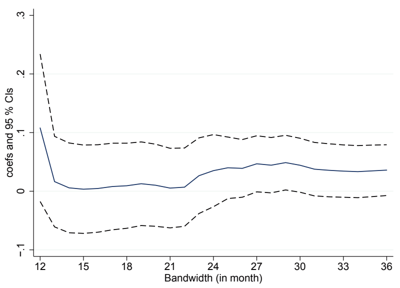
(b) Quadratic: 1st Grade



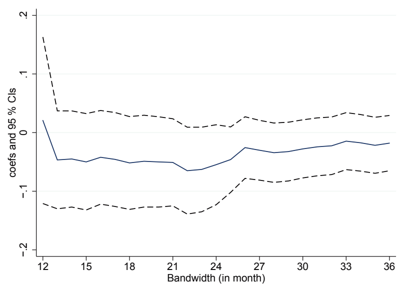
(e) Quadratic: 4th Grade



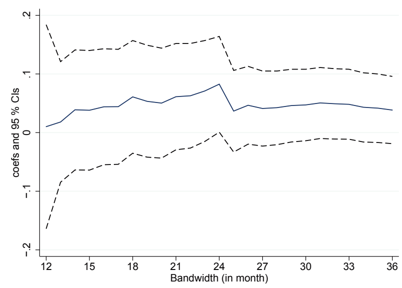
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



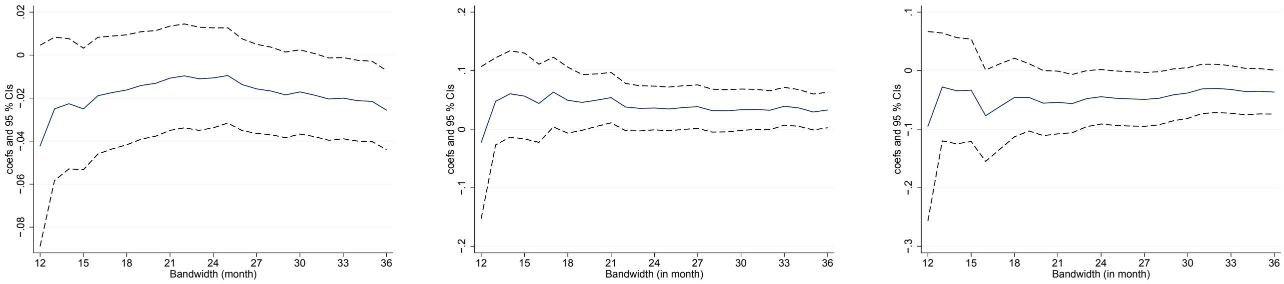
(f) Cubic: 4th Grade



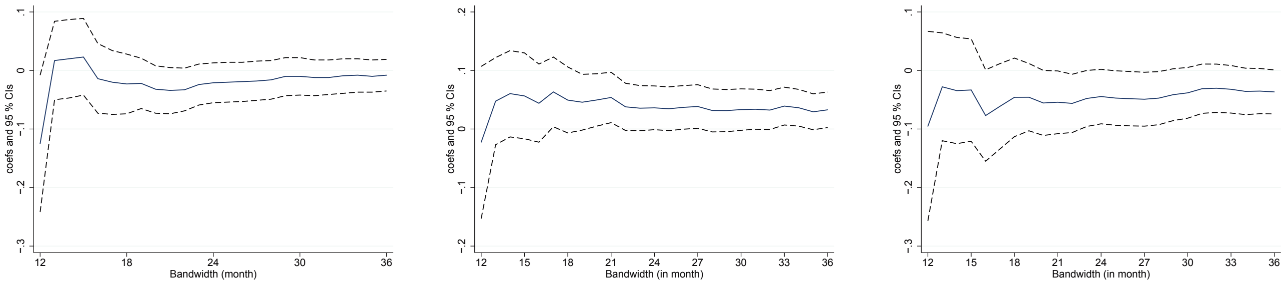
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

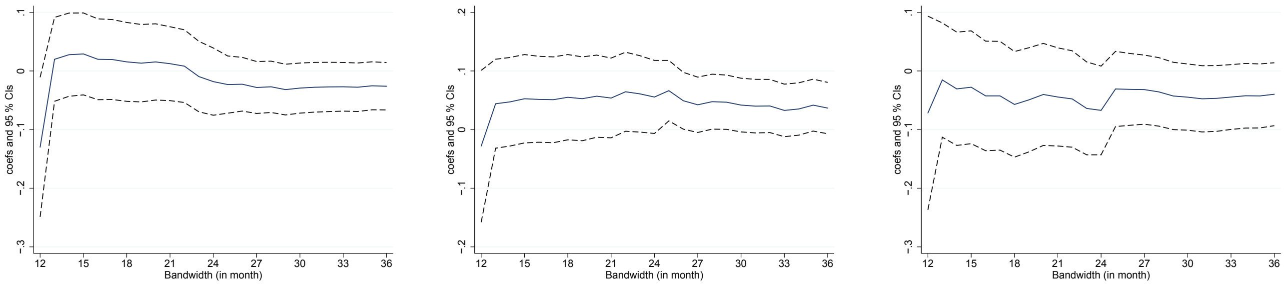
Figure C2: Probability of Receiving Parental Care



(a) Linear: 1st Grade (d) Linear: 4th Grade (g) Linear: 5th Grade



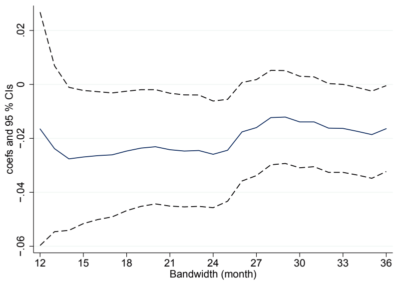
(b) Quadratic: 1st Grade (e) Quadratic: 4th Grade (h) Quadratic: 5th Grade



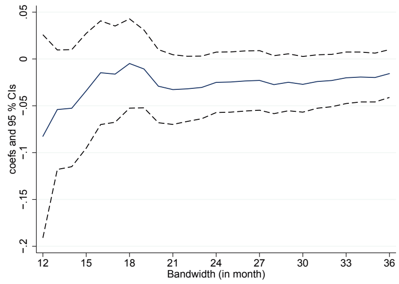
(c) Cubic: 1st Grade (f) Cubic: 4th Grade (i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

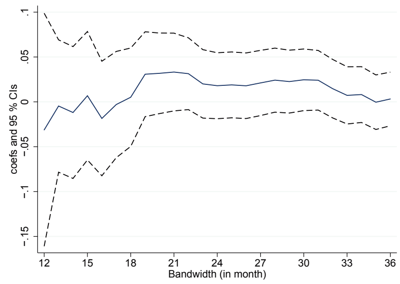
Figure C3: Probability of Utilizing Daycare Center



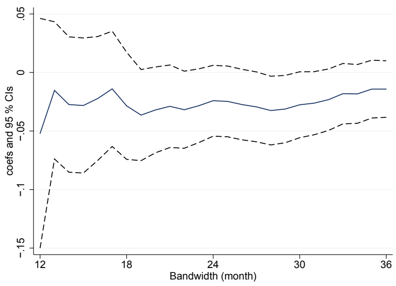
(a) Linear: 1st Grade



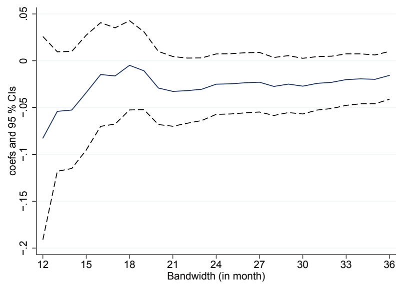
(d) Linear: 4th Grade



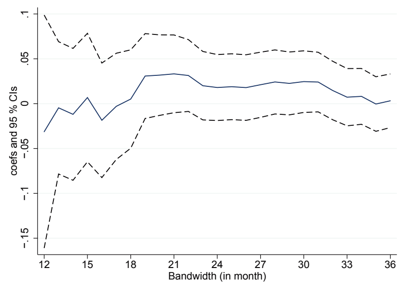
(g) Linear: 5th Grade



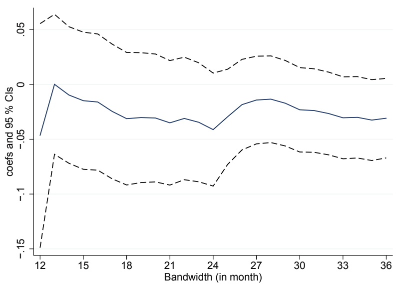
(b) Quadratic: 1st Grade



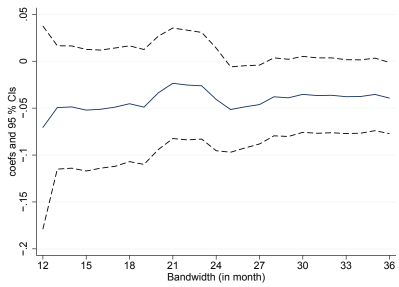
(e) Quadratic: 4th Grade



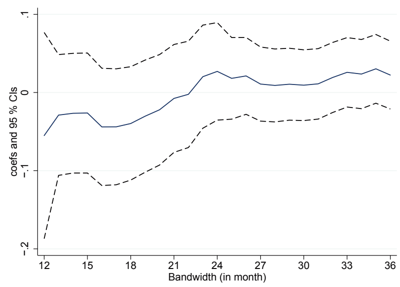
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



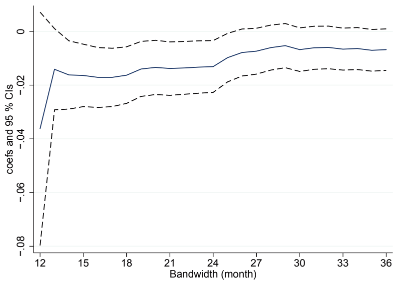
(f) Cubic: 4th Grade



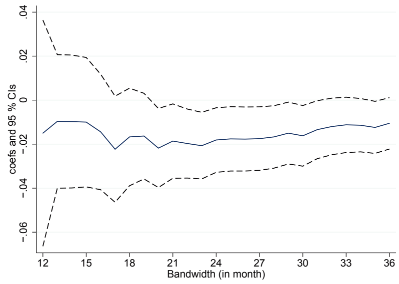
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

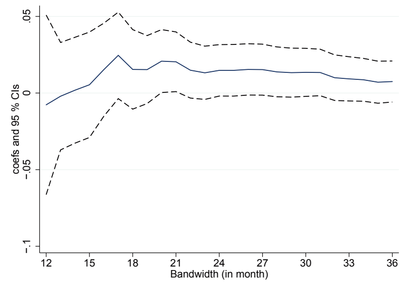
Figure C4: Probability of Having Any symptoms



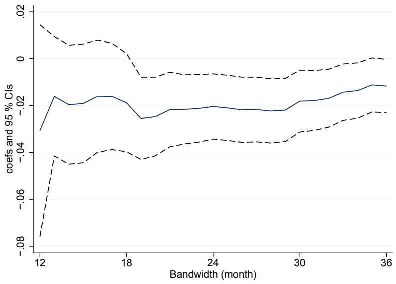
(a) Linear: 1st Grade



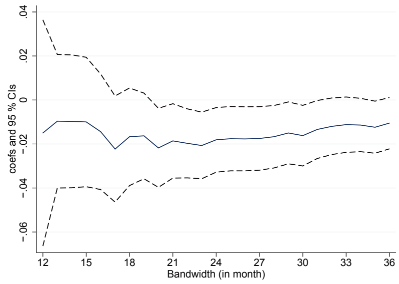
(d) Linear: 4th Grade



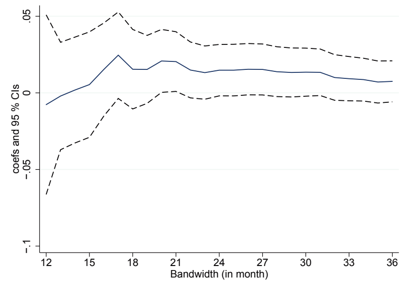
(g) Linear: 5th Grade



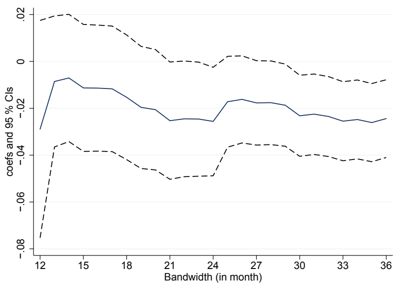
(b) Quadratic: 1st Grade



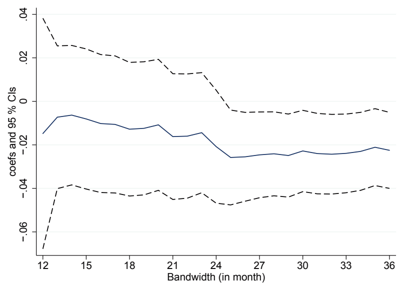
(e) Quadratic: 4th Grade



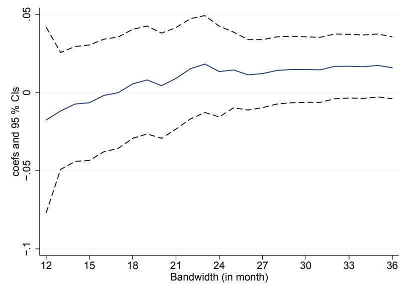
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



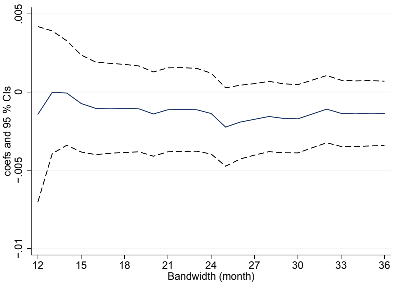
(f) Cubic: 4th Grade



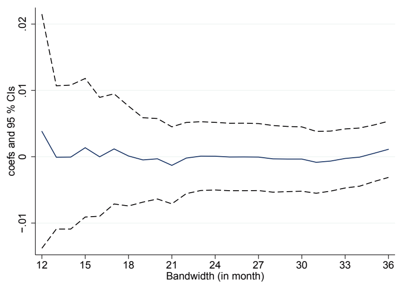
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

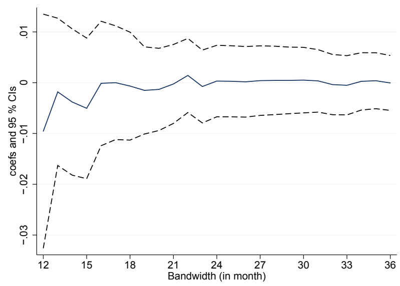
Figure C5: Probability of Taking a Fever



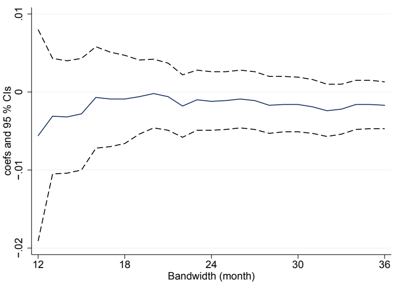
(a) Linear: 1st Grade



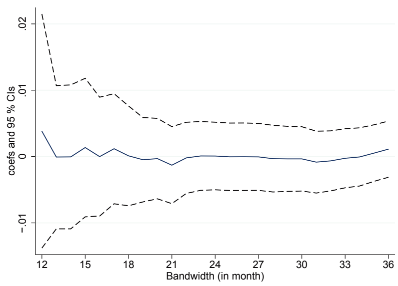
(d) Linear: 4th Grade



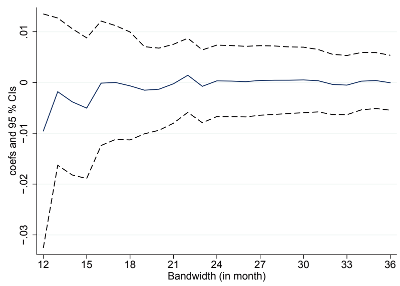
(g) Linear: 5th Grade



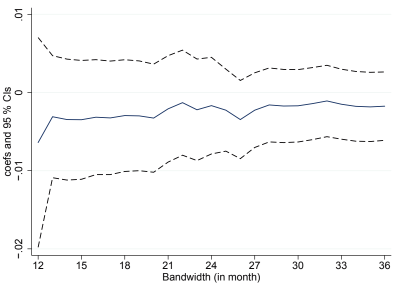
(b) Quadratic: 1st Grade



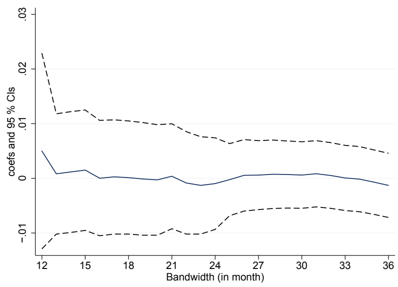
(e) Quadratic: 4th Grade



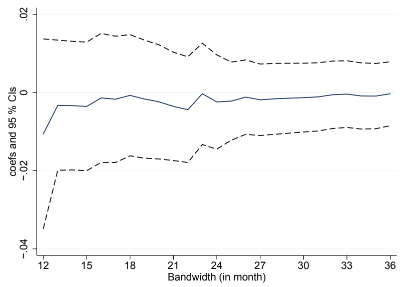
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



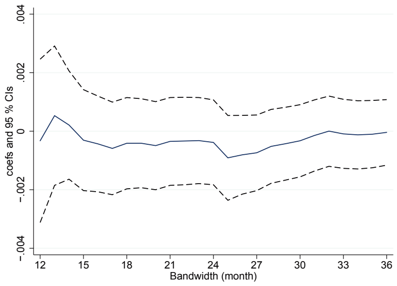
(f) Cubic: 4th Grade



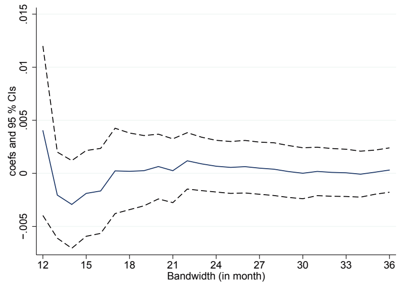
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

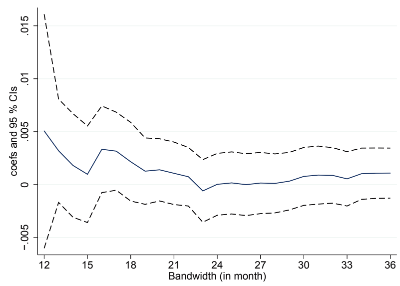
Figure C6: All Cause Injuries



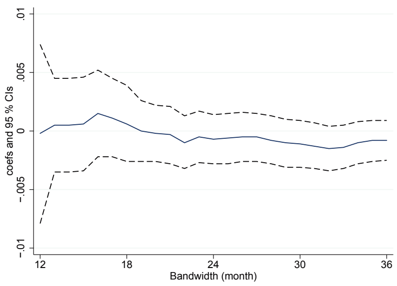
(a) Linear: 1st Grade



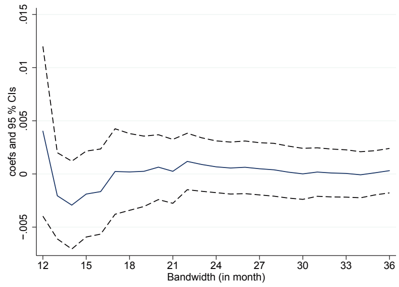
(d) Linear: 4th Grade



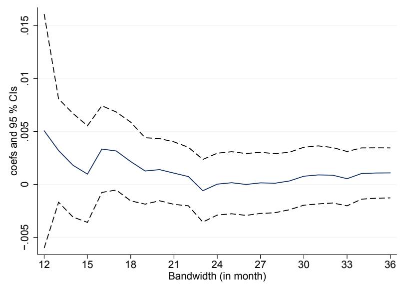
(g) Linear: 5th Grade



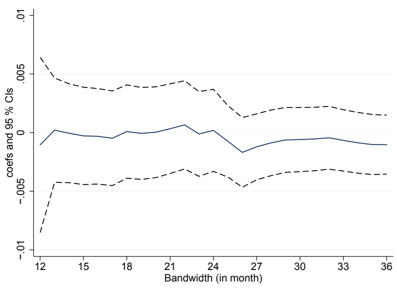
(b) Quadratic: 1st Grade



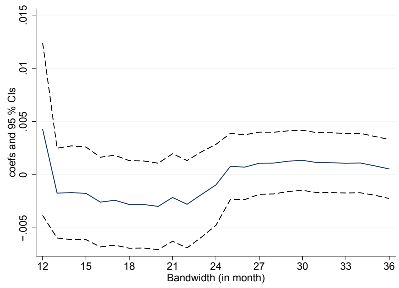
(e) Quadratic: 4th Grade



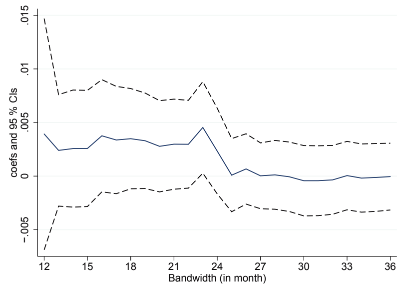
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



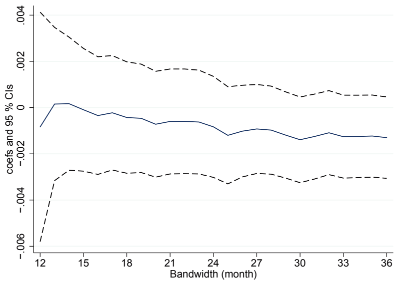
(f) Cubic: 4th Grade



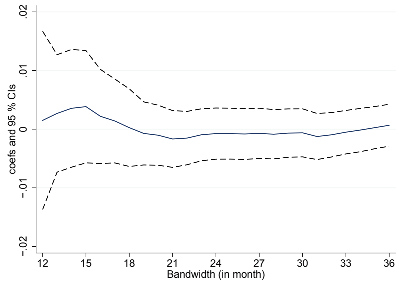
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

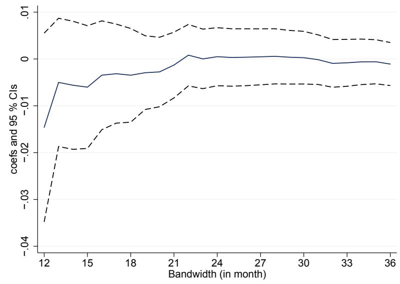
Figure C7: Fracture



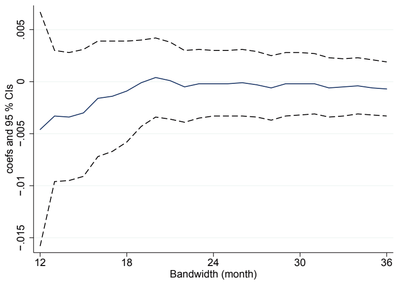
(a) Linear: 1st Grade



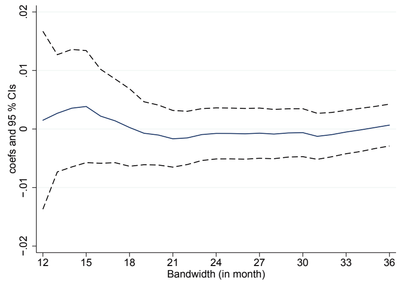
(d) Linear: 4th Grade



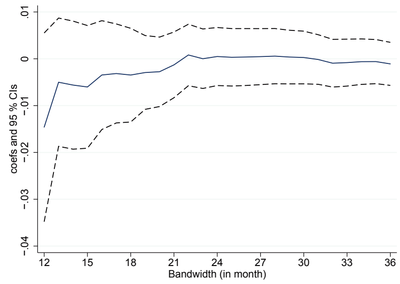
(g) Linear: 5th Grade



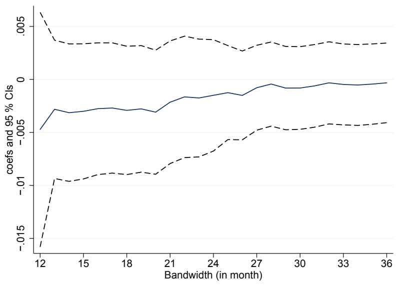
(b) Quadratic: 1st Grade



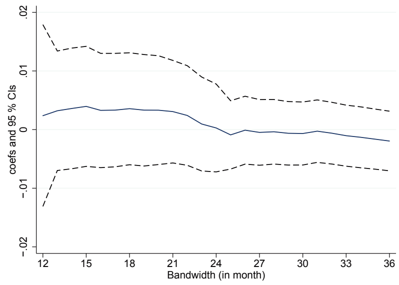
(e) Quadratic: 4th Grade



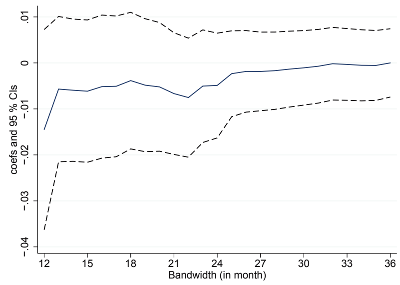
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



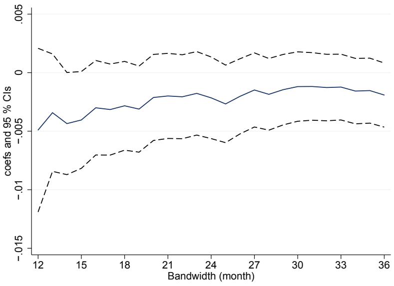
(f) Cubic: 4th Grade



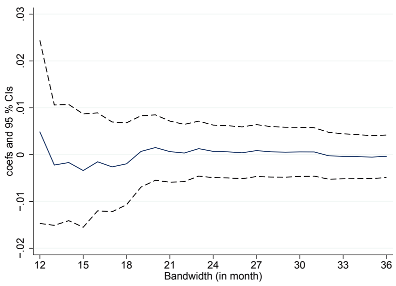
(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

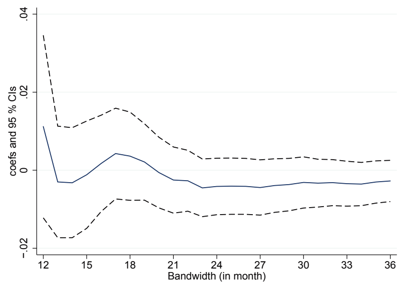
Figure C8: Other Injuries



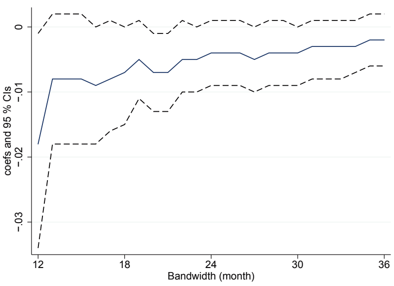
(a) Linear: 1st Grade



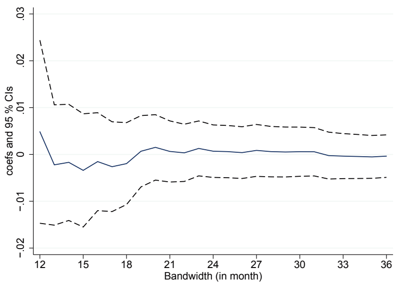
(d) Linear: 4th Grade



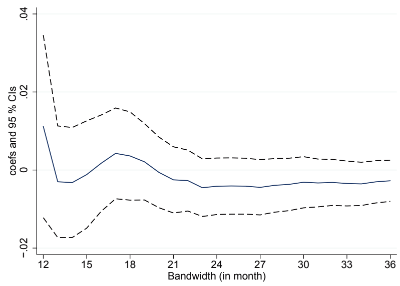
(g) Linear: 5th Grade



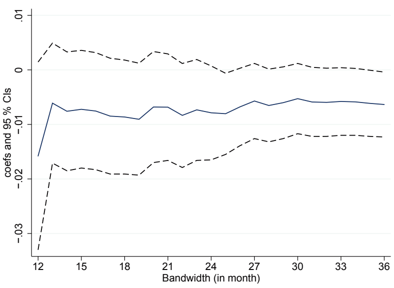
(b) Quadratic: 1st Grade



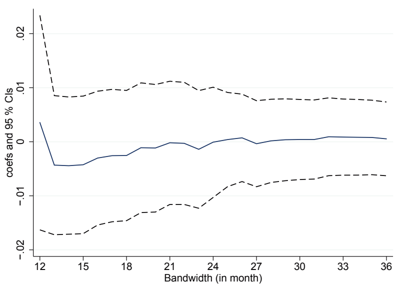
(e) Quadratic: 4th Grade



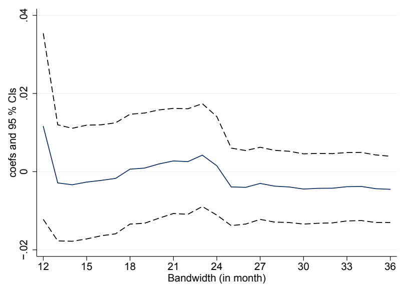
(h) Quadratic: 5th Grade



(c) Cubic: 1st Grade



(f) Cubic: 4th Grade



(i) Cubic: 5th Grade

Note: Horizontal axis represents the bandwidth which is changed from 12 months to 36 months, one by one. Solid line represents RD estimates and dash lines represent 95 percent confidence interval. The results in figure (a) to (c) are based on the "real" treatment which exploits the timing of the eldest child's school entry. From Figure (d) and (f), as a placebo test, timing of the treatment is changed to when the eldest child becomes the fourth grade in elementary school. From Figure (g) and (i), another placebo treatment with the promotion to fifth grade is exploited. All estimates control birth quarter fixed effects and other covariates. The controlled polynomials (linear, quadratic, cubic) are noted in the title of these figure.

Figure C9: Hospitalization

Chapter 3

The Effect of Patient Cost Sharing on Health Care Utilization among Low-income Children*

Reo TAKAKU†

Abstract

This paper examines how health care utilization among low-income children is affected by a reduction of the coinsurance rate, exploiting an institutional change in the Medical Subsidy for Children and Infants (MSCI) system in Hokkaido Prefecture, the north island of Japan, as a natural experiment. In 2004, the maximum age for MSCI recipients in Hokkaido Prefecture was raised from 3 years to include all children of preschool age. In this age group, the coinsurance rate was reduced from 30% to 0% in low-income children only, whereas it was reduced from 30% to 10% in higher-income children. As a result, the amount of co-payment reduction differed by 10 percentage points between low-income and middle- or higher-income children. A standard *difference-in-differences* technique was applied to analyze the effects of this policy change. The implied arc price elasticity among low-income children is −0.16, which is congruent with the commonly cited value (−0.2) presented in the RAND health insurance experiment and other experimental studies that cover middle- or higher-income populations. Nevertheless, the behavioral responses to cost sharing were found to differ across a variety of services and children's characteristics.

Keywords : price elasticity, cost sharing, difference in differences.

JEL classification : I10, H75

*I am grateful to Yukiko Abe, Michihito Ando, Janet Currie, Nobuyuki Izumida, Yoshikazu Kenjoh, Ayako Kondo, Sayaka Nakamura, Shuzo Nishimura, Wataru Suzuki, Atsuhiro Yamada, Michio Yuda, the seminar participants at the National Institute of Population and Social Security Research, and the session participants at the 9th World Congress of the International Health Economic Association for their helpful comments and advice. This project was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (Grant Number 24330097, Principal Investigator: Nobuyuki Izumida) The views expressed in this paper do not necessarily reflect the views of the Ministry of Education, Culture, Sports, Science and Technology. All errors are my own.

†Institute for Health Economics and Policy. Email:reo.takaku@ihep.jp

1 Introduction

The causal effect of insurance cost sharing on health care utilization and finance is not a new issue in health economics. Almost 30 years ago, authoritative evidence from a randomized controlled trial (RAND Health Insurance Experiment, or RAND HIE) was presented (Manning et al, 1987). Since then, numerous studies have examined how people respond to cost sharing, with some replicating and confirming the results of the RAND HIE (See Zweifel and Manning (2000) and Swartz (2010)). Despite the existence of "gold standard" results from the RAND HIE, there is a risk that these results may have little relevance in light of current socioeconomic environments and insurance systems; this indicates a need to produce additional estimates to better reflect current situations. In addition, it would be beneficial for researchers and policymakers to understand the heterogeneity of the real-world impact of cost sharing across individual characteristics (e.g., age, sex, income, and education) and a variety of health care services (e.g., outpatient care, drug prescription, and dental care) in order to design effective and appropriate cost sharing systems.

With the fulfillment of these needs as motivation, this paper shows how health care utilization in children from low-income households (hereafter referred to as low-income children) responds to changes in cost sharing by exploiting the exogenous reduction of coinsurance rates in Japan in 2004 as a natural experiment. Eighty percent of health care costs for preschool children in Japan is financed through public health insurance plans, while the remaining 20% is paid by patients as an out-of-pocket expenditure. However, local governments may subsidize this co-payment on a discretionary basis in order to reduce the financial burden on their constituencies. This subsidy program, designated the Medical Subsidy for Children and Infants (MSCI), has been dramatically expanded in the last decade.

By observing an institutional change in the MSCI in Hokkaido Prefecture, the north island of Japan, I estimate the arc elasticities of health care demand for preschool children. Given that children are unable to make decisions regarding visits to pediatricians and that these decisions are made by their parents or guardians, patterns in the demand for health care in children may be similar to those of adults. Empirically, Manning et al (1987) have also revealed that there is no difference in outpatient demand behavior between children and adults. This suggests that the price elasticity of health care demand in preschool children may be low in Japan, as Kan and Suzuki (2010) have reported the relatively low price elasticity among the Japanese adult population. In addition, Bessho (2012) finds no evidence that the MSCI has had any impact on the medical demand of preschool children, although that study was conducted using a cross-sectional framework[1].

In contrast to the previous studies conducted in Japan, this paper reports an arc elasticity of -0.16 in low-income children in Japan, which is similar to the central estimate of the RAND experiment (-0.2). To interpret this estimate, it should be noted that the estimated elasticity of low-income children in this paper (-0.16) is similar to those estimated in higher-income populations in other settings. This is con-

[1]With regard to school-age children, Bessho (2012) demonstrated that the MSCI had sizable effects on outpatient health care utilization.

sistent with the results of Chandra et al (2014), who report the price elasticity among low-income groups in Massachusetts. Indeed, the findings of this study indicate that low-income patients are very similar to their higher-income counterparts with respect to sensitivity to cost sharing, despite the well-known hypothesis that low-income patients are more susceptible to changes in co-payment than higher-income patients (Baicker and Goldman, 2011).

To estimate arc elasticities, I exploit an institutional change in MSCI coinsurance rates in Hokkaido Prefecture as a natural experiment. In October 2004, the prefectural governor Harumi Takahashi expanded MSCI eligibility from children under the age of 3 years to include all preschool-age children. Under this revision, the coinsurance rate for children in this age group was reduced from 30% to 0% for low-income children, whereas it was reduced from 30% to 10% in children from higher-income families. As a result, the reduction in co-payment differed by 10 percentage points between low-income children and middle- or higher-income children. As a baseline specification, a standard difference-in-differences (DD) framework is applied with low-income children as a treatment group and higher-income children as a control group. Since this paper utilizes the data of a public health insurance plan in a city which covers mainly vulnerable populations such as part-time workers and self-employed, the treatment and control group would be sufficiently comparable in their health care demand.

From the perspective of health policy, the importance of this paper lies in the fact that it focuses on the health care demand for low-income children. In addition to the declining birth rate in Japan, child poverty has recently become a policy issue attracting considerable public attention because the relative poverty rate among children in Japan has surpassed the OECD average and has been consistently increasing throughout the last decade(OECD, 2012). Hence, it is of particular importance to examine whether a reduction of co-payment actually improves access to health care in underprivileged children. In this regard, the estimates provided in this study may be useful to policymakers because it captures the sensitivity to cost sharing among low-income children through an analysis of a policy change designed mainly to benefit low-income households, rather than middle- or higher- income households.

Furthermore, the impact of changes to health insurance coverage for children is widely studied in both developed and developing countries because health statuses in childhood play a crucial role in human capital development and health statuses in later years. Several articles have revealed the impact of universal health coverage on children's medical care utilization, health statuses, and mortality; as well as the impact of children's health status on educational attainment and future productivity (Chen and Jin, 2012). In addition, the expansion of health insurance coverage in the United States has given rise to many studies on its impact on child health and health care utilization (Currie and Gruber, 1996a,b; Dafny and Gruber, 2005; Miller, 2012b). For example, Miller (2012b) reports that the Massachusetts Health Care Reform in 2006 reduced emergency room visits and improved children's subjective health. This paper can contribute to the comparative understanding of these issues by providing evidence from a recent Japanese experiment.

The remainder of this paper is structured as follows: Section 2 presents an overview of the existing literature. Section 3 summarizes the health care reform implemented in October 2004 in Hokkaido Prefecture and explains the research design employed in this study. Section 4 describes the data and summary statistics. Section 5 reports the results. Lastly, Section 6 offers a brief discussion and conclusion.

2 Prior Literature

Before introducing the study strategy and results, I present a brief literature review of the impact of cost sharing on health care utilization. Since there is already a substantial volume of studies that address the impact of out-of-pocket cost on utilization, I focus on studies that have analyzed the impact of cost sharing on outpatient care utilization[2] and studies that rely on experimental identification strategy. The first important contribution on this issue is from the RAND HIE (Newhouse, 1993), which has provided three main findings that are highly relevant to this study: (1) the central estimate of arc price elasticity of outpatient health care demand is -0.2, (2) children are as responsive as adults in outpatient care utilization, and (3) cost sharing has no significant impact on children's health and the physiologic measure of health (Valdez et al, 1985). Regarding the second issue, the similarities in sensitivity between adults and children appear logically sound, as parents—and not children—decide whether or not to visit a doctor. For this reason, as a first approximation, I compare the health care demand of children with that of adults in previous studies without assumptions of any substantial differences between the two.

Although these three findings have generally been considered a gold standard for health economic studies over the last three decades, several studies that exploit policy changes in a real-world setting have demonstrated price elasticity in various institutional settings and countries such as France (Chiappori et al, 1998), Belgium (Van De Voorde et al, 2001; Cockx and Brasseur, 2003), and Germany(Winkelmann, 2004). Among these studies, Chiappori et al (1998) present results that are markedly different from the results of the RAND HIE. They show that GP office visits are not elastic to small changes in cost sharing in France, and emphasize the importance of non-monetary cost, rather than out-of-pocket expenditures, on health care utilization. Other studies, however, report significant and moderate responses to cost sharing congruent with the results from the RAND HIE. Recently, two important studies from the United States also present results similar to those of the RAND HIE. Chandra et al (2010) report the price elasticity of GP office visits to be -0.1 for public employees in the state of California. Chandra et al (2014) study the impact of greater cost sharing for the low-income enrollees in Massachusetts' Commonwealth Care program. Given that Baicker and Goldman (2011) point out that "the evidence to support the contention that low-income groups are more price sensitive$\cdots$ seems less than fully reliable", it is of particular importance that the elasticity (-0.16) of low-income individuals reported in Chandra et al (2014) is very close to the estimates of the RAND HIE and studies that include populations from various income groups; this implies that the elasticity of low-income groups may not be different from that of more affluent groups[3].

[2] Many papers that address the impact of cost sharing focus on the impact on prescription drugs rather than on outpatient care. On the former issue, Goldman et al (2007) provides a comprehensive review.

[3] In this regard, we should take into account the possibility that the relationship between income and sensitivity to price

In comparison with this abundance of studies in various countries, there are, unfortunately, few studies based on quasi-experiments in Japan. Notable exceptions are Kan and Suzuki (2010), who exploit a policy change in the coinsurance rate in 1997 as a natural experiment; and Shigeoka (2013), who employs regression discontinuity design to focus on the discontinuous change in coinsurance rate for patients aged 70 years and above[4]. Among these studies, the results from Kan and Suzuki (2010) are directly comparable with those in this paper because both adopt a similar identification strategy. However, the elasticity (-0.05) reported by Kan and Suzuki (2010) is somewhat small, given that Shigeoka (2013) reports -0.17 for the price elasticity of outpatient care among an elderly population. Kan and Suzuki (2010) appear to attribute the reason for this small elasticity partly to the presence of patients who are not sensitive to cost sharing. After showing a complete lack of price elasticities among those with chronic diseases, they write "why do Japanese people visit physicians so frequently in comparison with other countries? $\cdots$ one possibility is that patients need to consult a physician every time a prescription drug is bought. For this reason, patients particularly those with chronic illnesses consistently need to visit a physician regardless of cost sharing" (Page 10). If this is indeed true, however, the elasticity among the elderly population would be expected to be much lower than their estimates because chronic diseases are more common among elderly people than in working-age adults. Although I do not have any definitive evidence to disentangle this contradiction, producing additional estimates would be necessary to contribute to understanding the impact of cost sharing in Japan.

Among the previous studies that address the impact of MSCI on children's health care utilization, Bessho (2012) suggests that MSCI does not have any impact on health care utilization in preschool children, basing this conclusion on a cross-sectional framework with nationally representative data from 2008. However, since there are many confounding factors that correlate with the expansion of MSCI and health care utilization (e.g., MSCI has been expanded mainly in affluent regions), the use of a cross-sectional framework would not appear sufficient to explicate the causal effect of cost sharing on health care utilization.

3 Research Design

3.1 Health Care Reform in Hokkaido Prefecture in October 2004

In Japan, the total fertility rate has been in decline over the last three decades, and child poverty has increased since the mid-1990s (OECD, 2012). In 2010, Japan's poverty rate of children under 18 years was similar with those of Canada and Italy but almost double the rates of Germany and Sweden. As the importance to solve these structural problems is widely acknowledged, it has been argued that the cost of child care, especially for low-income households, should be reduced. Following this argument, many local governments have expanded the policy which reduces out-of-pocket expenditure of child's health care

changes may differ by the type of disease. Chernew et al (2008) present a study of the impact of cost sharing in patients with diabetes mellitus and congestive heart failure, and report that "patients in low-income areas were more sensitive to co-payment changes than patients in high- or middle-income areas".

[4] In Japan, the coinsurance rate in outpatient health care services is 30% for the population under 70 years of age. This is reduced to 10% after the 70th birthday, except for the elderly in high-income households. These institutional settings generate discontinuous changes in health care utilization shortly before and after the 70th birthday.

utilization. Although medical costs for preschool children require a 20% co-payment throughout Japan[5], almost all municipalities provide an additional medical subsidy through appropriations of local tax revenue under the MSCI program (in Japanese: *Nyūyōji iryōhi josei*). The maximum age of children covered by this subsidy has been dramatically raised in the last decade. The share of municipalities that expanded the age criterion beyond preschool age was only 9.9% in 2000, but this increased to 98% by 2011[6]. In Hokkaido prefecture, as in the other prefectures in Japan, MSCI eligibility has been also expanded during the last decade. In particular, the reform in October 2004 was one of the largest reforms which reduced out-of-pocket expenditure for preschool child's health care utilization.

From the viewpoint of an *experimental ideal*, we should exploit pre-post changes in the MSCI in multiple municipalities to explore the impact of MSCI because of the wide range of inter-municipal variations[7]. However, it is difficult to conduct such an analysis due to the low availability of health insurance claims data in Japan. Instead, I focus on data from a municipality in Hokkaido Prefecture (designated Y City) and exploit an institutional change in Hokkaido Prefecture's MSCI as a natural experiment[8]. In this prefecture, eligibility for the MSCI was expanded from children aged under 3 years to include preschool-age children[9]. For low-income children, the coinsurance rate was reduced from 30% to 0%. For the other children, however, the rate was reduced from 30% to 10%. By exploiting this institutional change, I can estimate the impact of an additional 10 percentage points reduction in the coinsurance rate, as was done in previous studies such as Chiappori et al (1998), Cockx and Brasseur (2003) and Kan and Suzuki (2010).

On the definition of "low-income", Hokkaido prefecture defines low-income households as those in which all household members are exempt from residence-based taxes, and the city implements various support programs for these households based on this definition. The lowest annual taxable income threshold for residence-based tax is 1.29 million JPY for a household comprising a married couple with two children. If the taxable income exceeds this threshold, the parents of the household are required to pay residence-based taxes and their children would then be categorized in the middle- or higher income group. The criteria for defining " low-income " are revised every July.

Table 1 summarizes the changes in coinsurance rates in Y city. The coinsurance rate for children of

[5]The coinsurance rate for preschool children was reduced from 30% to 20% in 2008 throughout Japan by an initiative of the national government.

[6]For the expansion of MSCI in Tokyo Prefecture, see Nishikawa (2010) and Nishikawa (2011).

[7]In general, the system of MSCI differs across municipalities in four aspects. First, municipalities can freely set the eligible age for the MSCI within their jurisdiction. For children older than the upper-limit age, the municipalities offer no benefits. Second, municipalities can restrict eligibility for children from high-income households. In 2012, 25.6% of all municipalities adopted a form of restriction based on household income ceilings (Ministry of Health and Welfare, 2013). Third, municipalities can also choose in-kind transfers or repayment for the subsidization. The majority of municipalities adopt in-kind transfers in which the out-of-pocket payment in clinics and hospitals is reduced directly, whereas some municipalities adopt repayment for their MSCI. Under the repayment system, patients (or specifically, their parents) have to pay the co-payment when they obtain health care from pediatricians, but the co-payment is eventually reimbursed in full or in part. Finally, the amount of subsidy varies across municipalities. Some municipalities charge very small out-of-pocket costs for health care utilization to promote the "appropriate" use of pediatric services, whereas the co-payment is rendered completely free in 54% of all municipalities.

[8]Since this study uses data from a city, migration from other municipalities would be a potential threat. However, this is not likely because MSCI eligibility was concurrently expanded in almost all municipalities in Hokkaido prefecture, and the reform that raised the eligibility age of MSCI in Y city was also implemented in its neighboring municipalities.

[9]In Japan, children admit to elementary school on April at the age of 6.

married couples was 30% before the reform, but it was reduced to 0% for low-income children whose household members were exempt from residence-based taxes. On the other hand, coinsurance rates for middle- and higher-income households decreased to 10%. As for children with lone parents, their coinsurance rate remained unchanged they are already recipients of a more generous welfare program[10]. Also, there is no deductible for the first visit to a doctor in Y city due to the provision of an additional subsidy, which was suitable for my empirical analysis. In other municipalities in Hokkaido prefecture, patients must pay a 580 JPY deductible for their first visit. In addition, MSCI in Y city does not restrict eligibility based on household income. With regard to the administrative system for the subsidy, Y city adopts in-kind transfer for MSCI.

3.2 Difference-in-Differences

The identification strategy used in this study is fairly straightforward. A standard DD technique is exploited to estimate the impact of a quasi-experimental change in cost sharing. The treatment group comprises low-income children whose coinsurance rate was reduced to 0%, while the control group comprises higher-income children whose coinsurance rate was reduced to 10%. My sample consists of monthly data before and after October 2004. In addition, children aged 36 months to 72 months when the reform started are included in the analysis.

I begin by estimating DD models of the form,

$$M_{it} = \alpha_0 + \alpha_1 Low_{it} + \alpha_2 Post_{it} + \delta(Low_{it} \times Post_{it}) + X_{it}\gamma + Ind_i + Time_t + \epsilon_{it}, \quad (1)$$

where M_{it} is the medical utilization of individual i in a month t. Low_{it} is a dummy variable that is equal to 1 if a child i belongs to a low-income household in a month t. $Post_{it}$ takes the value of 1 for the period after October 2004[11]. α_0 is a constant term. α_1 and α_2 are the coefficients for Low_{it} and $Post_{it}$, respectively. Ind_i is an individual fixed effect, and $Time_t$ is a monthly dummy variable. The term X_{it} is a vector of time-varying observables that affect health care utilization in children. In this equation, δ is an adjusted DD estimator of policy change and tracks the behavioral response to a 10% change in cost sharing. In addition, this δ can be interpreted as sensitivity to cost sharing in low-income children, rather than middle- or higher-income children. Given that child poverty has gathered increasing public attention, the sensitivity of health care demand for underprivileged children may be of importance from a policy perspective.

Following Finkelstein et al (2012), I report the results on the decision to have a positive visit (extensive margin) and the total number of visits. This separation is based on a two-part demand model, which was first incorporated in Duan et al (1983). Since the standard principal-agent theory in health care utilization

[10] The Medical Subsidy for Children with Single Mothers (MSCSM) provided free health care services for children of single mothers only. Single fathers were not eligible for the MSCSM before October 2004 in Hokkaido. After the reform, however, the MSCSM was renamed “Medical Subsidy for Children with Single Parents” and children with single fathers also became eligible for the public medical subsidy.

[11] In the empirical analysis, this variable is absorbed in monthly dummies.

predicts that a patient determines whether to visit a physician and that the physician determines the entire treatment schedule after the first contact, conventional wisdom leads us to estimate the extensive margin and the total effect separately. If we find no significant effect on the former, it means that the overall effect of co-payment reduction is due to behavioral changes in physicians. In contrast, if the extensive margin is highly responsive to co-payment reduction compared to the changes in the total number of visits, we can conclude that behavioral changes in patients, not physicians, account for the overall effect of the reform. In addition to the analysis on the number of visits, the results on spending per visit are presented, which is observed only for children who visit pediatricians more than once in each unit of time. This variable appears to represent the intensity of care. As in Kan and Suzuki (2010), the natural log of spending per visit is calculated to address the skewed distribution. Finally, the overall effect of the October 2004 reform is considered as the sum of the effect on the total number of visits and spending per visit.

Equation (1) is estimated using a fixed effects model to eliminate time-invariant unobservables that affect children's health care utilization. In particular, it can be reasonably assumed that the characteristics and preferences of parents and children did not undergo any substantial changes during the short periods before and after the reform. Then, the results without X_{it} are presented as my central estimates because there are no plausible variables for X_{it} in my data after controlling for individual fixed effects and time effects. Instead, the whole sample is divided into several subgroups based on the characteristics of children and estimate the heterogeneous impact of coinsurance rate reduction across subgroups.

Nevertheless, there are potentially serious threats to the accuracy of my identification strategy; First, the underlying trend between the treatment and control groups might not be parallel in the absence of the October 2004 reform. If this is the case, the naïve DD estimator would necessarily fail to capture the true impact of the reform. With regard to this point, it should be noted that income is a potentially strong predictor of health care demand. Since the seminal work of Case et al (2002), who investigated whether the association between socioeconomic status (SES) and health could be found among children, numerous papers have confirmed that household income plays a crucial role in children's health and health care demand, although recent studies have presented mixed results (Apouey and Geoffard, 2013). Therefore, it is questionable that the parallel trend assumption holds true. Second threat is that the sensitivity for the price changes may differ across income status. Since the treatment and control groups were both affected by the reform, differential responses to the coinsurance rate reduction would potentially result in the biased estimate to the price elasticity. On these two threats, however, it should be noted that this study uses the data of enrollees from Citizens' Health Insurance (CHI) which generally covers vulnerable populations such as self-employed and part-time employees[12]. Hence, even if the definition of the treatment (low-incomes) and control (middle- and higher- incomes) are based on the household income, we can assure these groups would be sufficiently comparable with each other because they share many socio-economic backgrounds.

[12]Details in the insurance coverage for children are summarized in Appendix A.

4 Data and Descriptive Statistics

4.1 Data

This paper uses insurance claims data of CHI enrollees in Y city, located in Hokkaido Prefecture. With regard to the characteristics of CHI enrollees in this city, it should be noted that their income status is generally lower than the Japanese average. First, although I am unable to disclose the geographic location of Y city, the characteristics of this city are consistent with those of a small city. The population is approximately 40,000 and the major industries are fishing and tourism. Second, CHI mainly covers the proportion of the population who are likely to be vulnerable, such as farmers, the self-employed and unemployed, as well as their dependents.

For the construction of the data set, insurance claims data from Y city from April 2000 to March 2011 are utilized. These claims are monthly bills, which include health care costs, number of visits, clinical departments of the clinic or hospital where care was provided[13], a code denoting an impatient or outpatient episode, and age and sex of the patient. In addition to these bills, I use a list of CHI enrollees in Y city that includes the years and months when a person was enrolled in and withdrawn from the CHI. This CHI enrollment list is matched to the insurance claims data set by using the household ID, patient age and sex.

Among the complete data set, I focus on the period from April 2003 to March 2006, spanning the 18-month duration before and after October 2004[14]. Subsequently, children aged 36 months to 72 months when the reform started are included in the analysis. When a child was over 72 months or below 36 months, the observations are excluded even if the child was aged 36 months to 72 months in October 2004 because children's health condition may change when they enter elementary school in April at the age of 6. Next, I exclude children who have more than two siblings and children from lone parents since the number of such cases included in my data is small. Two children whose outpatient health care costs were extremely high are also excluded since their out-of-pocket expenditures reaches stop-loss amounts[15]. Finally, monthly data sets from 39 low-income children (the number of observations is 858) and 142 middle- or higher-income children (the number of observations is 3,734) are constructed.

The details of the variables are summarized as follows:

Health Care Utilization

On the health care utilization, a dummy variable is created that takes the value of 1 if a child utilizes outpatient care more than once in a month and a value of 0 otherwise. If a child is enrolled in CHI in a

[13] The data do not include disease names.

[14] As is mentioned previously, a low-income child is defined as one whose household members are exempt from residence-based tax. Hence, any tax reform may change the definition of low-income households in this paper. With regard to this point, it is impossible to use data from before April 2002 due to a raise in the lower threshold of taxable income for residence-based tax. In addition, residence-based tax was drastically reformed in 2006 through comprehensive revisions of local taxes and intergovernmental transfers implemented by then-Prime Minister Junichiro Koizumi, who had intended to give local governments more fiscal autonomy. In order to avoid the influence of these institutional changes, my analysis focuses on the period from April 2003 to March 2006.

[15] In Y city, out-of-pocket expenditure for outpatient care never exceeded 12,000 JPY due to stop-loss.

given month but their bill is not found in insurance claims, I can reasonably confirm that they did not use health care services in that month. To calculate spending per visit, total health care costs are divided by the number of visits. This variable is not observed if a child did not utilize any health care services in that month.

Income Status

To identify the household income level, I use a code in the insurance claims that shows whether family members of a patient pay residence-based taxes. If none of the child's family members pay residential tax, the child can be reasonably categorized as being from a low-income household. Since this code is included in insurance claims, and not the enrollee list, it is impossible to confirm the household income level of a child whose family members have never visited a doctor. However, this problem may be negligible since it would be quite rare for none of the family members to have seen a doctor at least once a year. If someone in the household (father, mother or child) visits a doctor, income status of the household is identified with a high level of certainty.

Household Characteristics

Since the insurance claims data and enrollment list do not include accurate marital statuses, it is impossible to identify directly if children are from households with married couples or lone parents. Instead, a child is categorized as being from a lone-parent household if their household includes only one person aged 20 to 60 years; otherwise, a child is categorized as being from a household with a married couple.

Finally, two limitations related to using the insurance claims data of Y city should be noted. First, the sample does not represent the entire Japanese population. As mentioned previously, Y city is small and it is plausible that children and their parents who live in urban areas have different preferences regarding health care services. Second, the insurance claims data do not contain information on important personal characteristics such as educational attainment and employment status, although the longitudinal nature of the data may alleviate some of these problems, as mentioned in Kan and Suzuki (2010).

4.2 Descriptive Statistics

Table 2 presents the summary of my data according to children's household income status. The mean follow-up period is approximately 2 years for both groups. The proportion of children who visit doctors more than once a month is 48% in low-income households, while it is 52% in higher-income households. Visits to pediatricians account for the most frequent doctor visits in both groups. The number of visits per month is 1.58 in low-income households and 1.23 in higher-income households. Subsequently, health care spending per visit is calculated as a proxy for treatment intensity. The natural log of spending per visit is 6.16 for low-income children, which is almost identical to that of their higher-income counterparts. This suggests that children receive identical treatment once they visit the doctor, regardless of their household income status. However, there are significant differences in treatment intensity of dental care services by income status. The log of spending per visit is 6.17 for higher-income children, which indicates that their spending per visit is 20 percentage points higher than in low-income children.

As for the characteristics of children, low-income children are more likely to be firstborn. The proportion of firstborn children is 58% for low-income households-16 percentage points higher than for higher-income households. We can assume that firstborn children are different from later children in the use of health care services, because parents may not have much experience in dealing with children's health problems in their first child. Hence, it is helpful to present the results divided by birth order subgroups.

5 Results

5.1 Graphical Representation

Turning to the empirical results, I present the unadjusted sample means of the outcome variable by the treatment and control groups to check the key assumptions required for DD equation. Panel A in Figure 1 presents the average number of outpatient visits during April 2004 to October 2005, with the monthly raw data grouped into half-year average values. First, the number of visits among preschool children from low-income households, denoted using diamond-shaped markers with blue solid lines, was almost stable at approximately 1 to 1.25 days per month before the October 2004 reform was implemented; the number among preschool children from higher-income households, denoted using circle markers with red solid lines, decreased slightly during the same period. In the standard DD framework presented in Equation (1), the former is used as the treatment group and the latter as the control. The DD estimation appears to have surface validity as the trends were sufficiently similar during the pre-reform period. However, it is not the case for the spending per visit in Panel B. In Panel B in Figure 1, trends during the pre-reform period seems to be different because of the irregular increases of spending per visit in the low-incomes during October 2003 to March 2004. Therefore, my DD specification may be inappropriate for this variable.

5.2 Main Results

The estimation results based on Equations (1) are reported in Table 3. Columns (1) and (2) both present the DD estimates of the probability of doctor visits. Column (1) uses the raw monthly data in the estimation while Column (2) groups the data into pre-post terms to address the underestimation of standard errors in the DD estimator with long time series, as recommended by Bertrand et al (2004). In Column (1), the DD estimate of the probability is 0.1086 and significant, suggesting that the probability to use outpatient health care services more than once a month would increase by approximately 11% after the reform. Although the DD estimate is not significant in Column (2), this is largely due to a difference in the measurement of the "probability"; the dependent variable in Column (2) is a dummy variable that takes the value of 1 if a child saw a physician or pediatrician more than once in the pre-post reform period spanning 18 months.

Furthermore, the reform shows a significant impact on the total number of visits in Columns (3) and (4). The DD estimate based on monthly data indicates that a 10 percentage point reduction in cost sharing increases the total number of visits by 0.49 days per month. This result still demonstrates robustness even if the data are collapsed into the pre-post period (Column [4]).

In addition, there is no significant reduction in treatment intensity as a result of the reform. The coefficient of spending per visit in Column (5) is negative but insignificant. Although the DD estimate in Column (6) presents a significant result, it may be because the underling trends would not be similar on this outcome. For instance, Panel B in Figure 1 shows that spendings per visit among low-incomes irregularly increased during October 2003 to March 2004. This may lead to the severe exaggeration of the negative impact on spending per visit. Without such a irregular increase, there would be no systematic changes in the spendings per visit. Consistent with the RAND HIE (Leibowitz et al, 1985), which reveals that the average medical bill per episode did not differ among cost sharing and free plans, this paper may suggest that cost sharing does not affect the amount of services provided once parents take their children to the doctor and medical treatment is initiated[16], while the more accurate examination would be needed for this issue.

Using the estimated parameters, I calculate arc elasticity—which uses the midpoint rather than the initial point to measure the magnitude of changes—of outpatient health care demand (ϵ), applying the DD estimate from Table 3 to the following formula and taking into account the fact that the average number of visits in low-income children is 1.58, as shown in Table 2:

$$\epsilon = \frac{\frac{q_a - q_b}{(q_a+q_b)/2}}{\frac{p_a - p_b}{(p_a+p_b)/2}} = \frac{\frac{0.4904}{1.58}}{\frac{0-10}{(0+10)/2}} = -0.155 \quad (2)$$

,

where q represents the outcomes, p is the coinsurance rate, and a and b indicate the "after" and "before" periods, respectively. Given that the coinsurance rate in my control group decreased from 30% to 10%, I can assume the treatment group was affected by the *additional* reduction in the coinsurance rate from 10% to 0%. This assumption postulates that health care utilization in the treatment group in a counterfactual case (where their coinsurance rate in the post-reform period would decrease from 30% to 10% instead of the actual 0%) can be extrapolated using the actual trend of the control group. This is equivalent to assuming that price elasticity of the treatment group (i.e., low-income children) would be the same as the control group (middle- and high- income children). Given the contention that the former would be more sensitive to changes than the latter, this assumption would seem to be demanding. It is, however, a reasonable assumption since a recent review paper summarizes that there is no reliable evidence to support this contention (Baicker

[16] On this point, it should be noted that my results resemble partial equilibrium effects rather than the market-wide and general equilibrium effects of health insurance, which are outlined in Finkelstein (2007) and Kondo and Shigeoka (2013). Finkelstein (2007) notes the importance of supply-side changes that result from the expansion of public insurance, and shows that the overall increase in medical costs caused by the introduction of Medicare is underestimated if only demand-side responses are taken into account. In the same spirit as this paper, Kondo and Shigeoka (2013) study the impact of the introduction of universal health insurance in Japan and report a great increase in the number of hospital beds. The reason why my results resemble partial equilibrium evidence is that the number of newly qualifying children was relatively small compared to the general population of Y city and there was no market-wide change in health care services.

and Goldman, 2011). Furthermore, Chandra et al (2014) find price elasticity of low-income patients to be similar to those calculated for higher-income patients. Even if the assumption does not hold, the estimated elasticity of low-income patients would not vary because the denominator in the formula for arc elasticity always takes a value of -2 (which denotes a 200% decrease compared to the mid-point) if the co-payment of the treatment group is reduced to 0 in the post-reform period[17]. Taken together, I conclude that the arc elasticity of health care demand among low-income children is -0.16, which is essentially the same as the commonly cited value of RAND HIE, i.e., -0.2.

When comparing my results with those of a previous quasi-experimental study of Japanese adults (Kan and Suzuki, 2010), the arc elasticities in this paper are two or three times higher. In addition, I find significant effects on doctor visits in both the extensive margin and total visits, while Kan and Suzuki (2010) do not find any effect on the number of visits. Given that Kan and Suzuki (2010) attribute their low estimates partly to the presence of patients with inelastic demand, such as patients with chronic illnesses, the reason why my estimate among children is higher than that of Kan and Suzuki (2010) may be that chronic illness such as diabetes and hypertension are uncommon among children. However, a recent study has reported that price elasticity among elderly populations is also higher than that of Kan and Suzuki (2010), regardless of the fact that chronic illness would be more prevalent among the elderly population; Shigeoka (2013) reports the outpatient price elasticity of the elderly population to be -0.17, applying an age-based regression discontinuity technique to grasp the impact of co-payment reduction at the age of 70. On this point, I posit that my estimates are more credible than those of Kan and Suzuki (2010) because their results do not seem to be robust to alternative data sets. Although their results are based on analyses of two data sets (2-year data and 3-year data[18]), the results on the number of visits contradict each other (Table 6 in their paper). The effect of co-payment increase on the number of visits is estimated to be significantly *negative* in the 2-year data, whereas the effect was estimated to be *positive* in the 3-year data. This contradiction is possibly due to the violation of the parallel trend assumption. If time trends of outcome variables are not similar between treatment and control groups, these two data sets would invariably present different results; however, we are unable to confirm whether the assumption would be valid or not since Kan and Suzuki (2010) do not show the comparison of trends in the pre-reform period. In contrast, the trends of the number of visits in my study are clearly shown in Figure 1. Hence, any bias arising from the violation of the parallel trend assumption would be unlikely to have an effect on my analysis.

In addition, it is also of particular importance that this paper presents the elasticity among low-income children. Although conventional wisdom suggests that the poor may be more sensitive to cost sharing than more affluent individuals, this paper does not support this conclusion. Instead, the estimated elasticity is very close to those calculated for middle- or higher-income children, as is also indicated in Chandra et al (2014), who examine how low-income populations in Massachusetts respond to cost sharing.

[17] Namely, $denominator = \frac{p_a - p_b}{(p_a+p_b)/2} = \frac{0-p_b}{(0+p_b)/2} = -2$.

[18] Two-year data include each year before and after the co-payment increase in September 1997, and 3-year data cover the period from April 1996 to March 1999.

Lastly, my results are markedly different from Bessho (2012), who finds no impact of MSCI on health care utilization in preschool children using nationally representative data. It is possible that this difference is due to the differences in research design: a major difference is that Bessho (2012) is based on a cross sectional framework, while this study employs a natural experiment.

5.3 Results by Subgroups

Estimations by subgroup may provide detailed insight into the heterogeneous effects of MSCI reform. Table 4 presents the DD results based on monthly data according to sex and birth order. As shown in Columns (1) and (2), I find large differences in the responses to cost sharing between boys and girls. The coefficient of the number of visits from DD is approximately 0.8 among girls, which is equivalent to an arc elasticity of -0.3[19], while the coefficient among boys is not so high. In addition, the results suggest that birth order is an important factor in health care demand. The DD estimate in Column (3) shows that in firstborn children, there is an increase in doctor visits by 0.86 days per month after a 10 percentage point reduction in the coinsurance rate, whereas the comparable DD estimate of later children is not significant. In firstborn children, decision to see a doctor may be influenced by out-of-pocket since first-time parents would be less confident of their subjective evaluation of their child's health.

Next, the results on the overall health care utilization are presented, which is the sum of outpatient care, inpatient care, dental care and drug prescriptions. Given that cost sharing for outpatient services could reduce utilization of complementary services such as hospitalization[20] and emergency room (ER) visits[21], it is meaningful to examine its impact on overall health care costs.

In this subgroup analysis, I can also determine the types of outpatient clinics that a child visits, as insurance claims data contain information regarding the specialty of clinics such as pediatric care and dermatology. Although parents generally take their children to clinics that provide pediatric services, they are free to choose clinics offering specialized care (such as dermatology and orthopedics) when the symptoms of their children appear to require such care[22]. Regardless of whether care is sought at pediatric clinics or clinics for other specialties, out-of-pocket expenditure of parents remains the same. In this subsection, outpatient care is divided into four clinical specialties: pediatric care, internal medicine which specializes in primary care mainly for adults, dermatology, and other services.

[19] Bessho (2012) also finds that school-age girls are more elastic to cost sharing than boys, although his analysis is based on cross-sectional comparison.

[20] Chandra et al (2010) and Karaca-Mandic et al (2012) show significant offsets in the elderly population and in children with asthma in the United States. In particular, Karaca-Mandic et al (2012) investigate the relationship between out-of-pocket medication costs and asthma-related hospitalization, using insurance claims for children with asthma who had commenced asthma control therapy between 1997 and 2007 in the United States.

[21] Miller (2012a) argues that non-urgent ER visits would decrease if patients can afford to receive adequate outpatient treatments before the sudden deterioration of health, as seen in reduced ER usage after the Massachusetts Health Care Reform. Kolstad and Kowalski (2012) also find significant reduction in ER utilization and length of hospital stay in an examination of the same reform as Miller (2012a). In addition, Miller (2012b) obtains similar results using a sample of children under 18 years of age.

[22] In Japanese clinical settings, pediatric clinics provide primary care treatment for a variety of children's conditions such as the common cold, sore throat, diarrhea, asthma, vomiting, dry skin, and influenza. On the other hand, other clinical departments deal mainly with adults, but also treat children.

In order to ascertain the robustness of the results in this section, the same equation on the number of visits is estimated with Quasi-ML Poisson regression as in Finkelstein et al (2012) because the outcome variable takes non-negative integer values. By applying the Quasi-ML Poisson model, I can control the individual fixed effects and calculate clustered standard errors of the households in order to address the over-rejection of the null hypothesis. These additional results are presented in the Appendix.

Table 5 summarizes the results from OLS regressions. Column (1) presents the results of total health care services, which consists of hospitalization, outpatient care, dental care and drug prescription[23]. DD estimates suggest that cost sharing in outpatient care increases the total number of days that a patient utilizes health care services by 0.71 days per month. This implies that a subsidy for the utilization of outpatient care increases the total health care cost even if it might reduce the hospitalization and other health care services that complement outpatient care; however, I am unable to definitively show the results of hospitalization due to the small sample size[24].

Column (2) reports the results of pediatric care. All equations produced non-significant estimations, indicating that pediatric care for children is not responsive to cost sharing. The results are similar in Column (3). Dermatological care, dental care, and drug prescription, however, are utilized with relatively high elasticity. DD estimates in Column (4) show that the number of visits increased by 0.1 days per month. Moreover, DD estimates in Columns (6) and (7) show that the number of visits increases by 0.22 days for drug prescription. Regarding the "other" outpatient services, I find that the DD estimate for the number of visits is positive and significant in the Quasi-ML Poisson regression, as shown in the Appendix B. These results indicate that the utilization of health care services other than pediatrics and dental care services is elastic to cost sharing. In particular, it is likely that treatments for allergic diseases such as allergic dermatitis and hay fever are elastic to out-of-pocket expenditures because these diseases probably account for a large share of children's visits to dermatological clinics and "other" health care services. Furthermore, the expansion of the MSCI had a positive impact on drug prescription. Given that the common cold and fever are the major reasons for pediatrician visits, my results can be interpreted to show that children's health care demand is inelastic if they present with clear symptoms of illness. Nevertheless, the clinical symptoms of some diseases such as allergic dermatitis and nasal inflammation are sometimes unclear and the decision on whether to see a doctor tends to rely on the parents' subjective assessment of their child's health conditions. Hence, the co-payment policy can be an important factor in influencing health care utilization. After all, the response of the demand for health care for children to cost sharing varies across a variety of services, as shown in Duarte (2012).

[23]Insurance claims data do not contain information on the use or non-use of emergency room services.

[24]The data in this paper include 53 hospital episodes, which is too small a number to examine the determinants of spending per hospitalization. With regard to the probability for hospital admission, my findings do not show any significant impact of the October 2004 reform.

6 Discussion

In the face of an increasing child poverty rate and a declining birth rate, Japan has reduced patient cost sharing in the last two decades to support households with children and improve access to public health care services especially for underprivileged children, led by the initiative of local governments. In addition to the standard coinsurance rate set by the national government, many municipalities subsidize the co-payment of children's health care utilization on a discretionary basis. This paper examines the consequences of this subsidy program, named the Medical Subsidy for Children and Infants.

By exploiting a recent institutional change in MSCI in Hokkaido Prefecture, the impact of cost sharing on health care utilization in low-income children is identified by comparing pre-post changes of health care utilization between low-income children (treatment group) and middle- or higher- income children (control group). The implied arc price elasticity of outpatient health care demand is -0.16, which is almost the same as the value presented in the RAND HIE. Furthermore, it is particularly noteworthy that the estimated elasticity of low-income children in this paper (-0.16) is similar to those estimated in higher-income populations in other settings. This finding contradicts the well-known hypothesis that low-income individuals are more affected by changes in co-payment than higher-income individuals. Instead, my paper is consistent with the findings of Chandra et al (2014), in that low-income patients are very similar with their higher-income counterparts with regard to sensitivity to cost sharing.

In addition, my findings show that price elasticities are relatively heterogeneous across different characteristics of children and a variety of health care services. For example, among low-income children, health care demand for firstborns and girls are responsive to cost sharing, but subsequent children and boys are not. I also find that the demand for dermatology and drug prescription are highly elastic to the reduction of the coinsurance rate. These heterogeneous effects of cost sharing indicate that the consequences of recent drastic expansions of MSCI have not been identical across subpopulations and clinical departments. Similarly, the heterogeneity of the impact suggests that some populations would not respond to reduced cost sharing even if policymakers attempt to improve their access to health care through the elimination of financial pressure. In other words, the benefit of low cost sharing does not reach all target populations equally. If this is the case, cost sharing on its own would not be an adequate policy tool to enhance the access to health care services in underprivileged children.

References

Apouey B, Geoffard PY (2013) Family income and child health in the uk. Journal of Health Economics (forthcoming)

Baicker K, Goldman D (2011) Patient cost-sharing and healthcare spending growth. Journal of Economic Perspectives 25(2):47–68

Bertrand M, Duflo E, Mullainathan S (2004) How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics 119(1):249–275

Bessho Si (2012) Medical care subsidies for, hospital visits of,and health status of children. Quarterly of Social Security Research (in Japanese) 47(4):413–430

Case A, Lubotsky D, Paxson C (2002) Economic status and health in childhood: The origins of the gradient. American Economic Review 92(5):1308–1334

Chandra A, Gruber J, McKnight R (2010) Patient cost-sharing and hospitalization offsets in the elderly. American Economic Review 100(1):193–213

Chandra A, Gruber J, McKnight R (2014) The impact of patient cost-sharing on low-income populations: Evidence from massachusetts. Journal of Health Economics 33(C):57–66

Chen Y, Jin GZ (2012) Does health insurance coverage lead to better health and educational outcomes? evidence from rural china. Journal of Health Economics 31(1):1 – 14

Chernew M, Gibson T, Yu-Isenberg K, Sokol M, Rosen A, Fendrick A (2008) Effects of increased patient cost sharing on socioeconomic disparities in health care. Journal of General Internal Medicine 23(8):1131–1136

Chiappori PA, Durand F, Geoffard PY (1998) Moral hazard and the demand for physician services: First lessons from a french natural experiment. European Economic Review 42(3-5):499–511

Cockx B, Brasseur C (2003) The demand for physician services: Evidence from a natural experiment. Journal of Health Economics 22(6):881–913

Currie J, Gruber J (1996a) Health insurance eligibility, utilization of medical care, and child health. The Quarterly Journal of Economics 111(2):431–66

Currie J, Gruber J (1996b) Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. Journal of Political Economy 104(6):1263–96

Dafny L, Gruber J (2005) Public insurance and child hospitalizations: access and efficiency effects. Journal of Public Economics 89(1):109–129

Duan N, Manning J Willard G, Morris CN, Newhouse JP (1983) A comparison of alternative models for the demand for medical care. Journal of Business and Economic Statistics 1(2):115–126

Duarte F (2012) Price elasticity of expenditure across health care services. Journal of Health Economics 31(6):824–841

Finkelstein A (2007) The aggregate effects of health insurance: Evidence from the introduction of medicare. The Quarterly Journal of Economics 122(1):1–37

Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, Allen H, Baicker K, the Oregon Health Study Group (2012) The oregon health insurance experiment: Evidence from the first year. The Quarterly Journal of Economics

Goldman DP, Joyce GF, Zheng Y (2007) Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. JAMA 298(1):61–69

Ministry of Health L, Welfare (2013) Survey on the medical subsidy for children and infants, URL http://www.mhlw.go.jp/stf/houdou/2r9852000002xx3m.html, (Accessed on March 5, 2014)

Ikegami N, Yoo BK, Hashimoto H, Matsumoto M, Ogata H, Babazono A, Watanabe R, Shibuya K, Yang BM, Reich MR, Kobayashi Y (2011) Japanese universal health coverage: evolution, achievements, and challenges. The Lancet 378(9796):1106 – 1115

Kan M, Suzuki W (2010) Effects of cost sharing on the demand for physician services in japan: Evidence from a natural experiment. Japan and the World Economy 22(1):1–12

Karaca-Mandic P, Jena AB, Joyce GF, Goldman DP (2012) Out-of-pocket medication costs and use of medications and health care services among children with asthma. JAMA 307(12):1284–1291

Kolstad JT, Kowalski AE (2012) The impact of health care reform on hospital and preventive care: Evidence from massachusetts. Journal of Public Economics 96(12):909 – 929

Kondo A, Shigeoka H (2013) Effects of universal health insurance on health care utilization, and supply-side responses: Evidence from japan. Journal of Public Economics 99(0):1 – 23

Leibowitz A, Manning WG, Keeler EB, Duan N, Lohr KN, Newhouse JP (1985) The Effect of cost-sharing on the use of medical services by children: interim results from a randomized controlled trial. Rand Report, Rand

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A, Marquis S (1987) Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review 77(3):251–77

Miller S (2012a) The effect of insurance on emergency room visits: An analysis of the 2006 massachusetts health reform. Journal of Public Economics 96(11):893–908

Miller S (2012b) The impact of the massachusetts health care reform on health care use among children. American Economic Review 102(3):502–07

Newhouse J (1993) Free for All: Lessons from the RAND Health Insurance Experi-ment. Harvard University Press, Cambridge, MA.

Nishikawa M (2010) Income restriction and individual payment in the infant medical-expense support grant. Aoyama Journal of Economics (in Japanese) 62(3):195 – 214

Nishikawa M (2011) Spatial relation among municipal decision making about infant medical-expense support. Aoyama Journal of Economics (in Japanese) 62(4):87 – 111

OECD (2012) Oecd family database: Child well-being module, URL http://www.oecd.org/els/family/oecdfamilydatabasechildwell-beingmodule.htm, (Accessed on March 5, 2014.)

Shigeoka H (2013) The effect of patient cost-sharing on utilization, health and risk protection. American Economic Review (Forthcoming)

Swartz K (2010) Cost-sharing: Effects on spending and outcomes. Robert Wood Johnson Foundation Synthesis Project Report No. 20

Valdez RB, Brook RH, Rogers WH, Ware JE, Keeler EB, Sherbourne CA, Lohr KN, Goldberg GA, Camp P, Newhouse JP (1985) Consequences of cost-sharing for children's health. Pediatrics 75(5):952–961

Van De Voorde C, Van Doorslaer E, Schokkaert E (2001) Effects of cost sharing on physician utilization under favourable conditions for supplier-induced demand. Health Economics 10(5):457–471

Winkelmann R (2004) Co-payments for prescription drugs and the demand for doctor visits - evidence from a natural experiment. Health Economics 13(11):1081–1089

Wooldridge JM (2002) Econometric Analysis of Cross Section and Panel Data. The MIT Press

Zweifel P, Manning WG (2000) Moral hazard and consumer incentives in health care. In: Culyer AJ, Newhouse JP (eds) Handbook of Health Economics, Handbook of Health Economics, vol 1, chap 8, pp 409–459

Panel A. Average Number of Outpatient Visits

Panel B. Log Spendings per Visit

Note: These figures plot the 6-month averages of the numbers of outpatient visits (Panel A) and log spendings per visit (Panel B) by subgroups. The values for October 2004 are the mean values from October 2004 until March 2005. Diamond-shaped markers represent the averages of low-income children, while circle-shaped markers represent those of higher-income children. The vertical line denotes October 2004, in which the reform was implemented.

Figure 1: Difference-in-Differences Analysis

Table 1: Changes in Coinsurance Rates in the October 2004 Reform

	Low-Income		Higher-Income	
	Before	After	Before	After
Children of Married Couples				
Doctor Visits	30%	0%	30%	10%
Hospitalization	0%	0%	10%	10%

Note: Low-income children are defined as those whose parents are exempt from paying residence-based taxes; the remaining children are classified as middle- or higher-income children. The first visit to the doctor is free in Y city, although an initial deductible of 580 JPY is required in almost all other municipalities in Hokkaido Prefecture.

Table 2: Descriptive Statistics

	Low-Income				Higher-Income			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
				Prob. of Visits				
Total Outpatient Care	0.48	0.50	0	1	0.52	0.50	0	1
Pediatrics	0.31	0.46	0	1	0.35	0.48	0	1
Internal Medicine	0.13	0.34	0	1	0.10	0.29	0	1
Dermatology	0.08	0.27	0	1	0.08	0.27	0	1
Other	0.17	0.38	0	1	0.10	0.29	0	1
Dental Care	0.10	0.29	0	1	0.15	0.35	0	1
Drug Prescription	0.14	0.34	0	1	0.19	0.39	0	1
Total Health Care	0.53	0.50	0	1	0.59	0.49	0	1
				N. of Visits				
Total Outpatient Care	1.58	2.61	0	18	1.23	1.72	0	18
Pediatrics	0.78	1.56	0	11	0.80	1.47	0	18
Internal Medicine	0.25	0.85	0	12	0.14	0.51	0	6
Dermatology	0.10	0.38	0	4	0.12	0.50	0	7
Other	0.45	1.24	0	9	0.17	0.62	0	8
Dental Care	0.22	0.83	0	7	0.34	1.04	0	12
Drug Prescription	0.27	0.83	0	6	0.44	1.20	0	16
Total Health Care	2.13	3.37	0	22	2.09	3.00	0	31
				Ln (Spending / visit)				
Total Outpatient Care	6.16	0.43	4.62	7.72	6.13	0.46	3.76	7.98
Pediatrics	6.12	0.51	3.37	7.52	6.13	0.49	4.14	7.98
Internal Medicine	6.22	0.30	5.53	7.00	6.23	0.33	4.93	7.63
Dermatology	5.99	0.33	4.47	6.55	5.94	0.40	4.48	7.13
Other	5.97	0.90	1.25	7.72	6.05	0.74	1.95	7.84
Dental Care	5.97	0.66	4.20	7.65	6.17	0.63	3.91	8.26
Drug Prescription	5.57	0.47	4.57	6.70	5.69	0.45	4.55	7.03
Total Health Care	6.14	0.50	4.37	8.47	6.15	0.53	3.91	8.89
				Characteristics of Children				
Female	0.38	0.49	0	1	0.49	0.50	0	1
Firstborn	0.58	0.49	0	1	0.42	0.49	0	1
Number of Observations	858				3,734			
Number of groups	39				142			
Mean Follow-up Months	22.00				25.06			

Note: This table summarizes outpatient medical utilization and the characteristics of the sample. The data cover 36 months from April 2003 to March 2006. Low-income children are defined as those whose parents are exempt from paying residence-based taxes; the remaining children are classified as middle- or higher-income children. "Other" denotes the medical utilization of outpatient care excluding internal medicine, pediatrics, and dermatology.

Table 3: Main Results on Outpatient Health Care Utilization

	Prob. of Doctor Visits		N. of Doctor Visits		Ln (Spending / Visit)	
	Month	Pre/Post	Month	Pre/Post	Month	Pre/Post
	(1)	(2)	(3)	(4)	(5)	(6)
DD	0.1086**	0.0157	0.4904**	0.4240**	-0.0686	-0.1512**
	[0.048]	[0.048]	[0.206]	[0.198]	[0.060]	[0.075]
N. of children	181	181	181	181	181	181
Obs.	4,592	355	4,592	355	2,352	337
R Squared	0.033	0.024	0.031	0.050	0.048	0.166

Note: This table presents the estimated impact of the expansion of the MSCI according to the characteristics of children by employing DD technique. The sample includes the children of married couples and covers the period spanning 18 months before and after October 2004. Columns (1) and (2) present the results of the probability of visits. Columns (3) and (4), and (5) and (6) present the results of the number of visits and the natural log of health care spending per visit, respectively. Columns (1), (3) and (5) use the monthly raw data; the other columns use the data collapsed into pre- and post-reform periods. All equations are estimated by OLS regression with individual fixed effects. Time effects are controlled by monthly dummies in Columns (1), (3) and (5), and post-reform dummy variables in Columns (2), (4) and (6). Standard errors are clustered by household. ***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 4: Results by the Characteristics of Children

	Sex		Birth Order	
	Boy	Girl	First	Subsequent
	(1)	(2)	(3)	(4)
Prob. of Doctor Visits	0.0820*	0.1712*	0.1358*	0.0928*
	[0.041]	[0.096]	[0.075]	[0.053]
Obs.	2,453	2,139	2,064	2,528
N. of Doctor Visits	0.3112*	0.8276**	0.8602**	0.1384
	[0.183]	[0.414]	[0.343]	[0.285]
Obs.	2,453	2,139	2,064	2,528
Ln (Spending/ Visit)	0.0177	-0.1698**	-0.0903	-0.0961
	[0.067]	[0.083]	[0.075]	[0.092]
Obs.	1,215	1,137	1,114	1,238

Note: The sample used in the estimation of the results in this table includes all the children of married couples, and covers the period spanning 18 months before and after October 2004. Columns (1) and (2) divide the sample by sex. Column (3) includes only firstborn children and Column (4) includes subsequent children. All equations are estimated by OLS regression with individual fixed effects and monthly dummies. Standard errors are clustered by household. ***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 5: Results by Clinical Department

		Outpatient Care					
	Total	Pediatrics	Internal Medicine	Dermatology	Other	Dental Care	Drug
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prob. of Doctor Visits	0.0984**	0.041	0.0573	0.0629***	0.021	0.0103	0.0833**
	[0.043]	[0.049]	[0.046]	[0.020]	[0.040]	[0.030]	[0.034]
Obs.	4,592	4,592	4,592	4,592	4,592	4,592	4,592
N. of Doctor Visits	0.7108**	-0.0443	0.2057	0.1029***	0.0148	0.039	0.2182**
	[0.317]	[0.199]	[0.129]	[0.035]	[0.069]	[0.083]	[0.091]
Obs.	4,592	4,592	4,592	4,592	4,592	4,592	4,592
Ln (Spending/ Visit)	-0.043	-0.1137	-0.1156	-0.0619	-0.1508	0.4114*	-0.0524
	[0.057]	[0.137]	[0.109]	[0.059]	[0.108]	[0.246]	[0.121]
Obs.	2,665	1,570	469	363	505	626	817

Note: This table presents the estimated impact of the expansion of the MSCI according to clinical department by employing DD technique. The sample includes all the children of married couples, and covers the period spanning 18 months before and after October 2004. "Other" denotes the medical utilization of outpatient care excluding internal medicine, pediatrics and dermatology. All equations are estimated by OLS regression with individual fixed effects and monthly dummies. Standard errors are clustered by household. ***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

A Health Insurance Coverage for Children in Japan

In Japan, children are covered under the same health insurance plan as their designated household head. Broadly, there are three types of health insurance plans for working-age adults in Japan: society-managed insurance; a health insurance plan managed by the Japan Health Insurance Association (JHIA); and Citizens' Health Insurance (CHI), which is a residence-based health insurance plan. These three plans account for almost 90%[25] of health insurance for those under 75 years of age. Adults who work for large firms participate in society-managed insurance, whereas those who work for small and medium enterprises are included in health insurance programs managed by the JHIA. Other adults must obtain coverage through the CHI in their residential area[26]. Hence, children covered under CHI are the children of employees of small firms (with fewer than five employees), self-employed workers in the agricultural and retail/service sectors, or the unemployed. In general, children covered under CHI are from families with lower household incomes than those of children whose parents are covered under the other types of insurance. In addition, CHI enrolls those who are unemployed and are therefore no longer covered under employment-based plans. Accordingly, their children may also be shifted from employment-based plans to CHI coverage due to loss of employment of the household head. Although we do not know the exact volume of changes in health insurance plans due to employment transitions, this may be a reason why the financial situation of persons covered under CHI plans can be considered more fragile than those covered under employment-based plans.

B Quasi-ML Poisson Regression

To verify my central results from OLS regressions in Tables 3 and 4, I present the results of the number of visits with another specification. Here, a Quasi-ML Poisson regression is utilized in order to address the skewed nature of the outcome, as in Finkelstein et al (2012). Quasi-ML Poisson regression provides consistent estimates under the relatively weak assumption that the conditional mean is correctly specified (Wooldridge, 2002). Standard errors are clustered by household as well as OLS estimates in this paper to deal with the underestimation of standard errors.

The DD estimates and implied incident rate ratios (IRR) are presented in Table 6. The DD estimate on the total outpatient expenditure is 0.3445 and significant at the 95% confidence interval. The IRRs imply that the number of outpatient visits increases by about 40% in the treatment group. This impact is consistent with the result from OLS analysis[27]. For the estimates by clinical specialty, the results suggest that cost sharing does not affect the number of visits to pediatric and dental services, while other types of care are utilized with high elasticity.

[25] The remaining 10% is included in Mutual Aid Associations that cover those employed in the public sector.

[26] A comprehensive and historical review of the Japanese health care system is provided in Ikegami et al (2011) and other articles published in *The Lancet*, August 30, 2011.

[27] The average number of outpatient visits in low-income children was 1.36 before October 2004. Given that the estimated impact of the reform is 0.4904 in (Table 3), the implied rate of increase is 36% (0.49/1.36), which is almost the same as the IRR from Quasi-ML Poisson regression.

Table 6: Results from Quasi-ML Poisson Regression

		Outpatient Care						
	Total	Total	Pediatrics	Internal Medicine	Dermatology	Other	Dental Care	Drug
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coefficient	0.3294**	0.3445**	-0.0664	0.7834	0.8683**	0.5658**	0.2186	0.6391**
	[0.155]	[0.145]	[0.201]	[0.618]	[0.371]	[0.285]	[0.317]	[0.282]
IRR	1.390	1.411	0.936	2.189	2.383	1.761	1.244	1.895
Obs.	4,592	4,592	4,240	2,940	3,014	3,016	3,118	3,463

Note: This table presents the DD and DDD estimates of the number of visits using Quasi-ML Poisson regression. The sample includes all the children of married couples, and covers the period spanning 18 months before and after October 2004. "Other" denotes the medical utilization of outpatient care excluding internal medicine, pediatrics and dermatology. All equations include individual fixed effects and monthly dummies. Standard errors are clustered by household. ***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Chapter 4

Does Reduced Patient Cost-sharing Improve Child Health ?*

Reo TAKAKU†

Abstract

Although cost-sharing has been widely used in many developed countries to contain health care expenditure, the effect of cost-sharing on health for the general population has not been elucidated. In particular, there are few studies on the effect of cost-sharing on children's health. This paper investigates whether reduced cost-sharing leads to an improvement of health status among preschool and school-age children in Japan, exploiting regional disparities in the expansions of municipality-level subsidy programs for out-of-pocket expenditure. With the eligibility for this subsidy program, known as the Medical Subsidy for Children and Infants (MSCI), the coinsurance rate generally decreases from 30 or 20 percent to 0 percent for outpatient health care services and drug prescriptions. In order to uncover the impact of this program, I conducted an original survey for all municipalities to understand time-series evolution of the eligible age for the MSCI. The response rate is 55 percent, but it covers 75 percent of the population under fifteen years old. The probability of being eligible for the MSCI was calculated by the age, prefecture where a child lives, and year. These probabilities were matched to children's health data from the Comprehensive Survey of Living Conditions from 1995 to 2010. The results show that eligibility for the MSCI improves subjective measures of health status among preschool children. Numerically, MSCI eligibility decreases the probability of having any symptoms by 2.8 percentage points among preschool children. However, I find no such improvement among school-age children. In addition, MSCI eligibility does not reduce hospitalization either among preschool or school-age children. Taken together, this paper suggests that extensions of the MSCI beyond preschool age have no health benefits. Finally, all the results hold after inclusion of a variety of covariates and subsample analysis that excludes prefectures where the response rate for the author's original survey is low.

Keywords : cost sharing, children's health, subjective symptoms, hospitalization, Japan

JEL classification : I10, H75

*This paper was supported by a Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (Research Project Number: 25780199). The views expressed in this paper do not necessarily reflect the views of the JSPS and Institute for Health Economics and Policy. I am grateful for comments and suggestions from Yoko Ibuka at the meeting of the Japanese Economic Association and Miao Jiang at the annual meeting of the International Health Economic Association. I also thank Michihito Ando, Shun-ichiro Bessho, Yoshikazu Kenjoh and Atsuhiro Yamada for providing many comments and suggestions. All errors are my own.

†Institute for Health Economics and Policy. Email:reo.takaku@ihep.jp

1 Introduction

In the countries where universal health coverage has been already implemented, use of patient cost sharing is a primary tool to contain health care expenditure. In fact, many high income countries increased their level of patient cost sharing between 2000 and 2010 (Johnson et al, 2013). However our knowledge of the effect of such broad use of cost-sharing policy is not sufficient. In particular, several literature reviews consistently point out that the impact of patient cost sharing on health is still uncertain (Baicker and Goldman, 2011; Goldman et al, 2007; Kiil and Houlberg, 2014; Mann et al, 2014). Thus far, most of what we have ever known is from the RAND Health Insurance Experiment (Manning et al, 1987), which shows there was little difference in general health across various cost-sharing plans, but the experimental results from 30 years ago are not directly applicable to the situation today, especially to countries other than the United States. In addition, there are few studies on the impact of patient cost sharing on children's health. Since childhood health status is widely recognized as an important determinant for future achievement and health (Case and Paxson, 2011; Currie, 2009; Currie and Staible, 2006; Fletcher et al, 2010), making a new reliable evaluation is of particular importance.

With the fulfillment of these needs as a motivation, this paper presents new evidence on the impact of cost-sharing on children's health, exploiting the recent local-level policy changes in Japan. In Japan, the national government sets the coinsurance rate for preschool children at 20 percent everywhere. However, local governments can reduce this amount at their own financial expense. This subsidy program, named Medical Subsidy for Children and Infants (MSCI, in Japanese: *Nyūyōji iryōhi josei*), has been dramatically expanded in the last decade. Although the eligible age for the MSCI was under two years old in 97 percent of municipalities (sub-prefectural government) in 1998, roughly 50 percent of the municipalities expanded the eligibility to elementary-school-age children in 2010. With eligibility for the MSCI, the coinsurance rate decreases to 0 percent or a very small amount in most municipalities. Regional diversity is also extensive. For instance, municipalities in the Kanto region, which includes Tokyo prefecture, the capital of Japan, have rapidly extended the eligible age to over school-age since 2007, but many municipalities in the Kinki region, which is the second largest economic area in Japan, had restricted MSCI eligibility to preschool children by 2010.

The extensive regional diversity in the MSCI eligible ages provides plausible quasi experiments that may uncover the causal effect of reduced cost-sharing on health. However, there have been few studies on the MSCI because no public organization had compiled comprehensive data on the MSCI until 2011. Although some local governments publish fragmented information on the MSCI in their districts, it had been impossible to assemble all the data. To overcome this difficulty, I conducted an original survey for all municipalities to compile the precise system and time-series evolution of the MSCI from 1995 to 2012. The overall response rate is 55%, but the survey covers 75 percent of the population under fifteen years old since the response rate is higher in urban areas. Aggregating them at the prefecture level, the probabilities of being eligible for the MSCI were calculated for every child in a given age, year, and prefecture. In the empirical analysis, these probabilities were matched to children's health data. The data used here are the Comprehensive Survey for Living Conditions (CSLC), which is a nationally representative sample of the Japanese population.

On the effects on children's health, two findings are presented. First, the MSCI expansions are associated with the improvement of subjective measures of health status among preschool children. In addition, it is also associated with improvements of a variety of subjective symptoms such as "cough" and "stuffy nose" among preschool children. Numerically, the probability of having any symptoms decreases by 2.8 percentage points if the child is eligible for the MSCI, while I find no such an improvements among elementary school children. Second, regardless of significant effects on subjective measures of health status among preschool children, the effect on the probability of hospitalization is not significant among all ages. Moreover, I find no effect on the severe subjective symptoms such as role limitation due to health. Taken together, these results suggest that the MSCI for preschool children improves health status to a moderate extent, but to expand the eligible age far beyond the preschool age has no health benefits. Finally, all the results hold after a variety of robustness checks such as inclusion of additional covariates and subsample analysis that excludes prefectures with a low response rate in the author's original survey.

The rest of the paper is organized as follows. The next section provides a brief overview of previous studies. Recent expansions of the MSCI and the explicit system are explained in Section 3. Section 4 describes the analytical framework and specifies econometric models. Section 5 provides description of the data. Section 6 summarizes the main results. Finally, section 7 provides conclusions.

2 Prior Literature

Before introducing the study strategy and results, a brief literature review of the impact of cost sharing on health is presented. On this issue, many studies focus on the impact of patient cost sharing for outpatient care and prescription drugs on later hospitalization and ER visits. As is mentioned in the literature review by Goldman et al (2007), previous studies find that increasing (decreasing) patient cost sharing for prescription drug is associated with an increase (decrease) of later adverse events among patients with diabetes (Mahoney, 2005), schizophrenia (Zeber et al, 2007), asthma (Jonathan et al, 2011; Karaca-Mandic et al, 2012), acute myocardial infarction (Rahimi et al, 2007), hypertension (Atella et al, 2006), or high cholesterol levels (Goldman et al, 2006), although some studies find no effect on hospitalization among patients with acute myocardial infarction(Pilote et al, 2002) or depression (Wang et al, 2010). Studies on the elderly population are likely to find a significant substitution effect of the changes in cost-sharing in prescription drugs and outpatient care. (Chandra et al, 2010; Mojtabai and Olfson, 2003; Tamblyn et al, 2001; Trivedi et al, 2010) because of the high prevalence of chronic conditions among this population. In particular, Chandra et al (2010) investigates the effect of a 10% increase in the co-insurance rate on later hospitalization among retired public employees in California, showing hospitalization significantly increased after the reduction of drug prescriptions due to increased cost-sharing. Although there are limited studies on children, Karaca-Mandic et al (2012) also finds similar consequences among children with asthma.[1]

On the other hand, these observational studies may suffer from several biases such as selection bias (e.g. patients with bad health condition are likely to move into generous cost-sharing plans) and omitted variable bias (e.g. unobservable factors affect the control and treatment groups together). Hence, results from randomized controlled trials are of particular importance even if their relevancy to today's environmental background is limited. Again, the most notable and important contribution on this issue is from the RAND HIE (Newhouse, 1993), which shows no causal effect of higher patient cost-sharing on later increases in health care expenditures. In addition, the RAND HIE also suggests that cost sharing has no significant impact on children's subjective and physiologic measures of health (Valdez et al, 1985), although it would increase health care costs.[2] Valdez et al (1985) reports the effect on a variety of outcomes ranging from subjective health to anemia, hay fever, and hearing loss, but they do not find significant differences between free and cost-sharing plans.[3]

In addition, Taubman et al (2014) examine the effect of Medicaid eligibility on emergency department utilization in the Oregon Health Insurance Experiment (Oregon HIE) and show that Medicaid eligibility increases ER utilization, rejecting the view that generous health insurance could reduce emergency-department use and perhaps even total health care costs through improved access to primary care. Given that many observational studies find the opposite, results from randomization (Taubman et al, 2014) are of particular importance.

Among studies of the Japanese population, two recent papers have examined the effect of patient cost-sharing on health in the elderly population (Nishi et al, 2012; Shigeoka, 2014). Nishi et al (2012) show that reduction of cost-sharing improves a 24-point scale based on the Kessler-6 instrument for nonspecific psychological distress through reduced out-of-pocket costs, exploiting a discontinuous change in the coinsurance rate from 30 percent to 10 percent at the age of 70. Based on a similar research design, Shigeoka (2014) finds no improvement in subjective health and mortality, although health care utilization significantly increases. However, the relevancy of these results to the present study may be limited because determinants of health in the elderly population are greatly different from those of children because chronic diseases are not common among children and the bulk of their outpatient visits are for acute conditions.

Finally, Bessho (2012) investigates the effects of MSCI eligibility on children's health, using the 2008 CSLC. However the strategy exploited in Bessho (2012) is based on a cross-sectional framework and certainly suffers from several biases. In short, the purpose of this paper is to enhance the validity of Bessho (2012) with a repeated cross-sectional framework that exploits the large regional differences in the dynamics of

[1] In the context of the effect of health insurance coverage, some studies observe reduction of children's hospitalization after enrollment in Medicaid. For instance, Dafny and Gruber (2005) shows the increase of avoidable hospitalization due to Medicaid expansions have been mild compared with that in overall hospitalization, indicating efficacy effects of Medicaid. Aizer (2007) also shows enrollment in Medicaid before deterioration of health leads to a reduction in children ' s hospitalization.

[2] RAND HIE (Newhouse, 1993) shows that the central estimate of arc price elasticity of outpatient health care demand is -0.2 and children are as responsive as adults in outpatient care utilization.

[3] Although they find significant differences only in hay fever between free and cost-sharing plans, they find no difference within cost-sharing plans.

MSCI expansions.

3 Institutional Background

3.1 Medical Subsidy for Children and Infants

In Japan, the national-level coinsurance rate is 30 percent for school-age children and 20 percent for preschool children.[4] Of course, the amount of out-of-pocket costs implied by these co-insurance rates is not low, compared with other developed countries such as the U.K. and Nordic countries, but the access to physicians is generally easy without any stringent gatekeeping. Consequently, use of public health care services may not be constrained severely, regardless of comparably high out-of-pocket costs. However, there have been persistent proponents to reduce these coinsurance rates further. The background to this movement is the declining birth rate and increasing child poverty rate. In Japan, the total fertility rate has been in decline over the last three decades, and the child poverty rate has increased since the mid-1990s (OECD, 2012). In 2010, Japan's poverty rate among children under 18 years was 15.7% [5], which was similar to those of Canada and Italy but almost double the rates of Germany and Sweden. As the importance of solving these structural problems has been widely acknowledged, it has been argued that costs for children, especially for low-income households, should be removed. Since health care costs for childhood illness are sometimes high especially during the preschool age, a reduction of coinsurance rates through the appropriation of local tax revenue has been supported in the local assembly. This local subsidy program is the MSCI.

In general, the MSCI system differs across municipalities in four aspects. First, municipalities can freely set the eligible age for the MSCI within their jurisdiction. For children older than the upper-limit age, municipalities provide no benefits. Second, municipalities can restrict eligibility for children from high-income households. Of all municipalities in 2012, 25.6% adopted household income ceilings (MHLW, 2013). Third, municipalities can also choose the reimbursement method (in-kind transfer or refund). The majority of municipalities have adopted in-kind transfers and the out-of-pocket payments are reduced immediately. By contrast, under the refund system, children (or specifically, their parents) have to pay 100% of the co-payment, and the co-payment is eventually reimbursed by the municipality. Finally, the amount of subsidy varies across municipalities. Some municipalities charge very small out-of-pocket costs for health care utilization to promote the "appropriate" use of pediatric services, whereas the co-payment is rendered completely free in 54% of all municipalities (MHLW, 2013).

It is useful to check the geographical distribution of the eligibility criteria for the MSCI, based on newly published official data from the Ministry of Health, Labour and Welfare (MHLW, 2013). In Figure 1, a map of Japan is color-coded according to the eligible age for the MSCI for outpatient care in each municipality. Light pink areas, which are concentrated in the northeast and southwest regions, indicate municipalities where the MSCI has been expanded to include all preschool children. In contrast, most of the municipalities

[4] The coinsurance rate for preschool children was reduced from 30% to 20% in 2008 by an initiative of the national government.

[5] According to the Comprehensive Survey for Living Conditions in 2012 (MHLW, 2014), This rate increased to 16.3% in 2012.

in the central main island of Honshu are shown in light green, indicating expansions of the MSCI to include children up to 15 years of age. Differences in eligible age are large even among regions of similar income level. For instance, municipalities in the Kanto area extend the eligible age to above preschool age but those in Kinki region, the second largest economic region in Japan, do not extend the MSCI to school-age children. These extensive regional disparities in eligible age provide plausible quasi-experiments for uncovering the effect of reduced cost-sharing on children's health.

3.2 Historical Evolution of MSCI

Although the maximum age of eligibility for the MSCI has been dramatically raised in the last decade[6], there have been no comprehensive data on the precise system of MSCI before 2011. The MHLW published the eligibility criteria of all municipalities in 2013 (MHLW, 2013)[7], but the historical evolution of the MSCI in each municipality has not been revealed. This is because the national government has no incentive to grasp the precise system since the municipalities and prefectures have complete discretion on the MSCI system. To compensate for this shortcoming, I conducted an original survey to gain an accurate understanding of the MSCI system in all municipalities.[8]The survey consists of the following four major questions on the MSCI from 1995 to 2012: (1) eligible age for hospitalization, (2) eligible age for outpatient visit, (3) co-payment design for outpatient care utilization, and (4) reimbursement methods. Although the overall response rate is 55% [9], it covers 75% of the population under fifteen years old since the response rate is higher in large municipalities.[10]

A summary of the survey is presented in Table 1. Table 1 shows the average eligible age for the MSCI in all prefectures from 1995 to 2010, which is calculated as the population-weighted average in the municipalities that responded to the survey. The right column in this table reports population-weighted response rates. Although the average eligible ages[11] may be inaccurate for prefectures where the response rate is low, the response rates seems to be enough high to capture regional trends of the MSCI expansions; there are only two prefectures where the weighted response rate is below 50%.

Table 1 also reveals that the timing of expansions of MSCI greatly differs across prefectures. For instance, Tokyo prefecture, the capital of Japan, has preceded the other prefectures. The average eligible age for the MSCI in Tokyo prefecture was 11.1 years in 2007, which was the highest. In addition, prefectures around Tokyo also have extended the eligible age rapidly since 2007. Saitama, Gunma, and Tochigi prefecture in the Kanto region had raised the average eligible age to over ten years old by 2010. However, prefectures in

[6]For the expansions of MSCI in Tokyo Prefecture, see Nishikawa (2010) and Nishikawa (2011)

[7]Since reports are published two years after the data are collected, the data published 2013 are based on the system in 2011.

[8]The survey was held in October 2013. The questionnaire was sent to 1732 municipalities, excluding eight municipalities in Fukushima prefecture because of the consideration of damage in the Great East Japan Earthquake. For the municipalities that did not reply, we followed up with letters in December 2013.

[9]Number of responses is 949.

[10]The crude response rate is 69.8% in cities and 43.2% in towns. With municipality-level population data as of 2010, I calculated the share of children who are covered by my survey.

[11]For municipalities where the eligible age is set as "preschool" age, six is used for the calculation of average eligible age since children enroll in elementary school in April at the age of six.

Kinki region such as Kyoto and Osaka did not raise the eligible age over school-age children. The average eligible age was 6.3 years in Kyoto and 6.1 years in Osaka in 2010, indicating most municipalities in these prefectures restricted the eligibility for the MSCI to preschool children. Given that the Kinki region is the second-largest economic area in Japan, regional income levels seen to have a limited influence on recent expansions of MSCI. This paper exploits such regional differences in the timing of the expansions of MSCI to uncover the association between cost-sharing and children's health.

4 Data Description

4.1 Data Source

This study utilizes one of the most comprehensive databases of children's health status in Japan. The Comprehensive Survey of Living Conditions is a nationally representative stratified random sample survey of the Japanese population. This survey has been conducted every three years since 1986 and there are 11 rounds available under the permission of Ministry of Health, Labour and Welfare. Among all rounds, I use six rounds from 1995 to 2010, which together cover the entire period of rapid expansions of the MSCI.

In this survey, the health questionnaire investigates the health status of children with a variety of self-reported measures on symptoms, general health, and role limitations. Since all subjective variables may be answered by parents, rather than children, they necessarily reflect the evaluation from parents and can be different from the judgment of physicians (Baker et al, 2004; Johnston et al, 2009). Regardless of the limitation, subjective measures of health status are one of the standard indexes to evaluate health status[12] and likely to provide useful prediction of objective measures of physical health (Idler and Benyamini, 1997).[13]

Although the most general measure of health status may be a well-known 5-grade subjective health measure[14], my primary focus is on subjective symptoms because the CSLC surveys the 5-grade measure only for children aged six and over. Instead, I focus on the question "Do you have any symptoms of illness or injuries currently ?" If a respondent feels any symptoms such as "fever" or "headache," he or she answers yes. Using this question, a binary variable is created. Subsequently, the effects on individual items[15] are also investigated. In addition, the CSLC investigates whether a respondent is admitted to hospital at the

[12] Subjective health measures in CSLC are also frequently studied in previous studies such as Nakamura (2013).

[13] This is the conventional rationale for the usage of subjective measures as an outcome variable. Recent studies on the Oregon and Massachusetts HIEs have provided another meaningful interpretation. Finkelstein et al (2012) and Courtemanche and Zapata (2014) both show significant and large improvement in subjective health, but Finkelstein et al (2012) suggests this increase is not explained by the objective improvement of health, since improvement in subjective health preceded the increase in preventive care utilization. These results suggest the possibility that subjective measures of health can change even when there is no change in objective health status. Supporting this interpretation, Baicker et al (2013) find significant improvement in mental health measures in the first two years of the Oregon HIE but they find no effect on objective measures such as blood pressure and level of hemoglobin. Finkelstein et al (2012) argues that reduction of financial stress for medical expenditures may explain the observed improvement of subjective health. On the other hand, this argument may not be relevant to my study since children do not care for how much they pay for the outpatient consultation and prescription drugs.

[14] To measure it, a survey respondent answers the question, "How is your current health?", and chooses the answer among 5 choices ranging from "very bad" to "excellent."

[15] The CSLC surveys 20-30 items in every wave. Among them, I choose twelve items that were surveyed with the same wording and definition from 1995 to 2010. In the analysis of individual items, the effect of MSCI expansions on the probability of suffering from "fever," "fatigue," "cough,' "headache," "wheezing," "toothache," "stuffy nose," "constipation," "diarrhea," "stomachache," "rash," and "cut" are investigated.

time of the survey. If a child has been admitted to hospital, the parents would answer affirmatively and the subsequent questions about the child would be skipped. Finally, since the CSLC surveys self-reported role limitation due to health is only for children aged six and over, the effects are evaluated for elementary school-age children.

For the demographic characteristics of household members and other socio economic conditions, I use the household questionnaire, which surveys broad household characteristics such as composition of household, housing environment, job of parents and insurance status. Since these data also investigate home ownership and number of rooms, these time-varying variables are controlled for. In addition, household income is identified with the income questionnaire. However, the main estimates are reported without controlling for household income because the income questionnaire is asked for almost 10% of the total respondents.

A limitation of the CSLC is that it provides geographic information only at the prefecture level. Although I assemble municipality-level information of the MSCI in my original survey, the municipality-level eligibility rules cannot be matched to child health data from the CSLC one-to-one since the CSLC does not contain municipality-level residential information. Beyond this problem, however, the effects of cost-sharing on health can be sufficiently revealed by exploiting regional differences in the MSCI expansions since the prefectural-border is still an important determinant in regional differences in cost-sharing, as is shown in Figure 1.

4.2 Sample Construction

In this paper, the empirical analysis is divided into elementary school-age children and preschool children. Since children in Japan enroll in elementary school in April at the age of six and graduate in March at the age of twelve, I include children aged 75-147[16] months during the field period of the CSLC for the sample of elementary school-age children. For the sample of preschool children, children aged less than 74 months and over 12 months are included. The other inclusion criteria are as follows; (1) children of a married couple, (2) age of household head and spouse from 20 to 70. Children who receive public assistance are also excluded since they are not eligible for the MSCI. The number of observations is 122,331 for preschool children, and 142,243 for school-age children.

The descriptive statistics for school-age and preschool children are summarized in Table 4. In this table, Panel A summarizes descriptive statistics on outcome variables and Panel B reports those of basic characteristics of the sample. On the other outcome variables except hospitalization, the data are observed for children who are not hospitalized. In Panel C, household income status is summarized through five binary variables. It should be noted that the sum of the means of 5 variables is not equal to 100% since households with 10 million JPY and over are excluded in the regression that controls these variables.

Of preschool children, 27% have some symptoms during the field period, but this percent decreases to 19% in elementary school-age children, suggesting that the probability of having any symptoms decreases according to age during childhood. For hospitalization, the probabilities are less than 1% for both samples.

[16] 75=6*12+3 and 147=12*12+3. The reason why “3” is added is that the field period of the CSLC is June, three months later than March.

This is because children are not identified as hospitalized if they are admitted to hospital just at the time the survey is held. For the insurance status, children are classified into two insurances; Citizens' Health Insurance (CHI) and Employment-Based Health Insurances (EBHI).[17] Of preschool children, 25% are enrolled in CHI. In addition, 68% of school-age children live in a house their parents own. To control for the health status of parents, standard 5-grade measures of subjective health status are used. These variables take 1 for "excellent" health status and 5 for "very bad" health. Finally, in Panel C, the summaries for income status are reported. Since the CSLC asks about household income only for around 10 percent of the entire respondents, the number of observations is 13,991 for preschool children and 15,077 for school-age children.

5 Identification Strategy

5.1 Analytical Framework

I begin with simple specification, assuming that the causal effect of MSCI on health can be derived from following equation,

$$H_{it} = F(Elig_{it}, X_{it}, \epsilon_{it}), \tag{1}$$

where $Elig_{it}$ is a dummy variable that takes 1 if the child i is eligible for MSCI in year t[18], X_{it} is a vector of other independent variables and ϵ_{it} is an error term. In this equation, we have consistent estimates of the effect of $Elig_{it}$ if $Cov(Elig_{it}, \epsilon_{it}) = 0$. This assumption is likely to be somewhat demanding but reasonable for the first approximation because $Elig_{it}$ mainly depends on the seemingly exogenous factor for children such as age and municipality where they live. Although residential choice is endogenous if parents with sick children immigrate to the municipalities with generous MSCIs, this incentive may not be strong since the eligibility for the MSCI is temporal in that their child must age out someday.

On the other hand, one major and fundamental limitation of my study is that we cannot observe the precise value of $Elig_{it}$ since there is no national child database in Japan, as well as the CSLC, contains municipal-level geographical information, regardless of the fact that the eligible age for the MSCI differs across municipalities. In order to compensate for this shortcoming, the probability ($Prob_{npt}$) for each child i aged n in prefecture p to be eligible for the MSCI is used as the proxy for $Elig_{it}$. The formal definition of $Prob_{npt}$ is given as follows,

[17] Broadly, there are three types of health insurance plans for children in Japan: society-managed insurance; a health insurance plan managed by the Japan Health Insurance Association (JHIA); and Citizens' Health Insurance (CHI), which is a residence-based health insurance plan. These three plans account for almost 90% of health insurance coverage for children. The remaining 10% is included in Mutual Aid Associations and health insurance plans that cover those employed in the public sector. Coverage of CHI is based on the municipality where a child lives but coverage of other insurance is based on the employment status of household head.

[18] As is mentioned previously, the explicit system of MSCI greatly differs across municipalities. For instance, there are notable differences not only in eligible age, but also in method and amount of reimbursement. Here, following Bessho (2012), I focus on the regional differences in the eligible age, ignoring the other aspects of MSCI benefits.

$$Prob_{npt} = \sum_{m=1}^{N} W_{mt} * Elig_{nmt}, \tag{2}$$

where N is the number of municipalities in the prefecture p that replied to author's original survey and $Elig_{nmt}$ is a binary variable that takes the value of 1 if a child aged n is eligible for the MSCI in municipality m in year t. W_{mt} is the population weight of municipality m, which is based on the population under fifteen years old. After inserting $Prob_{npt}$ into equation (1), the following equation is derived,

$$H_{it} = G(Prob_{npt}, X_{it}, \epsilon_{it}). \tag{3}$$

Clearly, if all municipalities in prefecture p make child i eligible for the MSCI, we obtain $Elig_{it} = Prob_{npt} = 1$. On the other hand, if all municipalities in prefecture i restrict the eligible age for the MSCI under the age of child i, we obtain $Elig_{it} = Prob_{npt} = 0$. In both cases, there is no measurement error for the use of $Prob_{npt}$. The precision of the approximation by $Prob_{npt}$ is highly dependent on disparity of the eligibility criteria within the prefecture. If the disparity is not so large, the precision of the approximation by $Prob_{npt}$ may have enough certainty, even if we cannot observe in which municipality a child lives. Otherwise, $Prob_{npt}$ may be an imprecise approximation for $Elig_{it}$.[19]

Before starting regression analysis, it is useful to check the distribution of the value of $Prob_{npt}$ in each survey year. If the value takes roughly zero or one, the bias from mis-approximation does not seem to be large. If otherwise, $Prob_{npt}$ should be regarded as inappropriate proxy for $Elig_{it}$. Figure 2 summarizes the distribution of $Prob_{npt}$among school-age children every three years from 1995 to 2010. In all figures before 2004, $Prob_{npt}$ is concentrated at 0, showing there was no school-age child who was eligible for the MSCI. After 2004, however, they have become gradually eligible for the MSCI, although most children still had a low probability to be eligible even in 2010. On the other hand, preschool children were more likely to be eligible even before 2000 (Figure 3). In particular, the eligibility was polarized especially in 1998. In 1998, some preschool children were identified to have no eligibility with complete precision ($Prob_{npt} = 0$), whereas the others were also identified to be eligible with high certainly ($Prob_{npt} = 1$). Subsequently, almost all preschool children had been eligible for the MSCI by 2007.

[19]Further discussion on the interpretation of equation (3) may be noteworthy. If $Prob_{npt}$ is interpreted as an proxy variable for $Elig_{it}$, the estimated coefficient of $Prob_{npt}$ may underestimate the impact of $Elig_{it}$ because of classical measurement error (Bessho, 2012). However, $Prob_{npt}$ can be also interpreted as an instrumental variable that addresses potential endogeneity of $Elig_{it}$ in equation (1). For instance, region (state)-level eligibility rules are usually used as instrument for endogenous eligibility status in the literature on the effects of Medicaid(Currie and Gruber, 1996b, 2001; Gross and Notowidigdo, 2011). In the same spirit, equation (3) is interpreted as a reduced-form equation that directly correlates the instrument ($Prob_{npt}$) to outcome variable.

5.2 Benchmark Econometric Specification

Based on the discussion above, a parametric model is specified to identify the impact of MSCI expansions on children's health, using repeated cross section data from 1995 to 2010. Specifically, the next equation is estimated,

$$H_{it} = \alpha_0 + \alpha_1 Prob_{npt} + Z_{it}\gamma_0 + Z_{pt}\gamma_1 + Trend_p + Pref_p + Year_t + \epsilon_{it}, \tag{4}$$

where Z_{it} is a vector of individual level control variables, Z_{pt} represents prefecture level covariates, $Trend_p$ is a prefecture-specific trend, $Pref_p$ is prefecture fixed effect, $Year_t$ is a survey year effect and ϵ_{it} is an error term.

Next, I turn to relax the assumption on the exogeneity of $Prob_{npt}$ by implementing several specification checks. First, $Prob_{npt}$ may be correlated to the error term if unobservable regional characteristics such as region-level income affect both the expansions of the MSCI and child health. To address this issue, prefecture-level covariates such as prefecture-level per capita income, unemployment rate, physician density[20] are controlled for, as well as prefecture specific trends. If the point estimate of α_1 is stable with and without these covariates, we can assume that region-level unobservable factors may not bias our estimates so seriously. Second, calculation of $Prob_{npt}$ in this paper may not be completely accurate and this may result in a measurement error problem. This is potentially a serious threat since my original survey covers only 75% of the population under fifteen years old. Then, to check the robustness for this measurement error, prefectures where the weighted response rate was low are eliminated. This subsample analysis provides a clear check for the second threat. In addition to addressing these two threats, many exogenous covariates are controlled for in order to alleviate the potential omitted variable bias. In addition to basic demographic variables such as age, gender, number of household members, number of children, age of household head and spouse, many socio-economic variables are also controlled for. Specifically, the estimation controls for working status of head and spouse, household income, health insurance of household head, home ownership, number of rooms.[21] In addition, the 5-grade subjective health status of parents was also included to control for the potential evaluation bias of parents (i.e. depressed parents may evaluate children's health status negatively.). Although the health of parents may be endogenous if they are affected by the health of their children, inclusion of parents' health status may decrease the distortion due to parental evaluation of child health.

On household income, unfortunately, the CSLC surveys household income only for about 10 percent of the respondents. Because of this limitation, statistical power substantially decreases when household income status is correctly controlled for. However, to control for income provides important robustness checks since some municipalities do not provide the MSCI for children from high income households. In

[20] Number of physicians per 100 thousand.
[21] Our survey does not include educational attainment of parents.

2011, 25.7 percent of municipalities restricted the MSCI eligibility based on such an income ceiling (MHLW, 2013). With the inclusion of household income, children who are not eligible for the MSCI because of high income would be successfully excluded. Since the criteria of this income ceiling is based on that of the Child Allowance in most municipalities, children from the household with income of 10 million JPY are not included in the analysis.[22]Five binary variables that categorize income classes into equal sizes are generated and incorporated into the benchmark specification.

6 Results

6.1 Effects on Subjective Symptom

This subsection summarizes the results on subjective symptoms. First, the results from preschool children and elementary school-aged children are presented respectively. Next, the results on individual items such as "fever" are reported.

Table 3 presents the results in the sample of preschool children. Column (1) is a basic specification that controls for only demographic characteristics. Then, Column (2) additionally controls for socio-economic characteristics such as working status of parents and home ownership. Column (3) controls for 5-grade subjective health measures of parents in order to control for the tendency to evaluate children's health status according to the health status of their parents (e.g. parents in bad health condition may evaluate their children's health status negatively). Moreover, these variables roughly control for intergenerational transmission of health status through genetic inheritance and low birth weight(Currie and Moretti, 2007). In addition to individual level controls, three prefecture-level covariates (per capita income, physician density and unemployment rate) are incorporated in Column (4). These variables may capture prefecture-level differential changes in children's health status. Column (5) presents the preferred specification that controls for prefecture specific linear trends. Column (6) and (7) check the robustness of the results in Column (5). First, Column (6) excludes the data from the prefectures where the population-adjusted response rate for my original survey is below 70 percent. Though far from completely plausible, this specification checks the robustness for the potential measurement error in the probability of being eligible for the MSCI, namely in $Prob_{it}$. The second robustness check is to control for household income. To control for household income level, 4 income class dummies are added to the specification in Column (5). Although the sample size in Column (7) is considerably smaller than the full sample results because of data limitations, this specification allows us to exclude high-income households that would not be eligible for the MSCI. As mentioned previously, in Column (7), households with a total income of 10 million JPY and over are excluded.

Turning to the results, with basic demographic variables and year and prefecture FEs, the coefficient

[22] In the case of households with a married couple and two children, they are not eligible for child allowance if their household income is over 9.6 million JPY. This threshold varies with the number of dependents and the amount of deductible income. Since the income ceiling for the MSCI is also based on those criteria, for simplicity, I assume the household income of 10 million JPY is a reasonable threshold for my analysis.

of the probability of being eligible for the MSCI is -0.024 and significant in Column (1), suggesting the probability of having symptoms decreases 2.4 percentage points if a child is eligible for the MSCI with 100 percent certainty (i.e all municipalities in his prefecture expand the eligible age for the MSCI above his age). The impact may be stable from Column (2) to (4), but the coefficient in Column (5) exhibits a slightly larger value (-0.028), suggesting elimination of omitted variable bias through controlling for prefecture-specific trends alleviates underestimation of the impact of the MSCI. On the other hand, restricting the sample to prefectures with high response rate does not change the results. Again, the coefficient of $Prob_{npt}$ is -0.028, the same as the full sample results. These results give me some confidence that measurement error in the value of $Prob_{npt}$ may not be so serious for the main results, while the response rate for my original survey on the MSCI system is unfortunately far from complete. The reason is because the reduction of attenuation bias in the sub-sample analysis would have made the point estimate of $Prob_{npt}$ considerably larger, if the attenuation bias were so severe. However, the results from sub-sample analysis and full sample analysis are very similar. Though not completely plausible, this partly supports that the bias from potential mis-measurement of $Prob_{npt}$ may not be so large.

In Column (7), household income status is controlled for with four dummy variables. Although the sample size is considerably smaller than the benchmark specification in Column (5), the coefficient of $Prob_{npt}$ is negative and significant. Moreover, household income status is not associated with the children's subjective symptoms. This finding is in line with Nakamura (2013) who investigates the association between parents' income and children's health, using the CSLC from 1998 to 2007. According to her study, parental income is negatively associated with specific items of symptoms such as "wheezing" and "difficulties in hearing," but there is no association on "any symptoms" because "skin" problem and "injury" are rather prevalent among higher-income household.[23]

On the other control variables, parents' subjective health status is strongly related to the health of children. Subsequently, being a girl, larger household size, home ownership and enrollment in CHI plans reduce the probability, while spouse's employment increases it. Interpretation of these results is too specific and beyond my scope, but the estimates on the coefficient of $Prob_{npt}$ are still robust after controlling these covariates.

Next, the results on elementary school-aged children are reported in Table 4. On the contrary to the results for preschool children, all regression suggests that there is no association between MSCI expansions and the probability of having any symptoms. In addition, these results suggest that the effectiveness of reduced cost-sharing may vary across child age. When it comes to the effects on individual items such as "fever" and "wheezing," the conclusion is not changed to a large extent. Table 5 summarizes the results on 12 items among preschool (Panel A) and elementary school-age children (Panel B). On each symptom, I estimate a benchmark specification that controls for individual- and prefecture-level covariates and prefecture specific trends and report the coefficient of $prob_{npt}$. In the upper section in each panel, full sample results

[23]She shows there is a weak but significant relationship between parental income and subjective general health.

are presented. Next, the results from two robustness checks are reported in the middle and bottom section. First, the subsample analysis that corresponds the estimation in Column (6) in Table 3 and 4 are shown in middle section and the results that control for household income are in bottom section.

Among 12 items, I find significant and robust effects of MSCI expansions in "fever," "fatigue," "cough," and "stuffy nose" in Panel A. For these items, the impact of MSCI eligibility is consistently significant. If these symptoms suggest an infection with common cold, it is likely that the health benefits of the MSCI are limited. More importantly, the effect on "wheezing," which may be associated with asthma, is not significant. This suggests that observed improvement in health status in preschool children is not attributable to the alleviation in asthma-related symptoms. Given that the out-of-pocket burden for asthma-related treatment may discourage periodic visits, reduced cost-sharing should have improved respiratory symptoms through improved adherence. However, I find no such improvement. On the sample of school-age children, the coefficients of $Prob_{npt}$ exhibit no significant and robust effects for any item, suggesting the MSCI for school-age children has no significant health benefit.

6.2 Effects on Hospitalization

Although the previous subsection investigates the impact on various types of subjective symptoms, the results necessary suffer from measurement errors on health status because of mis-reporting of parents and ambiguity of the definition of individual symptoms. Instead, hospitalization is one of the most common and reliable measures for the evaluation of health status since the decision on whether a child should be admitted to hospital or not depends on the evaluation from a specialist. In my context, it is possible *a priori* that expansions of the MSCI reduce or increase hospitalization. On the one hand, the improved access to primary care would reduce preventable hospitalizations. Empirically, some studies support this hypothesis in the general population (Chandra et al, 2010; Dafny and Gruber, 2005; Kolstad and Kowalski, 2012) and patients with specific chronic diseases (Mahoney, 2005; Zeber et al, 2007; Jonathan et al, 2011; Karaca-Mandic et al, 2012; Rahimi et al, 2007; Goldman et al, 2006).

On the other hand, reduction of co-payments would increase hospitalization if the price elasticity of child hospitalization is negative.[24] Since the MSCI reduces the co-insurance rate for inpatient care, it is possible that hospitalization increases because of the lower out-of-pocket cost for inpatient care. However, timing of the expansions of the MSCI for inpatient care preceded that for outpatient care. For instance, 72 percent of all municipalities had already expanded the eligible age for the MSCI for inpatient care over the preschool age by April 2004, but the MSCI for outpatient care had been expanded in 44 percent of all municipalities over preschool age by the same time(MHLW, 2013). In other words, changes in the coinsurance rate for inpatient care would not directly affect my results since, in many municipalities, the coinsurance rate for inpatient care remained unchanged when the MSCI for outpatient care was just expanded.

[24] On this issue, the RAND HIE suggests child hospitalization was not elastic to cost-sharing (Newhouse, 1993). However Almond and Doyle (2011) shows post-natal hospitalization would be elastic to co-payments. A recent randomized experiment from Oregon presents no results on child hospitalization, but Finkelstein et al (2012) and Taubman et al (2014) reveal hospitalization and emergency room visits in the adult population increased through the eligibility for Medicaid.

The results on hospital admission are summarized in Table 6. On the contrary to the results on subjective symptoms, this table shows that expansions of the MSCI have no impact on hospitalization among both preschool and elementary school-age children. On the results of preschool children, the point estimates of the impact are consistently positive and not significant. More importantly, these results hold even after focusing on a sub-sample that includes prefectures with high response rates (Column 6) and controls for household income (Column 7). The results for school-age children also robustly exhibit no effect of the MSCI. Taken together, the results suggest that reduction of the coinsurance rate for outpatient care and prescription drugs from 30 or 20 percent to 0 percent in Japan would not save the cost of inpatient care.

These findings correspond to previous results from two randomized control trials (RAND HIE (Newhouse, 1993), and Oregon HIE (Taubman et al, 2014)) and several observational studies such as Pilote et al (2002) and Wang et al (2010). However, it should be cautioned that this study does not capture the long-term effect of reduced cost-sharing for children. If observed improvement in subjective symptoms among preschool children generates long-term health benefits for newly-eligible children, this will reduce costs for the future health care system.

6.3 Effects on Other Outcomes among School Age Children

Since the CSLC surveys several outcomes only for children aged six years old and over, I present the results on these outcomes among elementary school-age children. First, I examine the effect on subjective health via two binary variables. One is the probability of reporting "good" or "excellent" health and another is that of reporting "bad" and "very bad" health. Next, the effects on the probability of reporting role limitation due to health, which is measured by the question "Does your child feel difficulties in daily life due to health problems?", are summarized. Importantly, the RAND HIE also examines the effect of cost-sharing on these outcomes[25] find no discernible differences between free and cost-sharing plans (Valdez et al, 1985).

Consistent with their findings, I find no effect on general health and role limitation. In Table 7, the coefficients in all panels are not significant regardless of various alternative specifications. Although standard errors in Panel A are somewhat small, they are not enough to be significant, suggesting that health benefit of MSCI expansions for school-age children is very limited or negligible.

6.4 Robustness Checks

Two threats to identification are omitted variable bias (i.e. unobservables may jointly affect child health and the expansions of the MSCI) and measurement errors in the treatment status (i.e. probability of being eligible for the MSCI is not understood with complete accuracy since the response rate for the author's original survey is not 100%) . In order to check the robustness of the main results for these threats, I implement several robustness checks in Online Appendix A and B. In Online Appendix A, I control for

[25] On role limitation, Valdez et al (1985) examines the effect on the item that "Is this child limited in the amount or kind of other activities (such as playing , helping around the house, hobbies) because of health ?" and on subjective health, they examine the effect on General Health Rating, which is a quite similar item to 5-grade subjective health.

cross effects of districts and years, by incorporating district-year fixed effects. Since respondents of CSLC are chosen from the 5,530 survey districts and we can identify which district a child comes from, inclusion of district-year FEs may control for unobservable regional characteristics more accurately than prefecture-specific trends. Next, in Online Appendix B, I test whether the main results change if children from the prefectures with low response rates for the original survey are excluded more comprehensively than in the main tables. Details on these robustness checks are summarized in the appendix. All in all, I find the main results are robust for these checks.

7 Discussion

Exploiting recent dramatic expansions of the medical subsidy for children, this paper shows how co-payment reduction improves child health. On the empirical analysis, time-series evolutions of eligibility criteria that are investigated from an original survey for this paper are matched to nationally-representative data of children's health. The results suggest that reduced patient cost-sharing would lead to improvement of subjective symptoms in preschool children. Numerically, the probability of having any symptoms decreases 2.8 percentage points if a child is eligible for the MSCI. On the other hand, I find no such an improvement among school-age children. In addition, the effect on subjective health and role limitation due to health are not significant among school age children. More importantly, there is no association between MSCI eligibility and the probability of hospitalization both among preschool and school-age children. Given that hospitalization is one of the most common and reliable measures of health condition and an efficiency argument has stressed that reduced cost-sharing for outpatient care and prescription drugs may save total health care costs through reduced hospitalization, my finding on the effect on hospitalization in the general childhood population is of particular importance. Indeed, this paper suggests that recent rapid expansions of the MSCI in Japan have not led to efficient consequences, at least in the short run. In addition, this study finds no discernible effects on several health measures in school-age children, suggesting recent rapid expansions of the MSCI for this age group have not been associated with the improvement of health status.

As is mentioned in Currie et al (2008) who shows expansions of public health insurance has no effect on current health status of children but has a positive effect on future health, however, the importance of cumulative effects on health in childhood should not be underestimated. Similarly, it is possible that children's health status may improve in the long run since I have observed positive impacts on subjective symptoms at least among preschool children. When those preschool children grow up to be adults, there may be substantial cost saving effects, though these important research question are far beyond the scope of this paper.

There are several limitations to this paper. First, as that mentioned previously, eligibility status for the MSCI is not understood correctly in this paper. The reason why is because the data on past history of the eligible age for MSCI is based on the author's original survey for all municipalities. Although this survey covers 75 percent of the population under fifteen years old and provides the best knowledge to date on the rapid expansions of the MSCI in Japan, some potential measurement errors are inevitable. Since this paper shows a variety of robustness checks to discern the seriousness of this problem and all results hold after these checks, I believe that central findings of my study are still valid even if we have a complete set of eligibility criteria in all municipalities during the study period. Nevertheless, needless to say, more comprehensive knowledge on the dynamics of the MSCI for all municipalities is desirable. Second, this paper does not reveal the mechanism through which the MSCI may affect children's health. In particular the effects of MSCI on the utilization of preventive care and adherence to prescription drugs should be examined with other data. These issues are left for future studies.

However, regardless of these limitations, it is of particular importance to present new empirical results on the impact of cost-sharing on children's health. For instance, this study examines the effect of reduced cost-sharing on children, without interviewing pregnant women. This perspective is noteworthy because some previous studies that exploit Medicaid expansions in the U.S. (Currie and Gruber, 1996a,b; Dafny and Gruber, 2005) measure the impact of health insurance coverage *in utero* on health after birth and find comparatively large health benefits. On the contrary, the results of this paper are not affected by the reduction of cost sharing *in utero* since pregnant women are not eligible for MSCI. Although the impact of insurance coverage is not comparable to that of reduced cost-sharing, this difference may partly explain why the MSCI does not have large effects on child health.

References

Aizer A (2007) Public Health Insurance, Program Take-Up, and Child Health. The Review of Economics and Statistics 89(3):400–415

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

Atella V, Peracchi F, Depalo D, Rossetti C (2006) Drug compliance, co-payment and health outcomes: Evidence from a panel of italian patients. Health Economics 15(9):875–892

Baicker K, Goldman D (2011) Patient cost-sharing and healthcare spending growth. Journal of Economic Perspectives 25(2):47–68

Baicker K, Taubman SL, Allen HL, Bernstein M, Gruber JH, Newhouse JP, Schneider EC, Wright BJ, Zaslavsky AM, Finkelstein AN (2013) The oregon experiment: Effects of medicaid on clinical outcomes. New England Journal of Medicine 368(18):1713–1722

Baker M, Stabile M, Deri Armstrong C (2004) What do self-reported, objective, measures of health measure? Journal of Human Resources 39(4)

Bessho S (2012) Medical care subsidies for, hospital visits of,and health status of children. Quarterly of Social Security Research (in Japanese) 47(4):413–430

Case A, Paxson C (2011) The long reach of childhood health and circumstance: Evidence from the whitehall ii study. Economic Journal 121(554):F183–F204

Chandra A, Gruber J, McKnight R (2010) Patient cost-sharing and hospiization offsets in the elderly. American Economic Review 100(1):193–213

Courtemanche CJ, Zapata D (2014) Does universal coverage improve health? the massachusetts experience. Journal of Policy Analysis and Management 33(1):36–69

Currie J (2009) Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. Journal of Economic Literature 47(1):87–122

Currie J, Gruber J (1996a) Health insurance eligibility, utilization of medical care, and child health. The Quarterly Journal of Economics 111(2):431–66

Currie J, Gruber J (1996b) Saving babies: The efficacy and cost of recent changes in the medicaid eligibility of pregnant women. Journal of Political Economy 104(6):1263–96

Currie J, Gruber J (2001) Public health insurance and medical treatment: the equalizing impact of the medicaid expansions. Journal of Public Economics 82(1):63–89

Currie J, Moretti E (2007) Biology as destiny? short- and long-run determinants of intergenerational transmission of birth weight. Journal of Labor Economics 25:231–264

Currie J, Staible M (2006) Child mental health and human capital accumulation: The case of adhd. Journal of health economics 25:1094?1118

Currie J, Decker S, Lin W (2008) Has public health insurance for oder children reduced disparities in access to care and health outcomes? Journal of Health Economics 27(6):1567–1581

Dafny L, Gruber J (2005) Public insurance and child hospitalizations: Access and efficiency effects. Journal of Public Economics 89(1):109–129

Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, Allen H, Baicker K, the Oregon Health Study Group (2012) The oregon health insurance experiment: Evidence from the first year. The Quarterly Journal of Economics

Fletcher JM, Green JC, Neidell MJ (2010) Long term effects of childhood asthma on adult health. Journal of Health Economics 29(3):377–387

Goldman D, Joyce G, Zheng Y (2007) Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. JAMA 298(1):61–69

Goldman DP, Joyce GF, Karaca-Mandic P (2006) Varying pharmacy benefits with clinical status: The case of cholesterol-lowering therapy. Am J Manag Care 12:21–28

Gross T, Notowidigdo MJ (2011) Health insurance and the consumer bankruptcy decision: Evidence from expansions of medicaid. Journal of Public Economics 95:767 – 778

Ministry of Health, Labour and Welfare (2013) Survey on the medical subsidy for children and infants, URL http://www.mhlw.go.jp/stf/houdou/2r9852000002xx3m.html, (Accessed on March 5, 2014)

Ministry of Health, Labour and Welfare (2014) Comprehensive survery on living conditions 2012, URL http://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa11/dl/12.pdf, (Accessed on November 15, 2014.)

Idler EL, Benyamini Y (1997) Self-rated health and mortality: A review of twenty-seven community studies. Journal of Health and Social Behavior 38(1):pp. 21–37

Johnson R, Goodman M, Hornbrook M, Eldredge M (2013) The impact of increasing patient prescription drug cost sharing on therapeutic classes of drugs received and on the health status of elderly hmo members. Health Policy 112(1-2):35 – 44

Johnston DW, Propper C, Shields MA (2009) Comparing subjective and objective measures of health: Evidence from hypertension for the income/health gradient. Journal of Health Economics 28(3):540 – 552

Jonathan DC, Allen-Ramey F, Sajjan SG, Maiese EM, Sullivan SD (2011) Increasing pharmaceutical copayments: Impact on asthma medication utilization and outcomes. Am J Manag Care 17(10):703–710

Karaca-Mandic P, Jena AB, Joyce GF, Goldman DP (2012) Out-of-pocket medication costs and use of medications and health care services among children with asthma. JAMA 307(12):1284–1291

Kiil A, Houlberg K (2014) How does copayment for health care services affect demand, health and redistribution? a systematic review of the empirical evidence from 1990 to 2011. The European Journal of Health Economics 15(8):813–828

Kolstad JT, Kowalski AE (2012) The impact of health care reform on hospital and preventive care: Evidence from massachusetts. Journal of Public Economics 96(12):909 – 929

Mahoney JJ (2005) Reducing patient drug acquisition costs can lower diabetes health claims. Am J Manag Care 11:S170–S176

Mann B, Barnieh L, Tang K, Campbell D, Clement F, Hemmelgarn B, Tonelli M, Lorenzetti D (2014) Association between drug insurance cost sharing strategies and outcomes in patients with chronic diseases: A systematic review. PLoS ONE 9(3):e89,168

Manning WG, Newhouse JP, Duan N, Keeler EB, Leibowitz A, Marquis S (1987) Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review 77(3):251–77

Mojtabai R, Olfson M (2003) Medication costs, adherence, and health outcomes among medicare beneficiaries. Health Affairs 22(4):220–229

Nakamura S (2013) Parental income and child health in japan. Journal of the Japanese and International Economies (0):–

Newhouse J (1993) Free for All: Lessons from the RAND Health Insurance Experi-ment. Harvard University Press, Cambridge, MA.

Nishi A, McWilliams JM, Hashimoto HNH, Tamiya N, Kawachi I (2012) Health benefits of reduced patient cost sharing in japan. Bull World Health Organ 90(6):426–435A

Nishikawa M (2010) Income restriction and individual payment in the infant medical-expense support grant. Aoyama Journal of Economics (in Japanese) 62(3):195 – 214

Nishikawa M (2011) Spatial relation among municipal decision making about infant medical-expense support. Aoyama Journal of Economics (in Japanese) 62(4):87 – 111

OECD (2012) Oecd family database: Child well-being module, URL http://www.oecd.org/els/family/oecdfamilydatabasechildwell-beingmodule.htm, (Accessed on March 5, 2014.)

Pilote L, Beck C, Richard H, Eisenberg MJ (2002) The effects of cost-sharing on essential drug prescriptions, utilization of medical care and outcomes after acute myocardial infarction in elderly patients. Canadian Medical Association Journal 167(3):246–252

Rahimi A, Spertus J, Reid K, Bernheim S, Krumholz H (2007) Financial barriers to health care and outcomes after acute myocardial infarction. JAMA 297(10):1063–1072

Shigeoka H (2014) The effect of patient cost sharing on utilization, health, and risk protection. American Economic Review 104(7):2152–84

Tamblyn R, Laprise R, Hanley J, et al (2001) Adverse events associated with prescription drug cost-sharing among poor and elderly persons. JAMA 285(4):421–429

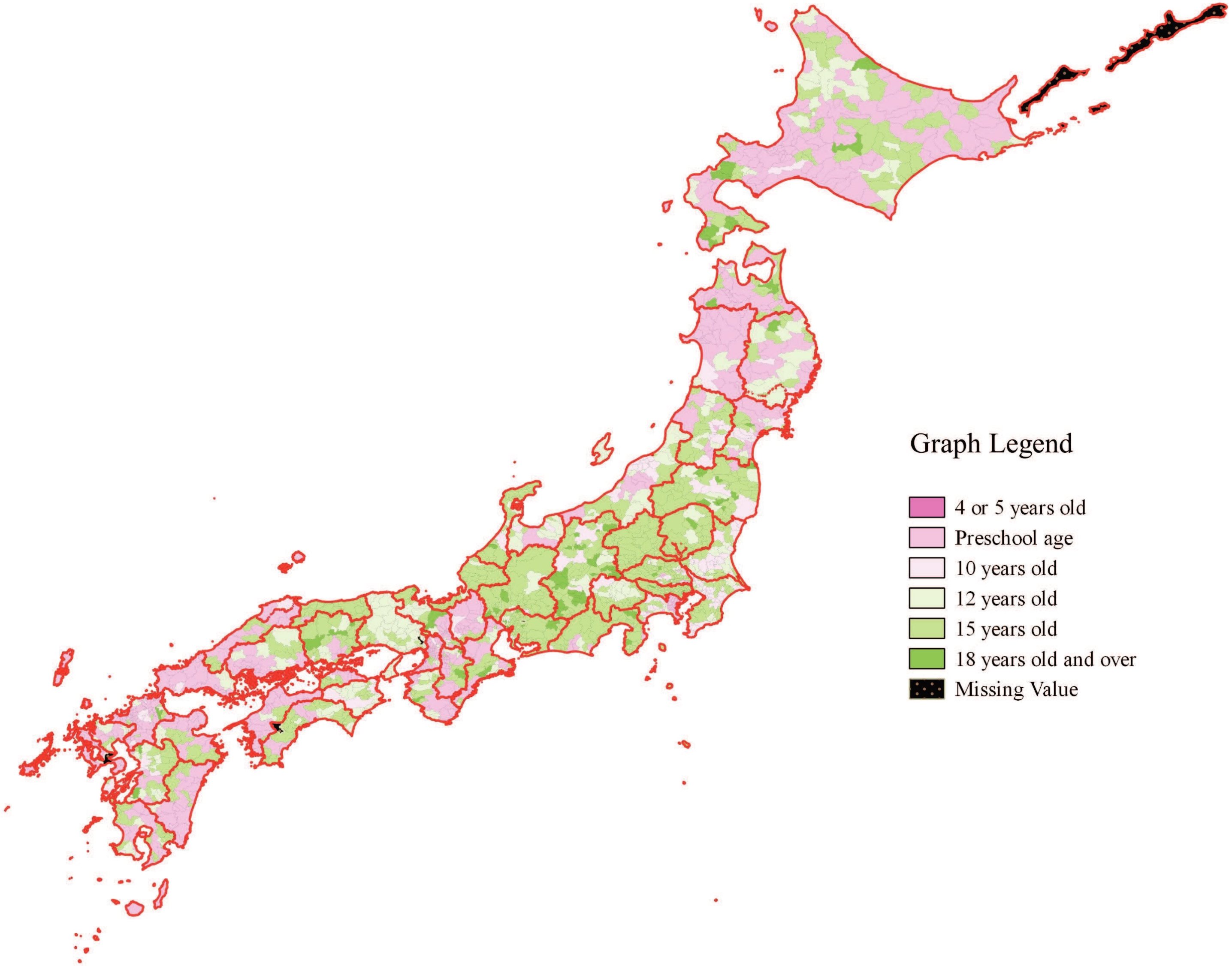
Taubman SL, Allen HL, Wright BJ, Baicker K, Finkelstein AN (2014) Medicaid increases emergency-department use: Evidence from oregon's health insurance experiment. Science 343(6168):263–268

Trivedi AN, Moloo H, Mor V (2010) Increased ambulatory care copayments and hospitalizations among the elderly. The New England Journal o f Medicine 362:320–328

Valdez RB, Brook RH, Rogers WH, Ware JE, Keeler EB, Sherbourne CA, Lohr KN, Goldberg GA, Camp P, Newhouse JP (1985) Consequences of cost-sharing for children's health. Pediatrics 75(5):952–961

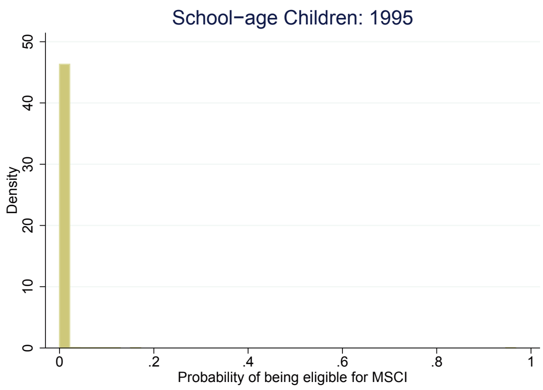
Wang PS, Patrick AR, Dormuth C, Maclure M, Avorn J, Canning CF, Schneeweiss S (2010) Impact of drug cost sharing on service use and adverse clinical outcomes in elderly receiving antidepressants. The Journal of Mental Health Policy and Economics 13(1):37–44

Zeber JE, Kyle L Grazier aMV, Blow FC, Lantz PM (2007) Effect of a medication copayment increase in veterans with schizophrenia. Am J Manag Care 13(2):335–346

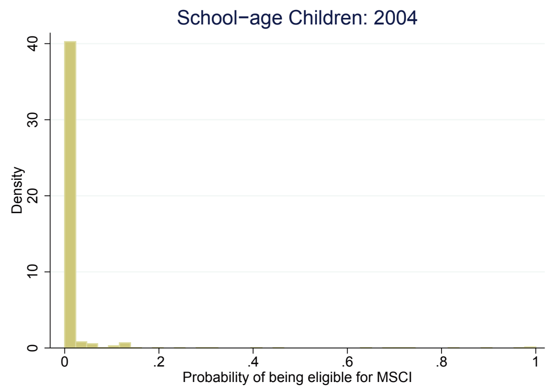


Note: This figure illustrates a municipal-level map of Japan shaded according to eligible age for the MSCI for outpatient care. Red lines represent prefectural borders.
Source: MHLW (2013)

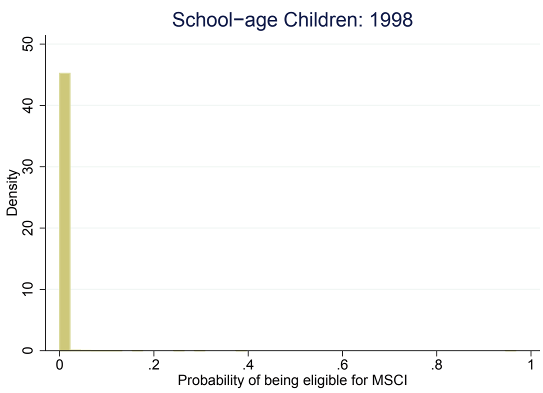
Figure 1: Geographical Distribution of the Eligible Age for MSCI: Outpatient Care



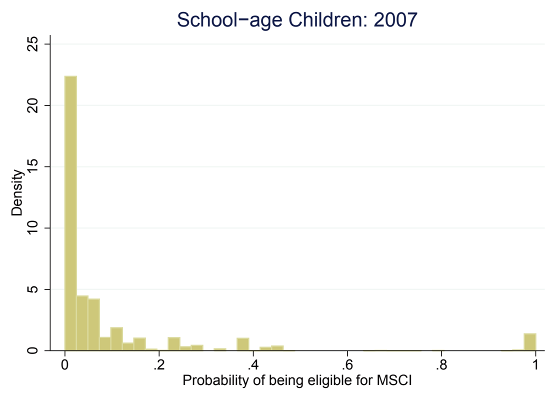
(a) 1995



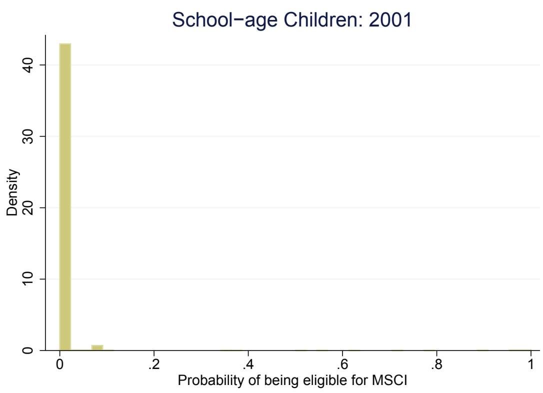
(d) 2004



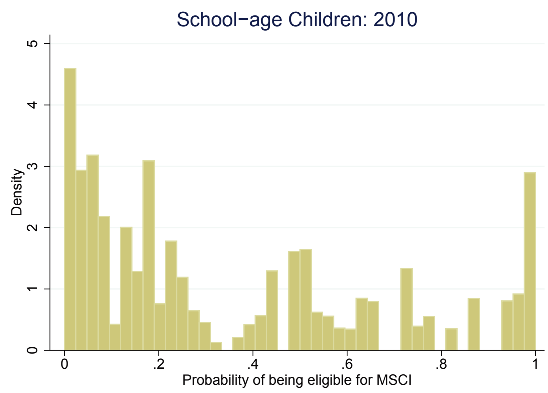
(b) 1998



(e) 2007



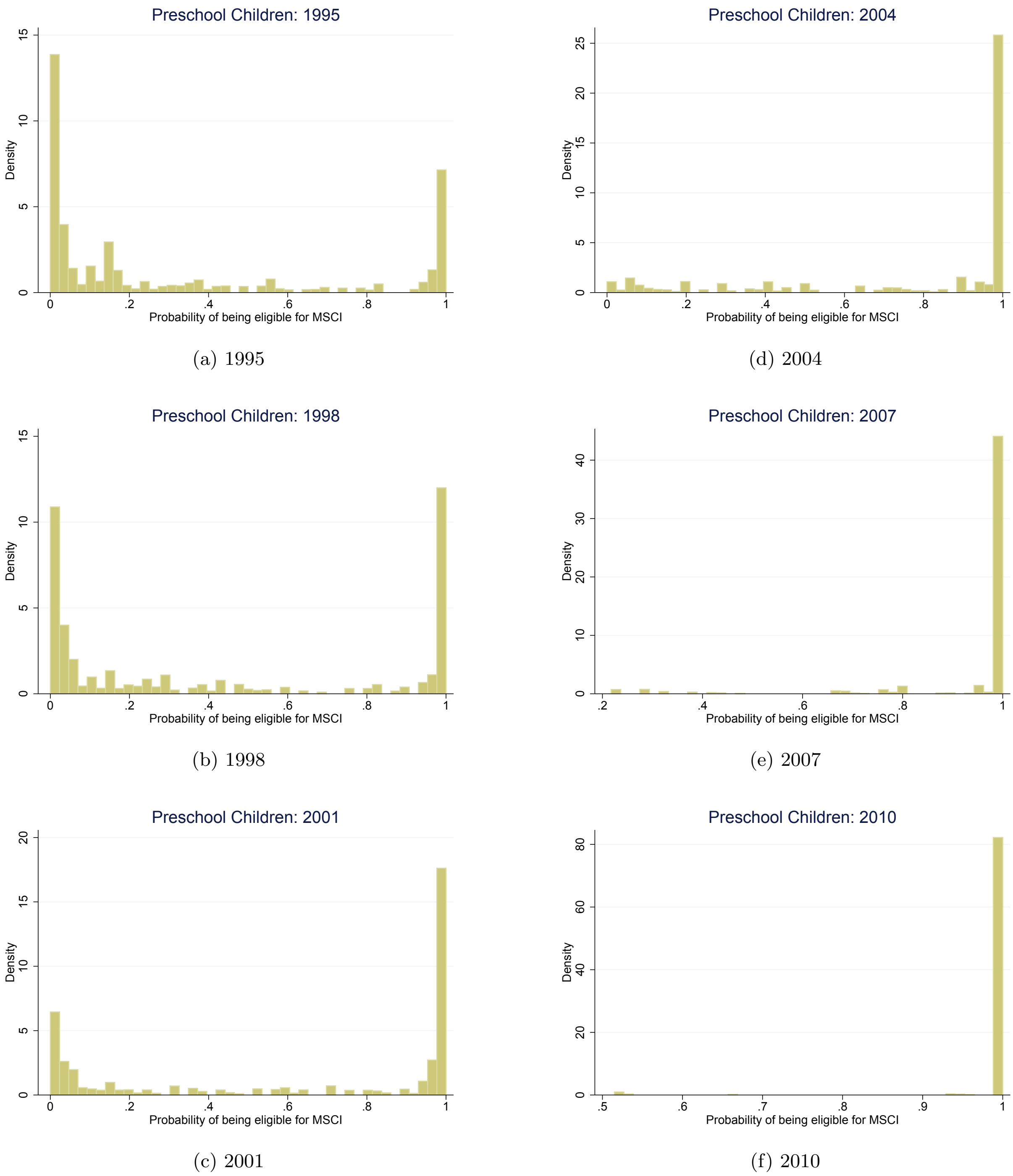
(c) 2001



(f) 2010

Note: The sample consists of school-age children enrolled in elementary school, roughly aged six to twelve.

Figure 2: Distribution of the probability of being eligible for MSCI: School-age Children



(a) 1995 (d) 2004

(b) 1998 (e) 2007

(c) 2001 (f) 2010

Note: The sample consists of preschool children over 12 months.

Figure 3: Distribution of the probability of being eligible for MSCI: Preschool Children

Table 1: Average Eligible Age for MSCI: 1995-2010

Prefectures	1995	1998	2001	2004	2007	2010	Response Rate
Hokkaido	1.5	1.5	2.3	3.7	6.2	7.1	75.0
Aomori	1.3	2.1	2.2	2.2	3.7	6.3	48.8
Iwate	0.6	1.1	2.8	3.1	6.0	7.3	90.6
Miyagi	0.9	0.9	0.9	1.7	3.2	4.4	94.9
Akita	0.3	0.5	1.4	1.6	6.0	6.5	48.7
Yamagata	1.1	2.2	3.3	4.2	5.6	5.3	85.8
Fukushima	2.0	2.1	5.1	5.1	7.4	10.1	77.1
Ibaraki	0.4	1.6	1.6	2.3	6.2	8.3	65.8
Tochigi	0.2	1.0	2.1	2.7	8.0	13.8	51.2
Gunma	1.7	2.9	5.5	5.7	6.4	14.2	69.9
Saitama	1.3	1.4	1.8	4.9	6.4	10.9	72.3
Chiba	0.3	1.7	2.1	2.5	4.9	8.3	89.6
Tokyo	2.6	4.8	5.5	5.8	11.1	14.0	70.0
Kanagawa	0.3	1.9	2.7	4.4	6.1	6.2	87.2
Niigata	0.3	0.3	0.9	2.6	5.9	9.7	54.1
Toyama	0.4	0.8	1.2	1.2	6.4	8.8	93.4
Ishikawa	0.8	1.5	1.8	5.2	8.6	9.1	72.3
Fukui	1.3	1.3	3.2	3.6	6.1	10.4	90.5
Yamanashi	0.7	1.8	3.8	4.8	8.0	11.5	57.1
Nagano	3.2	3.2	4.6	5.4	6.4	8.9	65.0
Gifu	1.3	1.9	2.4	3.7	8.9	11.8	68.6
Shizuoka	0.2	0.5	1.2	3.0	6.1	10.1	55.1
Aichi	1.9	1.9	2.4	4.7	6.2	12.6	94.7
Mie	0.5	1.0	1.0	2.0	4.6	7.2	57.0
Shiga	0.4	0.6	0.9	1.4	5.5	5.5	68.9
Kyoto	1.2	1.2	2.2	5.7	5.8	6.3	79.5
Osaka	0.9	1.3	2.1	3.4	5.0	6.1	67.9
Hyogo	1.8	1.9	2.2	5.4	7.9	8.8	90.7
Nara	0.7	1.7	1.7	2.3	5.1	6.6	60.4
Wakayama	1.5	1.5	1.5	2.1	6.1	6.4	68.4
Tottori	0.5	1.2	1.9	2.7	5.6	7.6	53.2
Shimane	0.2	0.2	0.2	0.5	5.9	6.3	64.0
Okayama	0.4	1.7	2.1	2.7	6.6	7.6	91.0
Hiroshima	0.3	0.4	0.8	1.6	6.4	10.1	79.1
Yamaguchi	1.2	1.2	1.2	3.9	6.0	8.6	91.0
Tokushima	1.0	1.9	1.9	2.8	6.4	10.3	24.5
Kagawa	1.0	1.2	2.2	4.2	4.6	4.6	24.5
Ehime	0.4	0.4	0.4	0.6	8.4	10.0	88.6
Kochi	0.0	0.1	0.1	1.0	4.0	6.4	69.0
Fukuoka	1.4	1.4	1.4	1.9	4.6	6.1	68.6
Saga	1.2	1.4	1.4	1.5	4.8	9.1	67.6
Nagasaki	1.0	1.0	1.0	1.5	6.0	6.0	74.8
Kumamoto	0.6	0.7	0.9	4.4	6.5	7.5	81.9
Oita	0.4	0.5	0.5	3.8	5.8	9.9	79.8
Miyazaki	1.7	1.7	3.4	4.3	4.6	10.2	85.7
Kagoshima	0.9	0.9	0.9	0.9	5.5	7.1	72.3
Okinawa	0.4	2.6	3.4	3.5	3.9	4.6	58.2

Note: Average eligible age for MSCI is calculated as a population-weighted average of eligible age in municipalities within a prefecture. The eligible age of municipalities is based on the original survey conducted by the author. The response rate of this survey is reported in the right column in the figure. The response rate is also calculated with population aged under fifteen years old of municipality as a weight. If the eligibility criteria is set as “preschool children,” I assign a value of six since children start elementary school in April at the age of six.

Table 2: Descriptive Statistics

	Preschool Children				Elementary School Children			
	Mean	S.D	Min	Max	Mean	S.D	Min	Max
Panel A. Outcome								
Having Any Symptom	0.27	0.44	0	1	0.19	0.39	0	1
Hospitalization	0.00	0.07	0	1	0.00	0.07	0	1
"Good" or "Excellent" Health	0.73	0.44	0	1	0.72	0.45	0	1
"Bad" or "Very Bad" Health	0.03	0.16	0	1	0.02	0.14	0	1
Difficulty in Daily Life	0.03	0.17	0	1	0.03	0.17	0	1
Panel B. Covariates								
Age	3.20	1.54	1	6	8.78	1.79	6	12
Female	0.51	0.50	0	1	0.51	0.50	0	1
Firstborn	0.33	0.47	0	1	0.41	0.49	0	1
N. of Child	2.07	0.80	1	9	2.34	0.77	1	9
N. of Household Members	4.34	1.12	3	13	4.61	1.04	3	13
Head Age	38.14	9.96	21	70	42.73	8.73	21	70
Spouse Age	35.79	9.41	21	70	40.17	8.33	21	70
Head Job	0.97	0.17	0	1	0.97	0.18	0	1
Spouse Job	0.38	0.49	0	1	0.51	0.50	0	1
Insurance Status: CHI	0.25	0.43	0	1	0.26	0.44	0	1
Insurance Status: EBHI	0.75	0.33	0	1	0.74	0.32	0	1
Home-ownership	0.52	0.50	0	1	0.68	0.46	0	1
N. of Rooms : Under 3	0.31	0.46	0	1	0.18	0.39	0	1
N. of Rooms : 4	0.27	0.45	0	1	0.26	0.44	0	1
N. of Rooms : 5	0.18	0.39	0	1	0.24	0.43	0	1
N. of Rooms : 6-7	0.16	0.37	0	1	0.23	0.42	0	1
N. of Rooms : over 8	0.07	0.26	0	1	0.08	0.28	0	1
Subjective Health: Head	2.35	1.00	1	5	2.37	1.00	1	5
Subjective Health: Spouse	2.29	1.01	1	5	2.32	1.00	1	5
Number of Obs.	117,522				136,921			
Panel C. Household Income								
Less Than 2 Million JPY	0.04	0.19	0	1	0.03	0.18	0	1
2-4 Million JPY	0.23	0.42	0	1	0.17	0.37	0	1
4-6 Million JPY	0.38	0.49	0	1	0.33	0.47	0	1
6-8 Million JPY	0.24	0.43	0	1	0.30	0.46	0	1
8-10 Million JPY	0.11	0.31	0	1	0.17	0.38	0	1
Number of Obs.	13,991				15,077			

Note: The sample size for household income in Panel C is smaller than full sample in Panel A and B since the CSLC only surveys household income for about 10 percent of respondents.

Table 3: Effects on Subjective Symptoms: Preschool Children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prob.	-0.024***	-0.025***	-0.024***	-0.024***	-0.028***	-0.028***	-0.035**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.015)
Age	-0.060***	-0.060***	-0.065***	-0.065***	-0.066***	-0.051**	-0.090**
	(0.013)	(0.013)	(0.012)	(0.012)	(0.012)	(0.018)	(0.042)
$(\text{Age}/10)^2$	1.509***	1.505***	1.634***	1.626***	1.608***	1.222*	2.422*
	(0.422)	(0.423)	(0.413)	(0.414)	(0.410)	(0.615)	(1.389)
$(\text{Age}/10)^3$	-14.855***	-14.755***	-15.676***	-15.601***	-15.317***	-12.590*	-23.472*
	(4.173)	(4.167)	(4.099)	(4.121)	(4.081)	(6.092)	(13.648)
Girl	-0.020***	-0.020***	-0.020***	-0.020***	-0.020***	-0.024***	-0.012
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.008)
Firstborn	0.018***	0.018***	0.019***	0.019***	0.019***	0.013**	0.012
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.008)
Number of Children	0.003	-0.001	0.001	0.001	0.001	-0.010	0.001
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.013)
Number of Household Member	-0.016***	-0.012***	-0.011**	-0.011**	-0.011**	-0.003	-0.017
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.011)
Head Age	-0.001*	-0.000	-0.001**	-0.001**	-0.001**	-0.001**	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Spouse Age	0.001	0.001*	-0.000	-0.000	-0.000	-0.001	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Household Income: Less than 2 Million JPY							
Household Income: 2-4 Million JPY							-0.027
							(0.033)
Household Income: 4-6 Million JPY							-0.020
							(0.029)
Household Income: 6-8 Million JPY							-0.001
							(0.029)
Household Income: 8-10 Million JPY							-0.012
							(0.029)
Head Works		-0.002	0.005	0.005	0.005	0.001	0.026
		(0.009)	(0.010)	(0.010)	(0.010)	(0.016)	(0.033)
Spouse Works		0.013***	0.013***	0.013***	0.013***	0.009**	-0.007
		(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.009)
Insurance Status: SMHI							
Insurance Status: CHI		-0.010**	-0.010**	-0.010**	-0.010**	-0.007	-0.005
		(0.004)	(0.004)	(0.004)	(0.004)	(0.007)	(0.009)
Home Ownership		-0.018***	-0.014***	-0.014***	-0.013***	-0.012**	-0.004
		(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.012)
Number of Rooms: Less than 3							
Number of Rooms: 4		0.009**	0.011**	0.011**	0.010**	-0.012**	0.028*
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.016)
Number of Rooms: 5-6		0.005	0.009	0.009	0.009	0.019***	0.010
		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.018)
Number of Rooms: 7-8		-0.004	0.001	0.000	-0.000	0.021***	-0.005
		(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.017)
Number of Rooms: Over 9		-0.015*	-0.006	-0.007	-0.007	0.010	-0.005
		(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.036)
Head's Subjective Health Status			0.062***	0.062***	0.062***	0.065***	0.048***
			(0.002)	(0.002)	(0.002)	(0.003)	(0.008)
Spouse's Subjective Health Status			0.030***	0.030***	0.030***	0.029***	0.034***
			(0.001)	(0.001)	(0.001)	(0.002)	(0.007)
Physician Density				-0.000	-0.001	-0.004	0.014**
				(0.002)	(0.002)	(0.003)	(0.006)
Per Capita Income				0.000**	0.000	0.000	-0.000
				(0.000)	(0.000)	(0.000)	(0.000)
Unemployment Rate				0.011**	0.019*	0.014	-0.009
				(0.005)	(0.010)	(0.013)	(0.014)
Prefecture Specific Trends					X	X	X
Exclude Low Response Rate Prefectures						X	
R squared	0.011	0.012	0.043	0.043	0.044	0.046	0.052
Obs.	115,019	115,019	115,019	115,019	115,019	52,787	13,122

Note: This table presents the estimated impact of the increasing probability to be eligible for the MSCI on subjective symptoms. The analysis includes all preschool children of a married couple, aged 12 months and over. From Column (1) to (5), the results with and without various covariates are reported. Column (6) presents the results of sub-sample analysis that excludes prefectures where the weighted response rate for my original survey was under 70 percent. To Column (5), Column (7) additionally controls for household income. Since the CSLC surveys household income only for 10 percent of respondents, the sample size in Column (7) is small. Standard errors are clustered at prefecture level. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 4: Effects on Subjective Symptoms: School-age Children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prob.	-0.001	-0.001	0.000	0.001	0.007	0.006	-0.016
	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.014)	(0.043)
Age	-0.095**	-0.096**	-0.088*	-0.088*	-0.086*	-0.078	-0.021
	(0.046)	(0.046)	(0.045)	(0.045)	(0.045)	(0.060)	(0.137)
$(Age/10)^2$	0.921*	0.941*	0.855	0.853	0.836	0.775	0.036
	(0.532)	(0.533)	(0.523)	(0.523)	(0.523)	(0.672)	(1.576)
$(Age/10)^3$	-3.050	-3.131	-2.859	-2.850	-2.791	-2.703	0.372
	(1.997)	(1.999)	(1.963)	(1.962)	(1.960)	(2.479)	(5.931)
Girl	-0.018***	-0.018***	-0.018***	-0.018***	-0.018***	-0.016***	-0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.007)
Firstborn	0.034***	0.033***	0.032***	0.032***	0.032***	0.036***	0.036***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.009)
Number of Children	0.008***	0.003	0.004	0.004	0.004	0.004	0.013
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.011)
Number of Household Members	-0.012***	-0.007**	-0.005**	-0.005**	-0.005**	-0.005	-0.017
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.010)
Head Age	-0.001***	-0.001**	-0.001***	-0.001***	-0.001***	-0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Spouse Age	0.001*	0.001**	0.000	0.000	0.000	0.000	-0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Household Income: Less than 2 Million JPY							
Household Income: 2-4 Million JPY							-0.025
							(0.019)
Household Income: 4-6 Million JPY							-0.012
							(0.017)
Household Income: 6-8 Million JPY							0.006
							(0.019)
Household Income: 8-10 Million JPY							-0.001
							(0.018)
Head Works		0.002	0.007	0.007	0.007	0.006	0.026
		(0.007)	(0.007)	(0.007)	(0.007)	(0.011)	(0.020)
Spouse Works		-0.010***	-0.009***	-0.009***	-0.009***	-0.006	-0.017**
		(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.007)
Insurance Status: SMHI							
Insurance Status: CHI		-0.013***	-0.015***	-0.015***	-0.015***	-0.015**	0.000
		(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.008)
Home Ownership		-0.019***	-0.014***	-0.014***	-0.014***	-0.011	-0.004
		(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.008)
Number of Rooms: Less than 3							
Number of Rooms: 4		0.009**	0.011**	0.011**	0.011**	0.020***	-0.001
		(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.012)
Number of Rooms: 5-6		0.010**	0.014***	0.014***	0.014***	0.020***	-0.005
		(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.012)
Number of Rooms: 7-8		-0.001	0.003	0.003	0.003	0.012	-0.012
		(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.014)
Number of Rooms: Over 9		-0.008	-0.002	-0.002	-0.002	0.003	-0.019
		(0.007)	(0.007)	(0.007)	(0.007)	(0.013)	(0.022)
Head's Subjective Health Status			0.049***	0.049***	0.050***	0.051***	0.051***
			(0.001)	(0.001)	(0.001)	(0.002)	(0.005)
Spouse's Subjective Health Status			0.027***	0.027***	0.027***	0.027***	0.029***
			(0.002)	(0.002)	(0.002)	(0.002)	(0.004)
Physician Density				-0.000	-0.002	-0.001	0.014**
				(0.002)	(0.002)	(0.004)	(0.006)
Per Capita Income				0.000	-0.000	-0.000	-0.000
				(0.000)	(0.000)	(0.000)	(0.000)
Unemployment Rate				0.004	0.008	0.017*	-0.009
				(0.004)	(0.006)	(0.008)	(0.014)
Prefecture Specific Trends					X	X	X
Exclude Low Response Rate Prefectures						X	
R squared	0.009	0.010	0.036	0.036	0.037	0.045	0.045
Obs.	133,855	133,855	133,855	133,855	133,855	61,593	14,180

Note: This table presents the estimated impact of the increasing probability to be eligible for the MSCI on subjective symptoms. The analysis includes all elementary school aged children. From Column (1) to (5), the results with and without various covariates are reported. Column (6) presents the results of sub-sample analysis that excludes prefectures where the weighted response rate for my original survey was under 70 percent. To Column (5), Column (7) additionally controls for household income. Since the CSLC surveys household income only for 10 percent of respondents, the sample size in Column (7) is small. Standard errors are clustered at prefecture level. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 5: Effects on Subjective Symptoms: Individual Items

	Fever	Fatigue	Cough	Headache	Wheezing	Toothache	Stuffy Nose	Constipation	Diarrhea	Stomachache	Rash	Cut
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A. Preschool Children												
Full Sample												
Prob.	-0.019**	-0.005*	-0.040***	-0.004	-0.007	-0.003	-0.035**	-0.005*	-0.011	-0.004*	-0.006	0.002
	(0.009)	(0.003)	(0.010)	(0.002)	(0.006)	(0.004)	(0.016)	(0.003)	(0.007)	(0.002)	(0.006)	(0.005)
Obs.	115,019	115,019	115,019	115,019	115,019	115,019	115,019	115,019	115,019	115,019	115,019	115,019
Low Response Prefectures Excluded												
Prob.	-0.011**	-0.003**	-0.013*	0.000	-0.003	-0.005***	-0.021***	0.000	-0.002	0.001	-0.004	0.003
	(0.005)	(0.001)	(0.007)	(0.002)	(0.003)	(0.002)	(0.007)	(0.001)	(0.002)	(0.001)	(0.003)	(0.002)
Obs.	52,787	52,787	52,787	52,787	52,787	52,787	52,787	52,787	52,787	52,787	52,787	52,787
Household Income Controlled												
Prob.	-0.019**	-0.005*	-0.040***	-0.004	-0.007	-0.003	-0.035**	-0.005*	-0.011	-0.004*	-0.006	0.002
	(0.009)	(0.003)	(0.010)	(0.002)	(0.006)	(0.004)	(0.016)	(0.003)	(0.007)	(0.002)	(0.006)	(0.005)
Obs.	13,728	13,728	13,728	13,728	13,728	13,728	13,728	13,728	13,728	13,728	13,728	13,728
Mean of Dep.	0.06	0.01	0.12	0.00	0.03	0.01	0.13	0.01	0.02	0.01	0.03	0.01
Panel B. School Age Children												
Full Sample												
Prob.	0.003	0.001	0.012**	-0.003	0.001	0.000	-0.001	-0.001	0.003	0.003	0.003	-0.001
	(0.004)	(0.003)	(0.006)	(0.003)	(0.003)	(0.003)	(0.007)	(0.001)	(0.003)	(0.003)	(0.005)	(0.004)
Obs.	133,855	133,855	133,855	133,855	133,855	133,855	133,855	133,855	133,855	133,855	133,855	133,855
Low Response Prefectures Excluded												
Prob.	0.009	-0.004	0.008	-0.010***	-0.002	0.006	0.006	-0.002	0.004	0.003	0.010*	-0.006
	(0.006)	(0.004)	(0.007)	(0.003)	(0.005)	(0.004)	(0.011)	(0.002)	(0.004)	(0.004)	(0.005)	(0.005)
Obs.	61,593	61,593	61,593	61,593	61,593	61,593	61,593	61,593	61,593	61,593	61,593	61,593
Household Income Controlled												
Prob.	-0.009	-0.005	-0.002	-0.002	-0.002	0.003	-0.007	-0.005	0.003	0.006	0.019	-0.015
	(0.012)	(0.006)	(0.022)	(0.009)	(0.011)	(0.012)	(0.026)	(0.004)	(0.008)	(0.011)	(0.013)	(0.011)
Obs.	14,782	14,782	14,782	14,782	14,782	14,782	14,782	14,782	14,782	14,782	14,782	14,782
Mean of Dep.	0.02	0.01	0.05	0.01	0.02	0.01	0.08	0.00	0.01	0.01	0.02	0.01
Prefecture Fixed Effect	X	X	X	X	X	X	X	X	X	X	X	X
Year Effect	X	X	X	X	X	X	X	X	X	X	X	X
Prefecture Specific Trends	X	X	X	X	X	X	X	X	X	X	X	X
Individual-Level Controls	X	X	X	X	X	X	X	X	X	X	X	X
Prefecture-Level Controls	X	X	X	X	X	X	X	X	X	X	X	X

Note: This table presents the estimated impact of the increasing probability to be eligible for the MSCI on subjective symptoms. Means of dependent variables are reported in the bottom row in each panel. Standard errors are clustered at prefecture level. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 6: Effects on Hospitalization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A. Preschool Children				
Prob.	0.000	0.000	0.000	0.000	0.000	0.001	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Obs.	117,522	117,522	117,522	117,522	117,522	53,975	13,374
			Panel B. School Age Children				
Prob.	-0.002	-0.002	-0.002	-0.002	-0.002	-0.004	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.006)
Obs.	136,921	136,921	136,921	136,921	136,921	62,999	14,466
Year Effect	X	X	X	X	X	X	X
Prefecture Fixed Effect	X	X	X	X	X	X	X
Demographic Controls	X	X	X	X	X	X	X
Socio Economic Controls		X	X	X	X	X	X
Parents' Subjective Health			X	X	X	X	X
Prefecture-level Controls				X	X	X	X
Prefecture Specific Trends					X	X	X
Exclude Low Response Prefectures						X	
Control Household Income							X

Note: This table presents the estimated impact of the increasing probability to be eligible for MSCI on the probability of hospitalization. The analysis includes preschool children in Panel A and elementary school children in Panel B. From Column (1) to (5), the results with and without various covariates are reported. Column (6) presents the results of sub-sample analysis that excludes prefectures where the weighted response rate for my original survey was under 70 percent. To Column (5), Column (7) additionally controls for household income. Since the CSLC surveys household income only for 10 percent of respondents, the sample size in Column (7) is small. Standard errors are clustered at prefecture level. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table 7: Effects on Other Outcomes Among School Age Children

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A. “Good” or “Excellent” Health						
Prob.	0.018	0.018	0.016	0.016	0.017	0.014	0.051
	(0.015)	(0.015)	(0.013)	(0.012)	(0.012)	(0.013)	(0.036)
Obs.	130,365	130,365	130,365	130,365	130,365	60,131	13,824
	Panel B. “Bad” or “Very Bad” Health						
Prob.	0.005	0.005	0.005	0.004	0.002	-0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.014)
Obs.	130,365	130,365	130,365	130,365	130,365	60,131	13,824
	Panel C. Role Limitation						
Prob.	0.003	0.003	0.003	0.003	0.000	0.001	-0.013
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.015)
Obs.	131,387	131,387	131,387	131,387	131,387	60,596	13,911
Year Effect	X	X	X	X	X	X	X
Prefecture Fixed Effect	X	X	X	X	X	X	X
Demographic Controls	X	X	X	X	X	X	X
Socio Economic Controls		X	X	X	X	X	X
Parents’ Subjective Health			X	X	X	X	X
Prefecture-level Controls				X	X	X	X
Prefecture Specific Trends					X	X	X
Exclude Low Response Prefectures						X	
Control Household Income							X

This table presents the estimated impact of the increasing probability to be eligible for the MSCI on subjective health and role limitation due to health. The analysis includes all elementary school children and results of preschool children cannot be reported since these outcomes are asked for children aged six years old and over. From Column (1) to (5), the results with and without various covariates are reported. Column (6) presents the results of sub-sample analysis that excludes prefectures where the weighted response rate for my original survey was under 70 percent. To Column (5), Column (7) additionally controls for household income. Since the CSLC surveys household income only for 10 percent of respondents, the sample size in Column (7) is small. Standard errors are clustered at prefecture level. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Online Appendix

A Inclusion of District-Year Fixed Effects

In order to address differential regional heterogeneity in children's health status that are not attributed to MSCI expansions, an additional robustness check is implemented. Since respondents of the CSLC are chosen from the 5,530 survey districts defined in the Census, it seems to be effective for eliminating unobservable regional and time-varying factors by including district-year fixed effects.[26] Since prefecture-level time-varying variables are absorbed into these district-year FEs, my benchmark specification is reduced to the following equation,

$$H_{it} = \beta_0 + \beta_1 Prob_{npt} + Z_{it}\delta_0 + DistXyear + \varepsilon_{it}, \tag{5}$$

where $DistXyear$ represents district-year FEs. Given that this specification allows us to compare the children's health outcomes with different eligibility status within the same district and survey year, incidental factors for each district in a given survey year such as local prevalence of hay fever may be absorbed into this term. In addition, district-year FEs also absorb the influences of the unobservable economic and political environment, which affects both children's health and the MSCI system. From an econometric perspective, it is useful to compare the point estimate of α_1 and β_1. If we observe that β_1 is greater in magnitude than α_1, it suggests that omitted variables bias in benchmark specification may underestimate the impact of the MSCI.

The results are summarized in Table A1. In this table, I report the coefficients of $prob_{npt}$ on 5 outcomes that have been already investigated, and on each outcome, results from subsample analysis that excludes low response rate prefectures are also reported, as well as those from full sample analysis. In Column (1) in Panel A, the estimated effect of MSCI eligibility is -0.032 and significant, suggesting MSCI eligibility decreases the probability of having any symptoms by 3.2 percentage points. Furthermore, this impact is not greatly changed in prefectures where the calculation of $prob_{npt}$ may be inaccurate because of low response rate for my original survey (Column 2). Compared with the impact from benchmark specification in Table 3, the impact after controlling for district-year FEs is somewhat larger than my benchmark results (2.8 percentage points), suggesting omitted variable bias may underestimate the impact of MSCI eligibility, while difference in the size of impact is not so large. On the other outcomes, as in the former specification, there are no discernible impacts of the MSCI on health outcomes either among preschool or school-age children.

[26] Since it is impossible to track a given district over several survey years because the codes that represent survey districts change year by year, we cannot control for district fixed effects.

Table A1: Robustness Checks with District-Year FEs

	Symptom		Hospitalization		"Good" or "Excellent"		"Bad" or "Very Bad"		Role Limitation	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Preschool Children										
Prob.	-0.032***	-0.036***	0.000	0.007						
	(0.008)	(0.012)	(0.001)	(0.007)						
Obs.	115,019	52,787	117,522	53,975						
Panel B. School Age Children										
Prob.	0.034	0.039	0.002	0.007	0.012	-0.007	0.003	0.007	0.006	0.006
	(0.024)	(0.031)	(0.005)	(0.007)	(0.023)	(0.031)	(0.009)	(0.011)	(0.011)	(0.014)
Obs.	133,855	61,593	136,921	62,999	130,365	60,131	130,365	60,131	131,387	60,596
Number of Clusters	19,929	19,929	19,929	19,929	19,929	19,929	19,929	19,929	19,929	19,929
District-year FEs	X	X	X	X	X	X	X	X	X	X
Individual Level Controls	X	X	X	X	X	X	X	X	X	X
Exclude Low Response Prefectures		X		X		X		X		X

Note: This table presents the results of robustness checks that control for district-year FEs. The analysis in Panel A includes preschool children and that in Panel B includes elementary school-aged children. Outcome variable is the probability of having any subjective symptom in Columns (1) and (2), hospitalization in Columns (3) and (4), "Good" or "Excellent" health in Columns (5) and (6), "Bad" or "Very Bad" health in Columns (7) and (8), and role limitation due to health in Columns (9) and (10). Even-numbered columns present the results of sub-sample analysis that excludes prefectures where the weighted response rate for my original survey was under 70 percent. Standard errors are clustered at the district-year cells. On three outcomes in Columns (5)-(10), the results on preschool children are not reported since these outcomes are surveyed only for children aged six years and over. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

B Robustness Checks for Response Rate for the Original Survey

One potential threat for my analysis is that the average probability of being eligible for the MSCI is not understood with complete accuracy because some municipalities did not respond to the author's original survey. In Appendix B, I construct subsamples according to the response rate of the original survey and check whether the main results from the full sample are not changed when prefectures with a low response rate are dropped. Specifically, I change the threshold response rate from 50% to 90% and test the robustness of the results.

The results are summarized in Tables B1 to B7. In Tables B1, the results on probability of having any symptoms among preschool children are presented. Panel A reports the results from the full sample and Panel B controls for household income. As is mentioned previously, the sample size is smaller in Panel B than Panel A since the CSLC surveys household income only for about 10% of the respondents. In Column (1) in Panel A, I report the same result with that of Column (6) in Table 3. With the same covariates, I restrict the sample according to the response rate for my original survey, namely from 50% in Column (2) to 90% in Column (10). The results in Panel A show that the coefficient of the treatment (Prob.) is very stable for alternative subsamples, ranging from -0.024 in Column (3) to -0.049 in Column (10). This suggests the response rate does not affect my results markedly. In addition, I run the same regression for the subsample for which the CSLC surveys household income in Panel B. Because of smaller sample size, standard errors are larger than those in Panel A. However, point estimates are very similar with the counterparts in Panel A.

From Table B2 to Table B7, I present similar tables on the other outcomes. In general, the results from the baseline specification are not changed according to the response rate, indicating that the main findings in this paper will be still valid if we have complete data on the eligibility criteria of MSCI from 1995 to 2010. It should also be noted that impacts among school age children become significant when we control for household income and restrict the sample to prefectures with high response rate (See Tables B3,B4 and B7). However, the sizes of the impacts are implausibly large, suggesting small sample sizes in these estimations may result in biased coefficients.

Table B1: Robustness Checks For the Results on Symptoms: Preschool

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	-0.028***	-0.028***	-0.024***	-0.026***	-0.029***	-0.028***	-0.026**	-0.024*	-0.029**	-0.049**
	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.009)	(0.010)	(0.012)	(0.012)	(0.018)
R2	0.044	0.043	0.044	0.043	0.044	0.046	0.047	0.049	0.050	0.056
Obs.	115,019	107,147	97,941	88,995	72,852	52,787	37,144	26,741	24,413	14,724
Panel B. Controlling for Household Income										
Prob.	-0.035**	-0.034**	-0.036**	-0.035*	-0.034*	-0.024	-0.033	-0.022	-0.031	-0.122*
	(0.015)	(0.016)	(0.017)	(0.018)	(0.019)	(0.023)	(0.031)	(0.040)	(0.043)	(0.052)
R2	0.056	0.052	0.053	0.052	0.052	0.055	0.056	0.063	0.066	0.066
Obs.	13,122	12,594	11,664	10,805	9,302	6,777	4,275	2,777	2,516	1,500

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. “50%” represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B2: Robustness Checks For the Results on Hospitalization: Preschool

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	0.000	0.000	-0.000	-0.000	-0.000	0.001	0.002	0.000	-0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)
R2	0.002	0.002	0.002	0.002	0.002	0.002	0.003	0.003	0.003	0.004
Obs.	117,522	109,488	100,080	90,942	74,449	53,975	37,977	27,334	24,947	15,022
Panel B. Controlling for Household Income										
Prob.	0.001	0.002	0.002	0.001	0.002	0.003	0.001	-0.009	-0.003	-0.007
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)	(0.004)	(0.006)
R2	0.004	0.013	0.013	0.014	0.014	0.013	0.015	0.019	0.024	0.024
Obs.	13,374	12,837	11,889	11,016	9,487	6,925	4,372	2,838	2,566	1,532

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. “50%” represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B3: Robustness Checks For the Results on Symptoms: School-age

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	0.007	0.008	0.007	0.005	0.011	0.006	0.014	0.008	0.013	0.003
	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)	(0.014)	(0.015)	(0.010)	(0.011)	(0.029)
R2	0.037	0.037	0.037	0.037	0.037	0.037	0.036	0.036	0.035	0.035
Obs.	133,855	124,438	113,930	103,761	85,157	61,593	43,055	31,292	28,368	16,966
Panel B. Controlling for Household Income										
Prob.	-0.016	-0.026	-0.030	-0.021	-0.035	-0.040	-0.139***	-0.160***	-0.148***	-0.109
	(0.043)	(0.045)	(0.046)	(0.048)	(0.053)	(0.067)	(0.040)	(0.024)	(0.027)	(0.072)
R2	0.035	0.047	0.047	0.047	0.047	0.048	0.048	0.043	0.047	0.050
Obs.	14,180	13,597	12,640	11,650	10,060	7,344	4,599	3,079	2,770	1,644

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. "50%" represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled or with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B4: Robustness Checks For the Results on Hospitalization: School-age

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	-0.002	-0.000	-0.001	-0.001	-0.002	-0.004	-0.002	-0.005*	-0.005*	-0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.007)
R2	0.002	0.002	0.002	0.003	0.003	0.003	0.003	0.003	0.004	0.005
Obs.	136,921	127,277	116,483	106,133	87,118	62,999	44,036	32,000	29,020	17,317
Panel B. Controlling for Household Income										
Prob.	-0.003	-0.002	-0.000	-0.002	-0.002	-0.006	-0.017***	-0.019***	-0.016**	-0.016**
	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.006)
R2	0.005	0.009	0.008	0.008	0.008	0.011	0.014	0.021	0.026	0.028
Obs.	14,466	13,864	12,881	11,874	10,249	7,479	4,681	3,130	2,815	1,671

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. "50%" represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B5: Robustness Checks For the Results on Good Health

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	0.017	0.011	0.012	0.015	0.018	0.014	0.004	0.001	-0.001	0.016
	(0.012)	(0.012)	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.015)	(0.018)	(0.045)
R2	0.153	0.154	0.154	0.155	0.159	0.160	0.153	0.150	0.150	0.145
Obs.	130,365	121,178	110,926	101,100	83,087	60,131	42,014	30,495	27,665	16,516
Panel B. Controlling for Household Income										
Prob.	0.051	0.068*	0.054	0.049	0.064	0.075*	0.100*	0.072	0.066	0.079
	(0.036)	(0.035)	(0.036)	(0.040)	(0.041)	(0.039)	(0.047)	(0.075)	(0.084)	(0.137)
R2	0.145	0.180	0.181	0.181	0.181	0.178	0.183	0.175	0.184	0.194
Obs.	13,824	13,258	12,326	11,370	9,840	7,214	4,520	3,017	2,709	1,602

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. "50%" represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B6: Robustness Checks For the Results on Bad Health

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	0.002	0.001	0.002	0.001	0.001	-0.001	-0.002	-0.005	-0.003	-0.016
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.013)
R2	0.008	0.008	0.008	0.008	0.007	0.008	0.007	0.007	0.007	0.007
Obs.	130,365	121,178	110,926	101,100	83,087	60,131	42,014	30,495	27,665	16,516
Panel B. Controlling for Household Income										
Prob.	0.004	0.005	0.006	0.002	0.002	-0.012	-0.013	-0.021	-0.019	-0.032
	(0.015)	(0.015)	(0.015)	(0.016)	(0.018)	(0.016)	(0.013)	(0.017)	(0.019)	(0.021)
R2	0.007	0.016	0.017	0.016	0.016	0.017	0.018	0.019	0.021	0.023
Obs.	13,824	13,258	12,326	11,370	9,840	7,214	4,520	3,017	2,709	1,602

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. “50%” represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Table B7: Robustness Checks For the Results on Role Limitation

	Full	50%	55%	60%	65%	70%	75%	80%	85%	90%
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Baseline Specification										
Prob.	0.000	-0.000	0.001	0.003	0.003	0.001	0.002	0.001	-0.000	-0.008
	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)	(0.014)
R2	0.007	0.007	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.008
Obs.	131,387	122,149	111,828	101,919	83,716	60,596	42,318	30,725	27,870	16,642
Panel B. Controlling for Household Income										
Prob.	-0.013	-0.012	-0.012	-0.012	-0.019	-0.027*	-0.036**	-0.049**	-0.051*	-0.058**
	(0.015)	(0.015)	(0.016)	(0.018)	(0.018)	(0.014)	(0.016)	(0.022)	(0.026)	(0.018)
R2	0.008	0.015	0.015	0.015	0.015	0.015	0.016	0.019	0.024	0.027
Obs.	13,911	13,341	12,398	11,437	9,891	7,248	4,538	3,029	2,720	1,609

Note: Column (1) reports the results from full sample analysis, but the other columns exclude children who live in the prefectures where the response rate for my original survey is low. The percentages reported in the upper row show the threshold. "50%" represent that the estimation excludes children from prefectures where the response rate is below 50%. The specification in Panel A is based on a baseline model, without controlling for household income. In Panel B, household income is controlled for with 5-grade categorical variables, although the sample size decreases. $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

Chapter 5

児童手当が両親の心理的健康に与える影響: 中低所得世帯における検証*

高久玲音

概 要

民主党政権下で 2010 年に導入された「子ども手当」は、手当の財源として各種控除が廃止されたことから中高所得者にはほとんど恩恵のない政策だった。しかし、中低所得世帯では手当の増加額が控除廃止による負担増を上回ったため、ネットでの可処分所得の増加がもたらされたと考えられる。本稿ではそのような流動性の付与が、中低所得世帯の両親の心理的健康にどのような影響を与えたか、日本家計パネル調査（JHPS）を用いて検証した。分析の結果、「子ども手当」導入による現金給付の拡充は両親の主観的健康を有意に向上させたことが分かった。年間 10 万円の現金給付の増加によって、母親が健康状態について「良い」と答える確率は 9% 上昇し、父親や健康状態について「悪い」と回答する確率は 6% 減少していた。また、父親では心身の自覚症状を示す指標が改善し、母親では「現在の生活に不満を感じる」と答えるサンプルが有意に減少するなど生活に対する満足度が上昇していることが確認できた。最後に、喫煙、飲酒といった嗜好品の消費行動や日常的な運動などの生活習慣への影響も分析したが、児童手当の影響は全くなかった。

キーワード : 児童手当、主観的健康
JEL classification : I18、H51

*本稿の執筆にあたり、文部科学省による科学研究費 (補助金番号 24330097、研究代表者　泉田信行) の補助を受けました。また、黒田祥子先生、権丈善一先生、佐藤一磨先生、直井道生先生、樋口美雄先生、別所俊一郎先生、山本勲先生、及び 2 名の匿名のレフェリーより頂きましたコメントと助言に感謝申し上げます。本稿の見解は文部科学省や所属組織の見解を反映するものではなく、全ての誤りは筆者に帰するものです。

1　はじめに

子育て世帯に対する現金給付を行う児童手当の拡充は、子どもの貧困を軽減するための最も基本的な方法と考えられ、日本を含めた先進諸国ではそうした政策の有効性が広く議論されている。その中でも、家計消費に対する児童手当の影響に関しては、既に我が国の経済学研究を展望しても多くの蓄積があり (田中, 2008; 両角, 2009; 小林, 2010; 宇南山, 2011)、児童手当に関する経済学研究の中心的トピックでもある。加えて、近年の諸研究はより広いアウトカムへの効果を推定するようになっている。例えば、カナダの児童手当や米国の EITC については、子どもの健康や学力への影響 (Milligan and Stabile, 2011; Dahl and Lochner, 2012)、出生時体重への影響 (Strully et al., 2010; Hoynes et al., 2012)、母親の主観的健康や幸福感への影響 (Evans and Garthwaite, 2010; Boyd-Swan et al., 2013) などが検証されている。また、健康状態に影響を与える消費行動として喫煙への影響も分析され、EITC の拡充は母親の喫煙率を低下させたことが指摘されている (Averett and Wang, 2012)。

多様なアウトカムに対する分析が行われるようになった背景には、発達心理学における研究展開がある。例えば、Yeung et al. (2002) は所得が子どもの発達に与える影響を二つの経路に分けて説明している。一つは子どもに対する金銭的・時間的な投資を通した経路である。この考え方は人的資本論 (Becker, 1981) によって提唱されている通りであり、子どもの人的資本は生物学的な生まれ持った能力と同時に、本や学習機材の購入といった両親による投資によって形造られると仮定する。よって、所得の高い家庭では子どものための様々な投資が可能であり、その結果として子どもは高い人的資本の蓄積が可能になる。一方、投資行動へ与える影響以外に重要な経路として、Yeung et al. (2002) では両親のストレスを通じた経路も強調されている。例えば、所得水準の向上は物質的欠乏に起因するストレスを緩和させ、両親の精神的健康状態や子どもに対する態度を改善させることによって、子どもの成長に寄与するかもしれない (Yeung et al., 2002; Mistry et al., 2002; Gershoff et al., 2007)。

そこで本稿では、民主党政権時代の 2010 年に導入された「子ども手当」が両親の心理的健康状態にどのような影響を与えたか、、2012 年版日本家計パネル調査（Japan Household Panel Survey）を用いて分析した。心理的健康を含めた両親の健康状態は、子どもの健康や人的資本の蓄積にとって極めて重要な要素だと考えらえることから (Case and Paxson, 2002; Propper et al., 2007; Wyman et al., 2007)、児童手当の政策効果を定量的に把握するためのアウトカムとして適切だと考えられる。

児童手当の政策効果を識別することに加えて、本稿の第 2 のモチベーションは所得と健康の因果関係についての頑健な知見を得ることにある。海外の研究を展望すると、多くの研究で高い社会経済的地位にある者ほど健康状態が良いことが指摘されている一方で、所得から健康への因果関係は必ずしも明らかになっていない (Deaton, 2003)。その点について、近年発表されている子どものいる世帯に対する現金給付政策の効果を扱った論文では、多くが両親及び子どもの健康指標に好ましい因果的効果を確認している (Evans and Garthwaite, 2010; Milligan and Stabile, 2011; Hoynes et al., 2012; Boyd-Swan et al., 2013)。これらの論文の強みは、外生性の高い所得変動を利用しながらダイレクトに所得と健康の因果関係を明らかにしているだけでなく、一般性の高い所得変動を用いていることにある。Evans and Garthwaite (2010) が指摘

するように、因果関係の識別を重視した諸研究では、所得のバリエーションとして特殊な状況を用いるものが多く[1]、広く一般に当てはまる推定値を導いているとは言えなかった[2]。しかし、所得保障政策の拡充は多くの人が影響を受ける政策であり、そこで得られた推定値は高い外的妥当性を持つことが期待される。

本稿の章立ては以下の通りである。まず 2 節では「子ども手当」の導入によって受給額がどのように変わったのか整理する。3 節は分析方法とデータの説明である。4 節は推計結果を、両親の主観的健康、心身症状指標、生活満足度・充実度の順に提示し考察を行う。最後に、5 節は結論と議論である。

2 「子ども手当」による純受給額の変化

まず、「子ども手当」をめぐる政策変遷を簡潔に把握するために、表 1 では給付額の変化を新旧児童手当と子ども手当で比較している。確認すると、小学生の場合で支給年額が 6 万円（旧児童手当）から 15.6 万円（子ども手当）、12 万円（新児童手当）と推移し、長期的には 6 万円だけ支給額が増加している[3]。同様に、中学生については旧児童手当の対象ではなかったことから、各種改革によって 12 万円支給額が増加している。この給付額の推移に加えて、分析する上で重要な点として、財源として住民税と所得税の扶養控除が廃止された。この控除廃止の影響は「子ども手当」による給付増を概ね相殺しており、所得税分の控除廃止（2011 年 1 月以降）と住民税分の控除廃止（2012 年 4 月以降）が完了する 2013 年以降についてはネットの給付増はほとんどないことが知られている (是枝, 2011)。例えば、所得税率が 5% であれば控除の廃止による負担増は所得税の扶養控除分で 1.9 万円（38 万円 × 0.05)、住民税で 3.3 万円（33 万円 × 0.1）の合計 5.2 万円となることから、児童手当が 6 万円拡大されても給付の純増は年間 8000 円程度に縮小する。さらに、高い所得税率に直面している中高所得層にとって、扶養控除廃止による所得増税の影響が強くなるため、改革による給付額の増加は全くないか、むしろマイナスになった (土居, 2010; 鈴木, 2011)。

一方、控除が廃止されるタイミングが給付拡大に遅行したため、2010 年と 2011 年については大幅に現金給付が増加した。特に、2012 年 4 月以前は 3.3 万円分の住民税の控除が存続していたことから、低所得者を中心に大きな給付増があったと見なせる。例えば、2011 年 1 月時点で考えると、給付額は年間 9.6 万円（15.6 – 6）増加する一方、増税は所得税の扶養控除分の 1.9 万円となることから年間で子ども 1 人あたり 7.7 万円の給付増が見込まれた。

以上のように、「子ども手当」の恩恵を大きく受けたのは扶養控除廃止の影響が小さかった低所得層のみであったと考えられる。また、給付規模の拡大は永続的ではなく、住民税分の扶養控除が廃止される 2012 年 4 月以前の期間でにおいてのみ、比較的大きな流動性の付与をもたらした。本稿ではそうした制度的背景を考慮して、分析対象を低所得世帯に限定し、改革直後の影響に着目した。

[1]例えば、宝くじ (Lindahl, 2005)、やドイツの統合 (Frijters et al., 2005)、19 世紀のワインの不作 (Banerjee et al., 2010) といったケースで分析されている。

[2]野口 (2008) 及び野口 (2011) は所得と健康の因果関係に対して明示的に操作変数を用いて識別する方法をとった数少ない我が国での研究である。ただし野口 (2011) では、操作変数が弱相関である可能性が排除できないことから、「第一段階での予測値が第二段階での IV として適切であるかについては，議論のあるところである。」としている。

[3]新児童手当以前に、「平成 23 年度における子ども手当の支給等に関する特別措置法」に基づき支給額が調整されている。この「特別措置法に基づく子ども手当」には「子ども手当」と同様に所得制限がなかったが、支給額は後の「新児童手当」と変わらない。更に、この法律には 2012 年 6 月から所得制限を復活させることが明記された。

3 分析方法

3.1 推定式

児童手当が家計に与える影響を識別する際には、児童手当の支給額が子どもの年齢と数によって決まっていることを考慮する必要がある。そのため、制度変更によって受給額が変動する場合にはじめて、児童手当の効果と子どもの年齢・数の効果を別々に識別することが可能となる (宇南山, 2011)。そこで、最初に受給する児童手当の額（$Cbenefit_{ht}$）がどのように決定されるか整理したい。前述の通り、児童手当は子どもの年齢に対して単価が決まっているため、世帯 h に含まれる i 歳の子どもの t 年における数と N_{iht} と表記し、その年齢に対応する児童手当の支給額を α_{it} とすると、世帯 h が受け取る児童手当の金額は、

$$Cbenefit_{ht} = \sum_{i=1}^{15} \alpha_{it} N_{iht} \tag{1}$$

と書ける。この算定式の中で、α_{it} の変動によってもたらされる児童手当額の変化に着目して、児童手当が両親の健康に与える影響を識別することが本稿の分析モデルである。つまり、N_{iht} をコントロールしてもなお $Cbenefit_{ht}$ が健康に影響を与えるとすれば、それは制度改正による受給額の変動の効果だと考えられる。

一方、本稿の分析デザインでは、この児童手当の増加に加えて所得控除廃止の影響も考慮しなければならない。現行の所得税制では 5% から 40% の累進性となっており、所得税率が増加するほど扶養控除の廃止に伴う増税額も増加する。しかし、高所得層では現金給付による主観的健康への影響があるとは想定しにくいことから、本稿の分析は所得税率が制度改正前年に 5% だったと考えられるサンプルに限定した [4]。このサンプルでは、制度改正後の増税額は児童 1 人あたり年額 1.9 万円で等しくなっている。このように、本稿では所得階層間の純受給額の違いではなく、低所得者層における世帯構成の違いに基づく純受給額の違いを用いて、児童手当の効果を識別した。中低所得者にサンプルを限定し、そのサンプル内での受給額の違いを用いて現金給付の効果を識別する点に関しては、米国の EITC の効果について検証した Hoynes et al. (2012) や Evans and Garthwaite (2010) とも共通する。

また、分析に際しては、児童手当受給額や所得控除をめぐる政策決定に関する期待形成の扱いが重要になる。特に、この期間では政策変更が年単位行われているため、調査時点の制度に忠実に純受給額を算出するか、予定された政策変更を織り込むかで算出された受給額に若干の変動がある。この点について、本稿では政策変更が極めて頻繁だったことと政治環境が流動的だったことから、将来の制度変更は分析に織り込まず、調査時点の制度に忠実に給付額（増税額）を算出し推計に用いた [5]。

[4] 所得税率は課税所得が 195 万円以下の場合に税率が 5% となるが、JHPS には正確な個人単位の課税所得のデータがない。そこで、「主な仕事からの収入（税引き前）」という調査項目を利用し、平均的な控除項目を指し引いた後の課税所得が 195 万円以下となるような税引き前所得として、世帯で最も高い所得を得ている者の「主な仕事からの収入（税引き前）」が 500 万円以下である世帯を対象とした。具体的には、基礎控除 (33 万円)、給与所得控除 (給与収入の 30% 程度)、配偶者控除（38 万円)、社会保険料控除（給与収入の 10% を想定すると、500 万円の年収で課税所得は 200 万円程度となる。

[5] この処理によって、例えば 2012 年 1 月調査では 2012 年 6 月徴収分からの住民税増税が予想されていた可能性があるが、一貫性を保つためにそうした将来の住民税増税については考慮していない。

以上の考えのもと、児童手当受給額から扶養控除の廃止に伴う増税分 (d_{ht}) を差し引いた後の児童手当の純受給額 ($NetCbenefit_{ht}$) は以下のように計算された。

$$NetCbenefit_{ht} = \begin{cases} \sum_{i=1}^{15} \alpha_{it} N_{iht} & if \;\; t \leq 2010 \\ \sum_{i=1}^{15} (\alpha_{it} - 1.9) N_{iht} & if \;\; t \geq 2011 \end{cases}$$

この式において、2011 年 1 月調査以降 $1.9 \times \sum_{i=1}^{15} N_{iht}$ 万円が差し引かれているのは、2011 年 1 月の改正によって児童 1 人あたり 1.9 万円分所得税が増税になったためである。受給額がこのように変更されても、推計の基本的考え方には影響がない。すなわち、N_{iht} をコントロールしてもなお $NetCbenefit_{ht}$ が健康に影響を与えるとすれば、それは一連の制度改正による受給額の変動の効果だと見なすことができる。なお、この分析モデルでは子どもの年齢・数（N_{iht}）は外生として扱われている。児童手当が出生行動に影響を与える場合そうした仮定が満たされない可能性もあるが、政策が流動的であったことなどを考えると、出生行動に大きな影響があったとは想定しにくく、妥当な仮定と言えるだろう。

以上の考察に基づき、2010 年 6 月の「子ども手当」支給開始前後のデータを用いて、次式を OLS 推定した。

$$Health_{it} = \beta_0 + \delta NetCbenefit_{ht} + \sum_{i=1}^{15} \beta_{it} N_{iht} + X_{it}'\gamma + \theta_i + \eta_t + \epsilon_{it} \tag{2}$$

ただし、β_0 は定数項、δ は児童手当がアウトカム変数に与える影響、β_{it} は N_{iht} の係数、θ_i は個人固定効果、η_t は調査年に固有の効果、ϵ_{it} は誤差項である。また、経時的に変化する属性をコントロールするために、説明変数として世帯員数、15 歳未満の子どもの数、子どもの通う学校種別（幼稚園、保育園、小学校、中学校)、昨年一年間における転居の有無を推計に加えた。さらに、本人の健康状態と関連の深い変数として、飲酒状況、喫煙状況、昨年一年間における入院の有無を加えた。飲酒、喫煙、入院に関しては、所得状況との相関関係も大きいと考えられることから、推計式に含める場合と含めない場合で児童手当純受給額の係数が変化するか確認した。また、本論文で用いる被説明変数には数段階の回答となっているものがあるが、そうした場合にも OLS による推定を行い、個人固定効果を明示的にコントロールした。

3.2 健康指標

次に本稿で主要な分析対象となる健康や生活満足度に関する指標について説明する。

3.2.1 主観的健康

主観的健康（Self-rated health :SRH）については「ふだんのあなたの健康状態はどうですか」という質問項目を利用した。回答は「よい (1)」「まあよい (2)」「ふつう (3)」「あまりよくない (4)」「よくない (5)」の 5 段階でなされ、数値が高いほど健康状態が悪いことを示している。このうち、本稿では 5 段階の質問

項目そのものだけでなく、上（下）の二つを 1 、他を 0 とするダミー変数を作成し分析している。主観的健康の決定要因は多くの研究で分析されているが、その理由は主観的健康がより客観的な健康指標（死亡率など）の信頼できる代理指標と考えられる点にある (Idler and Benyamini, 1997)[6]。本稿でも同様の考え方から、主観的健康の分析を通して、より客観的な健康指標への影響に関しても示唆が得られると期待している。

3.2.2　心身症状指標、及び生活満足度・充実度指標

JHPS では主観的健康の他に心身の自覚症状を把握するために Ben-sira (1982) を参考に、「頭痛やめまいがする」「動悸や息切れがする」「胃腸の具合がおかしい」「背中・腰・肩が痛む」「疲れやすい」「風邪をひきやすい」「イライラする」「寝つきが悪い」から成る 8 つの質問を行っている (石井, 2012)[7]。この心身症状指標の各項目は「よくある (1)」「ときどきある (2)」「ほとんどない (3)」「全くない (4)」の 4 段階で回答され、数字が高いほど望ましい状態にあることを表す。総合指標は 8 つの指標を合計した 8 から 32 までの値を取る指数であり、広く自覚症状の状態を把握できると考えられる。

以上の健康に関する質問の他に、よりダイレクトに生活への満足度に対する効果を分析するために、生活満足度や充実度に関連すると思われる項目について分析した。具体的には、「人と会うのがおっくうだ」「仕事への集中度がない」「今の生活に不満がある」「将来に不安を感じる」の 4 項目であり、これらの指数を合計した総合指標による分析も行った。回答方式は心身症状指標と同様であり、総合指標は 4 から 16 までの値を取り、数字が高いほど望ましい状態であることを表す。

3.3　データ

データは慶應義塾大学パネルデータ設計・解析センターが実施している日本家計パネル調査（Japan Household Panel Survey）の 2012 年版を用いた。JHPS は米国の Panel Study of Income Dynamics (PSID) や欧州の European Community Household Panel (ECHP) 等を参考に、特定の層に焦点を当てるのではなく、社会全体の人口構成を反映した家計パネル調査として設計されており、回答が得られなかったサンプルについても調査区と年齢群と性別から適合する予備調査群を設けるなど、サンプルの代表性に配慮がなされている。第 1 回調査である JHPS2009 では配偶者を含む 6,911 人に対して調査を行い、JHPS2012 ではそのうちの約 70 %にあたる 4,903 に調査が行われている。本稿では、そのサンプルから配偶者のいないものを除き [8]、「子ども手当」支給開始前の 2010 年 1 月調査において子どもが 1 人から 3 人いる世帯に対象を限定した。これらの世帯は政権交代による児童手当の増額の影響を受けたと考えられるが、2011 年 1 月調査以降に子どもが生まれた世帯については、旧児童手当の受給期間がないことからサンプルから除い

[6]ただし主観的健康には様々な問題が指摘されている。我が国でも、野口 (2011) では、様々な先行研究をサーベイし主観的健康の問題点をまとめている。

[7]ただし、調査対象に対して JHPS は現在の状況を聞いているのに対して、Ben-sira (1982) では過去一年間における状況を聞いており、両者には本質的な違いがあると考えられる。むしろ JHPS のこれらの調査項目は、国民生活基礎調査の健康票における「あなたはここ数日、病気や怪我で具合の悪いところ（自覚症状）がありますか」という質問に近いだろう。

[8]母子・父子家庭が調査対象に少なかったことから、分析を配偶者のいる世帯に限定した。

た。また、サンプルを中低所得層に限定するために、最多所得者の年収が 500 万円以下の世帯に分析を限定した。以上のサンプルについて欠損処理を行った結果、利用可能な対象者は配偶者を含めて 631 人、観測値は最大で 2327 となった。なお、サンプルの記述統計は表 6 にまとめている。

全てのパネル調査と同様に JHPS にも脱落（attrition）の問題が存在する。特に、本稿と関連するところでは赤林・野崎・敷島 (2013) によると、健康水準の低いものの継続率が第 3 回調査で低い点が指摘されている。そこで、本稿では Unbalanced-Panel に基づく推計をメインとしながら、別途 Balanced-Panel を作成し結果の頑健性を確認し、主要な結論には影響がないことを確かめた [9]。

4　分析結果

4.1　主観的健康

分析結果は表 2 にまとめた。まず、(1) 列から (6) 列までは父親の、(7) 列から (12) 列までは母親のサンプルにおける推定結果を掲載している。推定は、5 段階の主観的健康を被説明変数とした OLS 推定、健康状態が「よい」もしくは「悪い」かどうかの決定要因を分析した OLS 推定にわかれる。各々の被説明変数に対して推定結果は 2 種類掲載しており、はじめの推定は子ども数など必要最低限の変数を投入した場合の推定結果、もう一つは健康状態に影響を与えると思われる飲酒や喫煙の状況、そして入院経験を追加的にコントロールした場合の推定結果である。

まず、父親の推定結果をみると、5 段階の主観的健康や健康状態が「よい」と回答するかを分析した (1) 列から (4) 列における児童手当純受給額の係数はいずれも有意ではない。一方、健康状態が「悪い」と答えるかどうかを分析した (5) 列と (6) 列の推定結果は、いずれも有意に負となっている。推定値は (6) 列で - 0.006 であり、年間 10 万円の児童手当純受給額の増加に対して、健康状態が「悪い」と答える確率が 6% 低下すると解釈できる。

一方、母親の推定結果をみると、健康状態が「よい」かどうかを分析した (9) 列と (10) 列において、児童手当純受給額の係数が有意に正となっている。係数値はどちらも 0.009 となり、児童手当の年間 10 万円の増加に対して健康状態が「よい」と回答する確率が 9% 増加すると解釈できる。既婚の女性のサンプルについて、現金給付の増加が心理的健康を引き上げるという結論は、米国の EITC の効果を検証した Evans and Garthwaite (2010) や Boyd-Swan et al. (2013) と整合的であり [10]、我が国の児童手当でも同様の効果が観察できたと言えるだろう。ただし、全ての変数で一貫して有意な効果が得られているわけではなく、母親と父親の効果の異質性など、説明が容易でない点もあることに注意する必要がある [11]。

[9] データの作成方法は、制度変更前の 2010 年 1 月調査時点で子どもが 1 人以上 3 人以下いる 55 歳以下の男女にサンプルを限定し、分析に必要なデータが 4 年分すべてそろっている者のみを抽出した。その結果、491 人の 4 年間に渡るデータ（サンプル数は 1964）が作成された。推計結果については紙面の関係から省くが、リクエストがあれば筆者からお送りしたい。

[10] カナダの児童手当の効果を検証した Milligan and Stabile (2011) では母親の鬱指標に対する効果を確認した一方、主観的健康には影響がないとしている。米国の EITC の効果を検証した Evans and Garthwaite (2010) では既婚の母親について主観的健康を改善を報告している。また、同様に EITC の効果を分析した Boyd-Swan et al. (2013) は主観的健康について分析していないものの、鬱指標と主観的幸福感（self-reported happiness）の両方で母親に対する効果を確認している。

[11] 後述する心身症状指標に関する分析では、母親のサンプルでは「イライラ」の改善など心理的健康への影響が示唆されるのに対して、父親では「頭痛」や「風邪」などより身体的なアウトカムに効果が観察されている。それぞれの項目に関する考察は個別的に過ぎるため省略するが、仮に身体的に自覚症状がある場合に健康状態を「悪い」と回答すると考えるならば、主観的健康に関

その他の変数の係数についてみると、入院歴、喫煙、飲酒、転居のすべての変数について、一貫した健康状態との相関は確認できなかった。

4.2 心身症状指標

次に心身症状指標の推定結果は表 3 にまとめた。まず表 3 の (1)、(2)、(7)、(8) 列では、総合指標に対する児童手当純受給額の係数を報告している。結果を確認すると、父親のサンプルでは、(1) 列の係数は 0.074 で有意となっており、児童手当の増加によって心身症状が改善することを示している。この結果は、本人の入院歴や喫煙等をコントロールしても大きく変わらなかった（(2) 列）。しかし母親のサンプルでは総合指標に対する児童手当の効果は確認できなかった（(7)、(8) 列）。

次に、心身症状指標の内訳を確認する。まず、(3)、(4)、(9)、(10) 列では各質問項目に該当するかどうかに関する「よくある」から「全くない」までの 4 段階の回答を被説明変数としている。みると、父親では「頭痛やめまいがする」及び「風邪をひきやすい」といった質問項目に対して、児童手当の係数が有意に正と推定されている ((3)、(4) 列)。次に、4 段階の回答を 2 値変数に変換して推定するために、「ほとんどない」及び「ない」という回答を 1 とするダミー変数を作成し、OLS 推定を行った（(5)、(6)、(11)、(12)）。推定結果を確認すると、父親のサンプルではすべての推定で児童手当純受給額の係数は有意となっていないものの、母親では「イライラする」という質問項目に対して有意な効果が確認された。係数を解釈すると、児童手当純受給額の年間 10 万円の増加に対して母親が「イライラ」する確率は 10% 程度減少する。

4.3 生活満足度・充実度

次に、生活満足度に関連する質問項目を抽出し、それに対する効果を同様の方法で推定した。また心身症状指標の推計と同じく、4 段階の回答結果をそのまま被説明変数に用いた分析（(2)、(4) 列）、と 2 値変数への効果をみた分析（(3)、(6) 列）の両方の結果を掲載した。結果は表 4 にまとめた。まず、総合指標に対する効果を確認すると、父親 ((1) 列) 及び母親 ((2) 列) の両方で係数は正に推定され、生活に対する満足度や充実度が上昇していることが確認できた。

次に、内訳に対する効果を確認すると、父親のサンプルでは「仕事に集中できない」という質問に対する効果が確認され、母親では「今の生活に不満がある」という質問で児童手当の係数が有意となった。推定値を確認すると、児童手当の年額 10 万円の増加によって、父親が「仕事に集中できない」という質問を否定する確率は 9% 上昇し、母親が「今の生活に不満がある」と思わない確率を 13% 上昇させた。

4.4 生活習慣

次に飲酒や喫煙など生活習慣への影響も調べる。先行研究では現金給付の増加が健康にとって好ましくない消費活動を助長させる可能性を指摘しており、特に、嗜好品の原資として児童手当が用いられる場合、

する効果の違いもある程度説明されると考えられる。

多くの有権者は支給額が過大である可能性を指摘するだろう。近年の分析でも、政策的な現金給付がこうした好ましくない消費行動を喚起することが指摘されている。例えば、Gross and Tobacmanz (2013) では米国で 2008 年のリーマンショック後に給付された一時的な現金給付により、飲酒や薬物関連の入院が増加したと指摘している。また、児童手当に関しても Blow et al. (2012) では英国における政策的な児童手当の増額が、両親のアルコール消費量を増加させていることを指摘した。ただし嗜好品の消費に対する影響に関しては理論的に自明な関係があるわけではない。例えば、健康が正常財である場合、所得水準の上昇は健康への需要の拡大を通して嗜好品の消費を減少させるだろう (Grossman, 1972)。実証的にも米国の EITC が母親の喫煙に与える影響を分析した Averett and Wang (2012) では、EITC の給付額がより増加した母親で喫煙率が低下したと指摘している。

そこで、本節では健診の受診等も含めた、広く健康に影響する生活習慣への影響についてまとめた。被説明変数は、飲酒、喫煙、BMI、調査時点における人間ドック・健診・予防接種の受診、及び定期的な運動・ジム通い・サプリメントの摂取とした。推計結果は表 5 にまとめた。表 5 では児童手当純給付額の係数を父親・母親のべつに報告しているが、全ての推計で係数は有意ではなく、政策的な児童手当の増額はこれらの項目に対して影響しなかったことが分かった。特に、先行研究でも指摘されている飲酒への影響がなかった点も含めて、児童手当が好ましくない費目へ支出されたという証左は見つからなかった。

5　結論と議論

本稿では民主党政権下で導入された「子ども手当」とそれに伴う扶養控除の廃止が、両親の心理的健康にどのような影響を与えたか検証した。分析の結果、扶養控除分の増税を差し引いた児童手当の純給付額の外生的な増加は、両親の主観的健康や心身症状、及び生活に対する充実度や満足度を有意に改善させたことがわかった。児童手当純給付額の 10 万円の増加は、父親が自分の健康状態について「悪い」と答える確率を 6% 低下させ、母親が自分の健康状態について「良い」と答える確率を 9% 上昇させた。その他の変数については、必ずしも頑健で一貫した回答パターンの変化は示されなかったが、いくつかの質問項目で観察された効果を総合的に判断すると、児童手当の増額による心理的健康の改善を示唆していると考えられた。ただし、効果の大きさについては、推定結果が必ずしも頑健ではないことや、比較的小規模なサンプルによる分析であることも考慮し、さらに精緻な検証が必要といえる。また、飲酒や喫煙など健康状態と深く関連する消費行動に対する影響は全くなかった。

本稿の政策インプリケーションは 2 点ある。第一に、既存研究では児童手当に対して否定的な政策インプリケーションを得る研究が少なくなかったが (野口, 2008; 宇南山, 2011)、児童手当の好ましい効果を発見している点である。第二に、政策的な流動性の増加は健康を害する消費行動や倫理的に望ましくない消費行動を喚起することが指摘されてきたが、本稿の推定結果は「子ども手当」に関してそうした影響はなかったことを示している。給付水準が過大になるとこうした消費行動が誘発されるという立場に立つならば、本稿の分析結果は現行の児童手当の給付水準が過大とは言えないことを示唆しているだろう。以上の点に留意しつつ、児童手当の持つ様々な効果に十分に配慮して、政策的決定が行われる必要がある。

なお、本稿では子どもの成長を規定する重要な因子として両親の心理的健康に焦点を当てたが、よりダイレクトに子ども本人のアウトカム（健康、学力）に対する影響を精査する必要がある。その点は今後の課題としたい。また、本稿の限界として、子どもの数を児童手当の拡充に対して外生と仮定している点が挙げられる。「子ども手当」は控除の廃止を伴っており必ずしも家計に長期的な恩恵をもたらしたとはいえないことから、出生行動に対して大きな影響があったとは想定しにくいが、その点に関する実証的確認も必要だろう。

参考文献

Averett, Susan, and Yang Wang. 2012. "The Effects of Earned Income Tax Credit Payment Expansion on Maternal Smoking." *Health Economics*, 22 1344–1359.

Banerjee, Abhijit, Esther Duflo, Gilles Postel-Vinay, and Tim Watts. 2010. "Long-Run Health Impacts of Income Shocks: Wine and Phylloxera in Nineteenth-Century France." *The Review of Economics and Statistics*, 92(4): 714–728.

Becker, Gary S. 1981. *A Treatise on the Family.* Cambridge, MA: Harvard University Press.

Ben-sira, Zeev. 1982. "The Scale of Psychological Distress(SPD): Cross Population Invariance and Validity." *Research Communications in Psychology, Psychiatry and Behavior*, 7 329–346.

Blow, Larura, Ian Walker, and Yu Zhu. 2012. "Who Benefits From Child Benefit?" *Economic Inquiry*, 50(1): 153–170.

Boyd-Swan, Casey, Chris M. Herbst, John Ifcher, and Homa Zarghamee. 2013. "The Earned Income Tax Credit, Health, and Happiness." IZA Discussion Papers 7261, Institute for the Study of Labor (IZA).

Case, Anne, and Christina Paxson. 2002. "Parental Behavior And Child Health." *Health Affairs*, 21(2): 164–178.

Dahl, Gordon B., and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review*, 102(5): 1927–56.

Deaton, Angus. 2003. "Health, Inequality, and Economic Development." *Journal of Economic Literature*, 41(1): 113–158.

Evans, William N., and Craig L. Garthwaite. 2010. "Giving Mom a Break: The Impact of Higher EITC Payments on Maternal Health." NBER Working Papers 16296, National Bureau of Economic Research, Inc.

Frijters, Paul, John P. Haisken-DeNew, and Michael A. Shields. 2005. "The causal effect of income on health: Evidence from German reunification." *Journal of Health Economics*, 24(5): 997–1017.

Gershoff, Elizabeth T., J. Lawrence Aber, C. Cybele Raver, and Mary Clare Lennon. 2007. "Income Is Not Enough: Incorporating Material Hardship Into Models of Income Associations With Parenting and Child Development." *Child Development*, 78(1): 70–95.

Grossman, Michael. 1972. "On the Concept of Health Capital and the Demand for Health." *Journal of Political Economy*, 80(2): 223–55.

Gross, Tal, and Jeremy Tobacmanz. 2013. "Dangerous Liquidity and the Demand for Health Care: Evidence from the 2008 Stimulus Payments." *Journal of Human Resources (2013).*

Hoynes, Hilary W., Douglas L. Miller, and David Simon. 2012. "Income, the Earned Income Tax Credit, and Infant Health."(18206): .

Idler, Ellen L., and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior*, 38(1): pp. 21–37.

Lindahl, Mikael. 2005. "Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income." *Journal of Human Resources*, 40(1): 144–168.

Milligan, Kevin, and Mark Stabile. 2011. "Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions." *American Economic Journal: Economic Policy*, 3(3): 175–205.

Mistry, Rashmita S., Elizabeth A. Vandewater, Aletha C. Huston, and Vonnie C. McLoyd. 2002. “Economic Well-Being and Children’s Social Adjustment: The Role of Family Process in an Ethnically Diverse Low-Income Sample.” *Child Development*, 73(3): 935–951.

Propper, Carol, John Rigg, and Simon Burgess. 2007. “Child health: evidence on the roles of family income and maternal mental health from a UK birth cohort.” *Health Economics*, 16(11): 1245–1269.

Strully, Kate W., David H. Rehkopf, and Ziming Xuan. 2010. “Effects of Prenatal Poverty on Infant Health: State Earned Income Tax Credits and Birth Weight.” *American Sociological Review*, 75(4): 534–562.

Wyman, Peter, Jan Moynihan, Shirley Eberly, and Christopher Cox. 2007. “Association of family stress with natural killer cell activity and the frequency of illnesses in children.” *Archives of Pediatrics and Adolescent Medicine*, 161(3): 228–234.

Yeung, W. Jean, Miriam R. Linver, and Jeanne Brooks-Gunn. 2002. “How Money Matters for Young Children’s Development: Parental Investment and Family Processes.” *Child Development*, 73(6): 1861–1879.

宇南山, 卓. 2011. “児童手当が家計消費に与えた影響.” *RIETI Discussion Paper Series*, 11-J-021.

小林, 淑恵. 2010. “児童手当の家計への影響.” 季刊社会保障研究, 47(1): 67–80.

是枝, 俊悟. 2011. “新旧児童手当と子ども手当の比較分析.” *Legal Tax Report (*大和総研*).*

石井, 加代子. 2012. “JHPS2011 の概況.” *Joint Research Center for Panel Studies Discussion Paper Series.*

赤林, 英夫, 野崎 華世, 敷島 千鶴. 2013. “JHPS2012 調査の概況.” *Joint Research Center for Panel Studies Discussion Paper Series*(DP2012-006): .

田中, 慶子. 2008. “児童手当の拡充と子育て世帯の家計.” 家計経済研究, 80 39–44.

土居, 丈朗. 2010. “子ども手当導入に伴う家計への影響分析-JHPS を用いたマイクロシミュレーション.” 経済研究 (2): 137–158.

野口, 晴子. 2008. “世帯の経済資源が出産・育児期における女性の心理的健康に与える影響について　ー「消費生活に関するパネル調査」を用いた実証分析.” 経済研究, 59(3): 209–227.

野口, 晴子. 2011. “社会的・経済的要因と健康との因果性に対する諸考察-「社会保障実態調査」および「国民生活基礎調査」を用いた実証分析.” 季刊社会保障研究, 46(4): 382–402.

両角, 良子. 2009. “被服消費に着目した地域振興券のラベリング効果の検証-児童手当へのインプリケーション-.” 経済学研究, 58(4): 101–115.

鈴木, 克洋. 2011. “子どもに対する手当の増額と年少扶養控除廃止の影響　-世帯構成別及び所得別の影響試算.” 経済のプリズム, 96.

表 1: 児童手当と子ども手当の比較

	旧児童手当	子ども手当	特別措置の子ども手当	新児童手当
A.1 人あたり月支給額				
3 歳未満	1	1.3	1.5	1.5
3 歳以上小学生以下	0.5	1.3	1	1
中学生	0	1.3	1	1
B.1 人あたり年間支給額				
3 歳未満	12	15.6	18	18
3 歳以上小学生以下	6	15.6	12	12
中学生	0	15.6	12	12
政権党	自民党	民主党	民主党	民主党
実施時期	〜2010 年 3 月分	2010 年 4 月	2011 年 10 月	2012 年 4 月
所得制限	あり	なし	なし	あり

注：所得制限は夫婦と児童 2 人のケース。単位は万円。
資料：是枝 (2011) 及び厚生労働省資料から筆者作成

表 2: 主観的健康への影響

	父親						母親					
	5 段階指標		健康状態が良い		健康状態が悪い		5 段階指標		健康状態が良い		健康状態が悪い	
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
児童手当純受給額	-0.010	-0.010	0.000	0.001	-0.007***	-0.006**	-0.013	-0.011	0.009*	0.009*	-0.001	-0.001
	(0.010)	(0.010)	(0.006)	(0.005)	(0.003)	(0.003)	(0.012)	(0.012)	(0.005)	(0.005)	(0.004)	(0.004)
入院歴		0.243		-0.116		0.108*		-0.062		-0.047		-0.018
		(0.210)		(0.087)		(0.061)		(0.171)		(0.081)		(0.036)
喫煙ダミー 2 (毎日吸う)												
レファレンス												
喫煙ダミー 2 (ときどき吸う)		0.226		-0.201*		-0.058		0.209		-0.1		0.032
		(0.230)		(0.108)		(0.061)		(0.241)		(0.127)		(0.062)
喫煙ダミー 3 (やめた)		0.049		-0.002		0.007		0.119		-0.026		0.061*
		(0.151)		(0.071)		(0.033)		(0.163)		(0.094)		(0.033)
喫煙ダミー 4 (以前から吸わない)		-0.014		0.038		-0.038		0.155		-0.117		0.084*
		(0.272)		(0.125)		(0.065)		(0.201)		(0.099)		(0.043)
飲酒ダミー 1 (全く飲まない)												
レファレンス												
飲酒ダミー 2 (月に数回飲酒)		-0.002		-0.021		-0.008		-0.261**		0.168***		0.022
		(0.132)		(0.062)		(0.029)		(0.106)		(0.042)		(0.025)
飲酒ダミー 3 (週に 1〜2 回)		-0.114		0.008		-0.031		-0.118		0.088		0.038
		(0.179)		(0.084)		(0.042)		(0.160)		(0.067)		(0.040)
飲酒ダミー 4 (週に 3 回以上)		-0.367*		0.101		-0.07		-0.073		0.066		0.017
		(0.215)		(0.094)		(0.052)		(0.140)		(0.072)		(0.040)
転居		-0.113		0.024		-0.048**		0.129		-0.065		0.011
		(0.110)		(0.054)		(0.024)		(0.126)		(0.051)		(0.031)
子どもの学校種別	X	X	X	X	X	X	X	X	X	X	X	X
年齢別の子どもの人数	X	X	X	X	X	X	X	X	X	X	X	X
子どもの数	X	X	X	X	X	X	X	X	X	X	X	X
世帯員数	X	X	X	X	X	X	X	X	X	X	X	X
調査年	X	X	X	X	X	X	X	X	X	X	X	X
個体効果	X	X	X	X	X	X	X	X	X	X	X	X
R2	0.619	0.621	0.047	0.064	0.043	0.062	0.627	0.629	0.024	0.05	0.026	0.035
個人数	314	314	314	314	314	314	317	317	317	317	317	317
観測値	1,165	1165	1165	1165	1165	1165	1,162	1162	1162	1162	1162	1162

注：カッコ内は分散不均一に対して頑健な標準誤差。主観的健康指標は「ふだんの健康状態」が良い場合に 1 を、悪い場合に 5 を取る 5 段階の変数。(3)、(4) 列、及び (9)、(10) 列では健康状態が「まぁ良い」または「良い」場合に 1 をとるダミー変数を被説明変数としている。(5)、(6) 列、及び (11)、(12) 列では健康状態が「あまり良くない」または「悪い」場合に 1 をとるダミー変数を被説明変数としている。コントロール変数には、調査年ダミーと個体効果の他に、世帯員数、各歳別の子どもの数、子どもの学校種別、追加的なコントロール変数として転居の有無、飲酒、喫煙、入院歴を加えた。***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

表 3: 心身症状指標への影響

	父親						母親					
	総合指標		4 段階指標		ない = 1		総合指標		4 段階指標		ない = 1	
	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
心身症状指標	0.074**	0.068**					0.051	0.054				
	(0.034)	(0.034)					(0.035)	(0.036)				
頭痛やめまいがする			0.014*	0.013*	0.000	-0.001			0.005	0.006	0.000	0.000
			(0.008)	(0.008)	(0.004)	(0.004)			(0.009)	(0.009)	(0.004)	(0.004)
動悸や息切れがする			0.008	0.007	-0.002	-0.003			-0.004	-0.003	-0.004	-0.004
			(0.007)	(0.007)	(0.003)	(0.003)			(0.008)	(0.007)	(0.003)	(0.003)
胃腸の具合がおかしい			0.003	0.003	0.003	0.002			0.006	0.007	0.001	0.002
			(0.008)	(0.008)	(0.004)	(0.004)			(0.008)	(0.009)	(0.004)	(0.004)
背中・腰・肩が痛む			0.006	0.006	-0.001	(0.001)			0.010	0.011	0.008*	0.007*
			(0.010)	(0.010)	(0.005)	(0.004)			(0.010)	(0.011)	(0.004)	(0.004)
疲れやすい			0.011	0.011	0.001	0.000			0.003	0.004	0.001	0.001
			(0.011)	(0.011)	(0.005)	(0.005)			(0.009)	(0.010)	(0.004)	(0.004)
風邪をひきやすい			0.016*	0.015*	0.004	0.004			0.014	0.014	0.006	0.006
			(0.009)	(0.009)	(0.004)	(0.004)			(0.009)	(0.009)	(0.005)	(0.005)
イライラする			0.012	0.011	-0.006	0.004			-0.001	-0.001	0.011**	0.010**
			(0.008)	(0.008)	(0.004)	(0.004)			(0.010)	(0.010)	(0.004)	(0.004)
寝つきが悪い			0.005	0.005	0.004	0.003			-0.002	-0.002	0.000	0.000
			(0.008)	(0.007)	(0.004)	(0.004)			(0.009)	(0.009)	(0.004)	(0.005)
本人の入院歴		×		×		×		×		×		×
本人の喫煙		×		×		×		×		×		×
本人の飲酒		×		×		×		×		×		×
転居		×		×		×		×		×		×
子どもの学校種別	×	×	×	×	×	×	×	×	×	×	×	×
年齢別の子どもの人数	×	×	×	×	×	×	×	×	×	×	×	×
世帯員数	×	×	×	×	×	×	×	×	×	×	×	×
調査年	×	×	×	×	×	×	×	×	×	×	×	×
個体効果	×	×	×	×	×	×	×	×	×	×	×	×
個人数	314	314	314	314	314	314	317	317	314	314	317	317
観測値	1,165	1,165	1,165	1,165	1,165	1,165	1,162	1,162	1,165	1,165	1,162	1,162

注：係数は全て、児童手当の純受給額 (万円) に対する係数。カッコ内は分散不均一に対して頑健な標準誤差。心身症状指標における総合指標は「頭痛がする」」「動悸や息切れがする」「胃腸の具合がおかしい」などの 8 項目に対して「よくある」「ときどきある」「ほとんどない」「全くない」の 4 段階で回答し、その指数を合計したもの。インデックスの値が高いほど心身症状がないことを表す。(3)、(4) 列、及び (9)、(10) 列では個別の項目について ordered-probit 推計を行っている。(5)、(6) 列、及び (11)、(12) 列では「ほとんどない」及び「全くない」場合に 1 をとるダミー変数を作成し、OLS 推計している。調査年ダミーと個体効果の他に、世帯員数、各歳別の子どもの数、子どもの学校種別、転居の有無、飲酒、喫煙、入院歴を加えた。***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

表 4: 生活充実度・満足度への影響

	父親			母親		
	総合	4 段階	ない＝1	総合	4 段階	ない＝1
	OLS	OLS	OLS	OLS	OLS	OLS
	(2)	(4)	(6)	(8)	(10)	(12)
総合指標	0.038*			0.035*		
	(0.022)			(0.020)		
人と会うのが億劫になった		0.005	-0.001		-0.002	-0.002
		(0.008)	(0.003)		(0.009)	(0.004)
仕事に集中できない		0.023***	0.009**		0.01	0.005
		(0.009)	(0.004)		(0.010)	(0.004)
今の生活に不満がある		0.003	-0.001		0.018**	0.013***
		(0.010)	(0.005)		(0.008)	(0.004)
将来に不安を感じる		0.007	0.003		0.009	0.006
		(0.010)	(0.005)		(0.009)	(0.004)
本人の入院歴	×	×	×	×	×	×
本人の喫煙	×	×	×	×	×	×
本人の飲酒	×	×	×	×	×	×
転居	×	×	×	×	×	×
子どもの学校種別	×	×	×	×	×	×
年齢別の子どもの人数	×	×	×	×	×	×
世帯員数	×	×	×	×	×	×
調査年	×	×	×	×	×	×
個体効果	×	×	×	×	×	×
個人数	314	314	314	317	317	317
観測値	1,165	1,165	1,165	1,162	1,162	1,162

注 : 係数は全て、児童手当の純受給額 (万円) に対する係数。カッコ内は分散不均一に対して頑健な標準誤差。各質問に対して「よくある」「ときどきある」「ほとんどない」「全くない」の 4 段階で回答し、数字が高いほど望ましい状態であることを示す。総合指標は 4 つの質問の合計値。コントロール変数には、調査年ダミーと個体効果の他に、世帯員数、各歳別の子どもの数、子どもの学校種別、転居の有無、飲酒、喫煙、入院歴を加えた。***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

表 5: 生活習慣等への影響

	父親	母親
	(1)	(2)
週 3 日以上の飲酒（= 1）	-0.002	-0.001
	(0.003)	(0.005)
全く飲まない（= 1）	0.002	-0.002
	(0.003)	(0.002)
喫煙（= 1）	-0.004	-0.003
	(0.003)	(0.003)
BMI	0.015	0.004
	(0.015)	(0.014)
人間ドック・健診・予防接種の受診（= 1）	-0.007	0.004
	(0.004)	(0.004)
運動・ジム通い・サプリメントの摂取（= 1）	0.005	-0.002
	(0.004)	(0.005)
転居	X	X
子どもの学校種別	X	X
年齢別の子どもの人数	X	X
世帯員数	X	X
調査年効果	X	X
個体効果	X	X

注 : 係数は全て、児童手当の純受給額 (万円) に対する係数。カッコ内は分散不均一に対して頑健な標準誤差。コントロール変数には、調査年ダミーと個体効果の他に、世帯員数、各歳別の子どもの数、子どもの学校種別、転居の有無を加えた。***, $p < 0.01$. **, $p < 0.05$. *, $p < 0.1$.

A　記述統計量

表 6: 記述統計量

大項目	小項目	Obs	平均	標準偏差	最小値	最大値
児童手当	児童手当純受給額（万円）	2327	16.14	10.85	0	44.1
健康指標	主観的健康	2327	2.17	0.96	1	5
	健康状態が良い（=1）	2327	0.60	0.49	0	1
	健康状態が悪い（=1）	2327	0.07	0.26	0	1
	心身症状指標	2327	21.93	5.03	8	32
生活満足度・充実度	総合指標	2327	11.05	2.79	4	16
	人と会うのが億劫になった	2327	0.76	0.43	0	1
	仕事への集中力がなくなった	2327	0.79	0.41	0	1
	今の生活が不満がある	2327	0.50	0.50	0	1
	将来に不安を感じる	2327	0.39	0.49	0	1
生活習慣	週 3 日以上の飲酒（＝1）	2327	0.40	0.49	0	1
	全く飲まない（＝1）	2327	0.25	0.43	0	1
	喫煙（＝1）	2327	0.36	0.48	0	1
	人間ドック・健診・予防接種の受診（＝1）	2276	0.29	0.45	0	1
	運動・ジム通い・サプリメントの摂取（＝1）	2252	0.26	0.44	0	1
	BMI	2063	22.26	3.47	14.8	36.8
入院歴	入院歴	2327	0.03	0.18	0	1
喫煙	毎日	2327	0.34	0.47	0	1
	ときどき吸う	2327	0.03	0.16	0	1
	今は吸わない	2327	0.21	0.41	0	1
	以前から吸わない	2327	0.43	0.50	0	1
飲酒	全く飲まない	2327	0.36	0.48	0	1
	月に数回	2327	0.40	0.49	0	1
	週に 1 〜2 回	2327	0.25	0.43	0	1
	週に 3 回以上	2327	0.10	0.30	0	1
転居	転居	2327	0.08	0.27	0	1
子どもの通う学校	保育園	2327	0.26	0.56	0	3
	幼稚園	2327	0.15	0.37	0	2
	小学校	2327	0.63	0.72	0	3
	中学校	2327	0.25	0.48	0	2
世帯人数	世帯人数	2327	4.34	1.24	1	9
年齢別子どもの数	0 歳	2327	0.01	0.11	0	1
	1 歳	2327	0.09	0.29	0	2
	2 歳	2327	0.12	0.33	0	2
	3 歳	2327	0.13	0.33	0	1
	4 歳	2327	0.14	0.34	0	1
	5 歳	2327	0.13	0.34	0	1
	6 歳	2327	0.12	0.33	0	1
	7 歳	2327	0.12	0.32	0	1
	8 歳	2327	0.10	0.30	0	1
	9 歳	2327	0.11	0.31	0	1
	10 歳	2327	0.10	0.30	0	1
	11 歳	2327	0.11	0.31	0	1
	12 歳	2327	0.11	0.31	0	2
	13 歳	2327	0.10	0.29	0	1
	14 歳	2327	0.09	0.29	0	1
	15 歳	2327	0.08	0.27	0	2

Chapter 6

Hospital Response to Financial Incentives: Evidence From Nighttime EMS in Japan*

Reo TAKAKU†

Abstract

In many countries, reimbursement for hospital care is linked to the number of patient bed days, where a "day" is defined as the period from one midnight to the next. This "midnight-to-midnight" definition may incentivize health care providers to manipulate hospital arrival times in emergency patients, as patients admitted before midnight would have an additional day for reimbursement when compared with those admitted after midnight. Using administrative data of emergency transportations in Japan, we find significant manipulation of hospital arrival time. In addition, the manipulation was generally observed in prefectures where emergency medical services are provided mainly by private hospitals.

Keywords : Hospital arrival time; Manipulation; Emergency medical services; Bunching

JEL classification : I10, I13

*We thank the Fire and Disaster Management Agency of Japan for allowing access to administrative records of emergency transportations. Although this project was supported by the Institute for Health Economics and Policy, the views expressed in this paper do not necessarily reflect the official views of the institute. We gratefully acknowledge the helpful suggestions and comments from Rei Goto at the annual meeting of the Japanese Economic Association and from Minah Kang at the Tri-Country Joint Health Economics Symposium held in Taipei. We also thank Michihito Ando, Shun-ichiro Bessho and Atsushi Yamaoka for their advice. All errors are our own.

†Corresponding Author, Institute for Health Economics and Policy, Japan. Email: reo.takaku@ihep.jp

1 Introduction

A central theme in health economics research is the understanding of how health care providers respond to financial incentives. Policymakers also strive to understand the effects of alternative payment systems, such as fee-for-service (FFS) and capitation payments (CAP), which may offer some control over the behavior of health care providers. There is a consensus that, in theory, FFS encourages providers to abuse health care resources, whereas CAP incentivizes the underprovision of care (Ellis and McGuire, 1986). However, the results from empirical studies on the impact of payment systems on physician behavior are mixed: although some studies have reported a strong link between payment systems and health care provision (Delattre and Dormont, 2003; Devlin and Sarma, 2008; Shafrin, 2010; van Dijk et al, 2013), others have found no such associations (Hadley et al, 1979; Grytten and Sorensen, 2001).[1] Recent experimental-based studies have also suggested that although financial incentives may affect physician behavior, other elements, such as health benefits to patients, have considerable influence on medical practice (Hennig-Schmidt et al, 2011; Godager and Wiesen, 2013). In addition, extensive research has been conducted on the effects of payment systems on clinical practice, but the majority of these studies have focused on outpatient care and drug prescriptions. At present, our knowledge on the effects of payment systems on inpatient care remains limited (Echevin and Fortin, 2014).[2] Moreover, to the best of our knowledge, there are no studies in the current literature that examine the importance of financial incentives in emergency medical services (EMS).

With the aim of shedding light on this relationship, this paper presents an example of how EMS providers respond to payment systems by exploiting a discontinuous financial incentive that is directly dependent on the measured amount of hospital care provided. Specifically, we focus on the possibility that the traditional method of counting bed days may inadvertently encourage EMS providers to engage in undesirable behavior. In Japan, as in many countries, reimbursement for hospital services is tied to the number of patient bed days. An emergency patient who is admitted to a hospital at 23:55 and is discharged the following morning would have to pay out-of-pocket expenses for two nights rather than one, even if they are only hospitalized for one night.[3] This is because the calculation of bed days is based on a "midnight-to-midnight" method, where a "day" is defined as the period from one midnight to the next.

From a theoretical standpoint, this midnight-to-midnight definition used in Japan may encourage EMS providers to ensure that emergency patients arrive before midnight, as the starting "day" for hospital care directly affects reimbursements from insurers. A hospitalization episode begins when an emergency patient

[1] Given the possibility of publication bias on this issue, there may be an overrepresentation of studies that demonstrate a link between payment systems and physician behavior.

[2] Echevin and Fortin (2014) investigated the effects of alternative payment systems on length of stay and the probability of readmission by exploiting the introduction of a new reimbursement scheme in Quebec.

[3] This description is based on the Japanese system. The U.S. Medicare and Medicaid system excludes the day of discharge and death from the calculation of bed days. In addition, emergency patients in the U.S. are initially treated as outpatients, even if they are eventually hospitalized after clinical examination and observation. A new rule for reimbursements in the U.S. also uses the number of midnights to assign inpatient status. In 2014, the Centers for Medicare and Medicaid Services introduced a "2-midnight rule". Under this rule, Medicare Part A covers hospitalizations of 2 midnights ("2 days") or longer; in contrast, stays shorter than 2 midnights are regarded as outpatient episodes for observation, and are therefore not covered by Medicare Part A (Sheehy et al, 2013).

is received by hospital staff, and the day of arrival is therefore important to health care providers. While hospital staff are unable to manipulate the starting dates of hospitalizations for planned admissions (which generally begin in the daytime), this is not the case for emergency patients admitted at night. Anecdotal evidence also suggests that hospital staff may have the ability to shift the timing of emergency arrivals by a few minutes.[4] Hence, if hospital staff behave according to the conventional supplier-induced demand (SID) theory (Dranove, 1988), they may attempt to increase reimbursement by manipulating the arrival times of emergency patients such that they are processed before midnight. In addition, we also posit that this incentive may be larger for patients with mild and moderate symptoms than for more severe patients, because the financial margin of this manipulation would be larger with shorter hospitalization durations under the Japanese health care system.

In this study, we test the hypothesis that EMS staff manipulate patient arrival times at night through an analysis of administrative records of emergency transportations covering almost all emergency admissions in Japan from 2008 to 2011. The data contain records of emergency phone call times, on-scene arrival times of ambulances at each patient ' s location, arrival times at the hospital, as well as basic patient characteristics (age, sex and reason for emergency care). The entire sample comprised approximately 16 million emergency cases. Among these, we focused on 2.1 million cases where the emergency call times were recorded as being within 180 minutes before or after midnight.

Our findings can be summarized in the following 3 main points. First, we observed significant bunching in the number of hospital arrivals around midnight: the number surged just before midnight but fell after midnight had passed. Among patients who were thought to arrive during the short period from 0:00 to 0:04, the estimated proportion of manipulated patients (i.e., reported as having arrived before midnight) was 2.6%, which was statistically significant at the 95% confidence level. Second, we found no signs of manipulation for patients who were dead on arrival and those with severe symptoms, which was consistent with our predictions. This was likely because the addition of an extra bed day is not feasible for dead patients, and the financial margin may be too small or even non-existent for patients with severe symptoms who are expected to be hospitalized for longer durations. For patients with mild symptoms, we observed a large degree of bunching before midnight: the estimated proportion of manipulated patients in the period just after midnight increased to 3.6%, indicating that the financial incentive triggers this manipulation. Third, the manipulation of arrival time was generally observed in prefectures where EMS are provided mainly by private hospitals. For example, 70% of emergency patients in Chiba prefecture were treated at a private hospital, and the proportion of manipulated patients was as high as 5.6%. This suggests that hospital ownership and regional structures of emergency care provision are associated with the manipulation of hospital arrival time. When considered together, these findings demonstrate the existence of socially wasteful behavior among EMS providers in Japan, and that the magnitude of this behavior is modest but not negligible.

[4] We also conducted interviews with several physicians, who confirmed awareness of the benefits to their hospital associated with receiving emergency patients before midnight.

The remainder of this paper is structured as follows: the next section explains the background of EMS providers in Japan. Section 3 describes the theoretical framework of our study, and Section 4 introduces the data used. The empirical strategy to detect manipulation of hospital arrival time is presented in Section 5. The main findings are then described in Section 6. Finally, our concluding remarks are presented in Section 7.

2 Background

To provide insight into the contextual background of our study, this section briefly explains the EMS provision system and the associated payment systems in Japan.

2.1 Emergency Medical Services in Japan

The Japanese EMS system is composed of three tiers: (1) Primary EMS, which serves patients who do not require hospitalization, (2) Secondary EMS, which serves patients who require hospitalization and (3) Tertiary EMS, which provides intensive care for severely ill patients. The numbers of centers that provided primary, secondary and tertiary EMS in 2010 were 529, 3,231 and 221, respectively.[5] Due to Japan's universal coverage in health care, these services are designed to be affordable for all emergency patients. In addition, patients are not charged for emergency transportation to EMS providers.

Fire stations, which are managed by local municipal governments, provide emergency transportation services in Japan. There were 791 fire stations located throughout the country in 2012, and 61% of these served small populations with fewer than 100,000 people. As fire station staff are municipal employees, their salaries are based on the seniority system, and are not linked to the quality and volume of transportation. It is important to our study that emergency transportation services and care are provided by different institutions. As the remuneration of fire station staff is independent from the hospital sector, they would have no reason to manipulate hospital arrival times to increase the revenue of hospitals.

In addition, we postulate that hospital ownership may affect behavior in the manipulation of arrival time to increase revenue. Private hospitals are the main provider of EMS in Japan, supplying almost 60% of all emergency care. In 2012, 53.9% of all emergency patients were transferred to private clinics or hospitals, while 23.6% were transferred to hospitals owned by prefectural or municipal governments. Approximately 13% were transferred to not-for-profit hospitals such as those owned by the Japanese Red Cross Society. The remaining 6.4% of emergency patients were transferred to national hospitals. It should be noted that tertiary EMS is likely to be provided in government-owned hospitals such as emergency critical care centers (known in Japanese as *kyumei kyukyu sentaa*)[6], rather than in private and not-for-profit hospitals. Hence, patients in critical condition are likely to be transferred to a government-owned hospital.

[5] A general summary of the Japanese EMS system is also provided in Kitamura et al (2010).

[6] There are currently 266 emergency critical care centers located throughout Japan. Not all are government-owned, but the prefectural governments are generally responsible for the management of these centers.

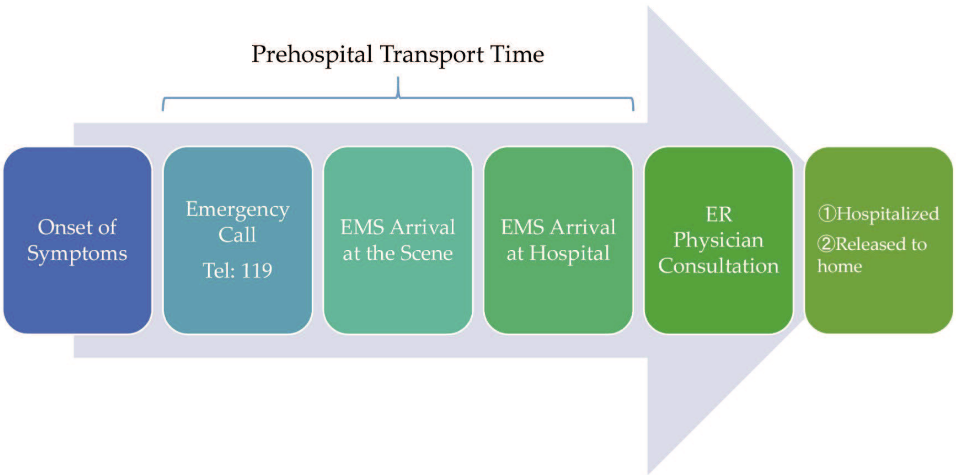


Figure 1: Timeline of Events in Emergency Transportation

The detailed timeline of emergency transportation is presented in Figure 1. Patients with emergency symptoms in Japan can call for an ambulance by dialing 119. The timing of the call is recorded by the receiving fire station through an automated recording system. After receiving a call, the fire station dispatches an ambulance to the scene, and the time of on-scene arrival is recorded by the ambulance crew. The ambulance then transports the patient to a suitable hospital. The arrival time of the patient at the hospital is designated "hospital arrival time", and the duration from the 119 call to hospital arrival is designated "prehospital transport time".[7] Upon arrival at a hospital, the patient is examined by an emergency care specialist. If the specialist determines that there is no need for hospitalization, the patient is charged as an outpatient, and is required to pay a copayment based on an outpatient care fee schedule. If the patient is hospitalized, the consultation is regarded as part of inpatient care and is billed accordingly. In this case, the starting time of hospitalization is defined as the moment that the patient passes through hospital reception, which is essentially equivalent to hospital arrival time.

2.2 Payment Systems

In this section, we explain the financial incentives in the provision of hospital care that are related to our research theme.

Until 2003, the FFS system was the principal reimbursement system for hospital care in Japan. Under this system, the amount of reimbursement to hospitals corresponds to the duration of hospitalization as well as several fundamental measures that reflect the quality of care. For example, reimbursement for hospital care is linked not only to the number of patient bed days, but also to the number of nursing staff per

[7] Unfortunately, our data did not include the time of symptom onset. Many previous studies such as Smolderen et al (2010) and Mooney et al (2014) define "prehospital transport time" as the period from the onset of symptoms to arrival at the hospital, whereas others such as Kleindorfer et al (2006) and Kitamura et al (2014) investigated the determinants of the time from the emergency call to arrival at the hospital. On this point, see Appendix B.

patient. Under the FFS system, longer hospitalizations result in larger reimbursements, which incentivizes the prolongation of hospitalization durations. However, the amount of reimbursement per bed day decreases for hospitalizations that extend beyond 14 days, and decreases further for hospitalizations that extend beyond 30 days. The base reimbursement is 40% higher in hospitalizations that last below 14 days when compared with hospital stays of 30 days or more. Because of this financial penalty for protracted hospitalizations, the incentive to prolong hospital stays may be diminished in severely ill patients.

In 2003, the prospective payment system (PPS) was introduced in Japan for acute care services. The Japanese PPS has two main characteristics: First, it is based on the patient diagnosis procedure combination (DPC) case-mix system, which is analogous to the DRG system in the U.S. Second, it is calculated "per diem", rather than "per episode".[8] The DPC/Per-Diem Payment System (DPC/PDPS) was initiated in 82 hospitals[9] in 2003, and since then, many other hospitals have adopted the system. The number of DPC/PDPS-compliant hospitals was 713 in 2008, which had doubled by 2012. These hospitals encompassed 55% of all regular hospital beds in Japan in 2012 (Ministry of Health and Welfare, 2014). The shift from FFS to DPC/PDPS has also affected EMS. However, the DPC/PDPS is also susceptible to the same financial incentive as the FFS system with regard to length of stay, as reimbursement increases with longer hospitalizations.

3 Theoretical Framework

This section provides a brief theoretical discussion on our hypothesis. Essentially, our model is based on the SID theory, as the manipulation of hospital arrival time can be broadly regarded as an example of SID. The basic theoretical model is derived from the utility function proposed by Ellis and McGuire (1986), in which a physician acts as a double agent to both his patients and the hospital, and chooses treatments that patients willingly accept.

In the application of this model to the Japanese context, it should be noted that hospital physicians work for a salary, which is not tied to their practice load (Ikegami and Campbell, 1995; Shigeoka and Fushimi, 2014). This distinguishes our model from the standard assumption of a conventional SID model in which physicians are concerned with their own income, but not institutional revenue. Based on He and Mellor (2012) and Shigeoka and Fushimi (2014), we argue that in this context, physicians consider the revenue of the hospital rather than their own income and the private costs/benefits of demand inducement, as described in the following utility function:

$$U = U(\pi(I), B(I)), \tag{1}$$

[8] The reason for the per-diem reimbursements is the wide variations in average length of stay durations among Japanese hospitals. If the reimbursement rates for hospitalizations were based on average hospital costs per episode, hospitals with longer hospitalizations would lose a considerable amount of revenue. As this may incentivize the premature discharge of patients, the per-diem payment system was adopted.

[9] These include university hospitals and some national hospitals.

where π is the net revenue of the hospital, and B is the patient's expected net benefit from demand inducement (I), which is evaluated by the physicians. We can also assume that $U_\pi > 0$, $U_B > 0$, $U_{\pi\pi} < 0$ and $U_{BB} < 0$.[10] The key assumption in this equation is that the inducement invariably reduces the net benefits of the patient, namely $B_I < 0$. In our context, the manipulation of emergency arrival times would result in additional out-of-pocket expenditures for patients without any improvement in health, indicating that $B_I < 0$. The equilibrium level of inducement is given when we solve the following first order condition:

$$\frac{\partial U}{\partial \pi}\theta + \frac{\partial U}{\partial B}\frac{\partial B}{\partial I} = 0, \tag{2}$$

where θ is a financial incentive factor of the inducement, which meets the condition $\pi = \theta I$.

Next, we discuss the relationships between demand inducement behavior, the financial incentive and the characteristics of the treatment provided. The last factor is particularly important because previous studies on SID have addressed various medical treatments such as coronary artery bypass graft surgeries (Yip, 1998), chemotherapy for lung cancer patients (Jacobson et al, 2010) and newborn treatments in neonatal intensive care units (Shigeoka and Fushimi, 2014). In our model, the characteristics of medical treatments are incorporated into the term B_I. If an inducement causes substantial harm to a patient, B_I takes a large negative value; otherwise, it takes a small value in the absolute term. Intuitively, the responsiveness of physicians to financial incentives varies according to the value of B_I. For example, physicians may treat their patients aggressively if B_I takes a slightly negative value. On the other hand, physicians may be less inclined to over-treat their patients if it causes harm, since they would generally be concerned about their patients' net benefits (B). Although a more detailed discussion on this point is provided in Online Appendix A, we briefly address two implications of the manipulation of hospital arrival time in Japan. First, the manipulation of hospital arrival time would unequivocally deteriorate each patient's net benefit, even without taking into account the possible effects on patient health. In particular, an emergency patient with mild symptoms who arrives at a hospital a few minutes before midnight would not receive any additional health benefits, but their copayment would increase substantially. We also note that EMS providers are likely aware of this situation when they rush to accept emergency patients before midnight. In these cases, B_I may take a large value, and EMS providers may refrain from manipulating the arrival times. On the other hand, it is also possible that we may observe a large degree of manipulation (as seen in previous studies) if EMS providers act selfishly as predicted by the standard economic theory of cheating. Second, our model indicates that copayment designs affect the degree of SID. For example, some forms of SID, such as upcoding, would not directly affect patient benefits in countries where patients pay fixed amounts regardless of total medical costs. However, SID induces higher patient costs in Japan as patients pay a fixed percentage of total health care costs (i.e. the amount of copayment is proportional to the health care costs). As a result, the equilibrium level of SID may be higher in countries with a fixed copayment system than in

[10] For the sake of simplicity, the private costs/benefits of the inducement are assumed to be negligible.

Japan.

4 Data

4.1 Administrative Records of Emergency Transportation

We utilized unique administrative data of emergency transportations in Japan, which are collected by the Fire and Disaster Management Agency (FDMA). Our data covered all emergency transportations throughout Japan (except for Tokyo prefecture) from January 2008 to December 2011.[11] Since the data are collected by the FDMA, they do not include follow-up data of patients after hospital arrival. However, these data include precise information on the timing of emergency events, such as the 119 emergency call and hospital arrival. Using these time records, we counted the number of emergency calls and hospital arrivals in 5 minute intervals.

The total number of records was 16,856,758, which covers 86.3% of all emergency transportations during the study period. Among these, we focused on 2,146,498 records in which the emergency call times were recorded as being within 180 minutes before or after midnight. The data also contain information on patient severity, which was measured by a physician at first contact; the first physician who examines an emergency patient evaluates their severity and informs the ambulance crew, who record this information. The classification of patient severity is highly standardized. In general, each emergency patient is grouped into 1 of the following 4 categories: dead, severe, moderate and mild. Dead patients are those who had died before the ambulance arrived at the hospital. Severe patients are those expected to be hospitalized for over 3 weeks, and moderate patients are those expected to be hospitalized for 3 weeks or less. If patients do not require hospitalization, they are categorized as mild cases. In addition to patient severity, the data also include causes of the emergency, which are divided into the following 11 categories: fire accident, natural disaster, water accident, traffic accident, workplace accident, sports and recreational activity accident, general accident, violence, attempted suicide, disease and hospital transfer. Information on patient characteristics includes age, sex and prefecture of residence. Finally, the place where the emergency occurred is also recorded using the following 5 categories: patient's home, public facility, workplace, road and other places.

As our study is dependent on the accuracy of the various reported times (call time, on-scene time and hospital arrival time), this issue should be addressed in greater detail. Call times are likely to have an extremely high level of accuracy without significant measurement errors, as the FDMA automatically records each call correct to the day, hour and minute. On the other hand, on-scene time and hospital arrival time may be more susceptible to misreporting because they are recorded by the ambulance crew. For example, ambulance crews have to push a button in the ambulance that informs their commander of their arrival at the hospital. It is a potentially serious threat to our assumptions and analytical approach if the ambulance crew systematically misreports hospital arrival time around midnight. The possibility this reporting bias

[11] Fire stations in Tokyo are managed by an organization independent of the national government, and as such are not included in the FDMA database.

has already been addressed by Almond and Doyle (2011) in the reporting of births; they observed that obstetricians avoid reporting 0:00 as the exact birth time because of a desire to clarify the exact day of birth. Although there are no incentives for fire station workers to behave in a similar way (because these small changes to the arrival times would have no impact on their salary or occupational reputation), we cannot fully disregard the possibility of reporting bias. If this is the case, we may confuse mere reporting bias in hospital arrival time with "manipulation" of hospital arrival time. To rule out this possibility, several checks were implemented as described below, and we discuss our measures to prevent reporting bias from affecting our findings.

5 Empirical Strategy

Our method to estimate the extent of manipulation of hospital arrival time is heavily reliant on recent advances in bunching estimation techniques in the context of progressive income taxation, although we adopted a modified method more suited to our data and research purpose. Among the empirical studies on bunching, seminal work by Chetty et al (2011) has provided a method to quantify the size of bunching. Einav et al (2013) builds further on Chetty et al (2011)'s strategy to investigate the bunching of drug expenditure around the kinks generated from the "donut-hole" in Medicare Part D. The strategies from these previous studies have relied on the derivation of the counterfactual density of outcomes, i.e., what the distribution would look like without any kinks and notches. In order to derive this density, Einav et al (2013) fit a polynomial function to the actual count data, excluding the data near the threshold points, and extrapolate the counterfactual density from the polynomial function. The size of bunching, termed "excess mass", is calculated as the difference between the counterfactual and actual counts.

In the possible manipulation of hospital arrival time in Japan, we define the hospital arrival time of patient i to be *manipulated* when T_{i0} (which is the hospital arrival time that would be obtained if there were no manipulation) is different from the actual hospital arrival time (T_i). In the standard approach, this research question is reduced to examining whether the total number of arrivals at a hospital in a certain time bracket m (N_m) is equal to that which would be obtained in the non-manipulated counterfactual cases (N_{m0}). If N_m is significantly different from N_{m0}, we can show that hospital arrival time is manipulated. Of particular importance is that the standard strategies directly compare N_{m0} with N_m by adopting a polynomial function to fit underlying trends around kink and notch points.[12] However, it may be difficult to ascertain a plausible underlying density, especially when sample sizes are small; this suggests that the standard strategies may not be appropriate in some cases.

We began our analysis to estimate the counterfactual arrival times (T_{i0}) for all emergency cases. To derive T_{i0}, we utilized the fact that hospital arrival time is the sum of emergency call time and prehospital transport time (e.g. hospital arrival time was calculated to be 12:30 if emergency call time was at 12:05 and

[12] The choice of the polynomial depends on model fit. For example, a seventh order polynomial was adopted in Chetty et al (2011) and a cubic polynomial in Einav et al (2013). Similarly, Jürges and Köberlein (2015) adopted a fifth order polynomial to estimate the size of DRG upcoding in neonatal care.

prehospital transport time was 25 minutes). This simple relationship tells us that the counterfactual value of hospital arrival time could be obtained once the determinants of prehospital transport time without any manipulation are explicitly specified. Hence, for the first step, we explored the determinants of prehospital transport time during nighttime based on the following equation:

$$Tran_i = Z_i'\gamma + Pref + Month + Minute + \epsilon_i, \tag{3}$$

where $Tran_i$ is the prehospital transport time of patient i, Z_i is a vector of covariates, $Pref$ is the prefecture fixed effects, $Month$ is the fixed effects from 48 (4 years * 12 months) monthly dummy variables, $Minute$ is the emergency call time fixed effects and ϵ_i is an error term. In this equation, we eliminated upward trends of prehospital transport time during nighttime due to decreased staffing in emergency transport systems by controlling for emergency call time fixed effects.

An important assumption is that the manipulation of hospital arrival time around midnight by a few minutes would not have a large effect on the estimated coefficients in Equation (4), which are presented in Online Appendix B. This is because the number of manipulated patients may be negligible in comparison to the total number of patients admitted during nighttime. We determined this assumption to be reasonable because our calculations using Equation (4) included all emergency patients who called for an ambulance from 21:00 to 3:00. The choice of this time interval was arbitrary, but the results of interest were robust to different ranges.[13]

After estimating Equation (4), we derived the counterfactual hospital arrival times for patient i by adding the predicted prehospital transport time to the emergency call time, and allowing for accidental deviation from the mean. In order to replicate accidental distribution of prehospital transport time, we used the root mean square error (MSE) in Equation (4), as it provides a consistent estimate of the distribution of ϵ_i. After 1,000 repetitions, we counted the number of hospital arrivals in 5-minute time intervals and defined the 95% confidence intervals (CIs). The CIs were defined as two points located in the 2.5th and 97.5th percentiles. Finally, hospital arrival times were designated as being manipulated if the actual count fell below the lower bound of the CI in the 5-minute interval just after midnight. While the results from this strategy were similar to those from standard methodology, it should be noted that we did not depend on explicit assumptions regarding the choice of the counterfactual density of the outcome variable.

This entire procedure can be summarized as follows:

Step 1 Estimate Equation (4) using the sample of all emergency patients who called for an ambulance from 21:00 to 3:00.

Step 2 Calculate counterfactual prehospital transport time, assuming that the distribution of ϵ_i is asymptotically equal to the root MSE in Equation (4).

[13] We found that our main results were largely unaffected for shorter (from 22:00 to 2:00) and longer time intervals (from 20:00 to 4:00). These results are available on request.

Step 3 Calculate counterfactual hospital arrival time by summing the call time of each patient and their predicted transport time.

Step 4 Repeat **Step 2** and **Step 3** 1,000 times.

Step 5 Count the number of hospital arrivals in 5-minute intervals, and determine the 95% CIs.

Step 6 Compare the actual number of hospital arrivals with the lower bound of the CI in the 5-minute interval just after midnight.

6 Main Results

6.1 Preliminary Analysis

We present a graphical representation of the raw data to facilitate intuitive understanding of the main results. First, the number of emergency calls in 5-minute intervals around midnight is shown in Figure 2-(a). The horizontal axis of this figure represents the 2 hours before and after midnight in 5-minute intervals, and the circles represent the total number of calls in each interval. This figure clearly demonstrates that the occurrence of emergency episodes was smooth during the night, which allows us to reasonably assume that the number of hospital arrivals would also be smooth if there were no manipulation. In addition, the findings were similar for the number of on-scene arrivals, as shown in Figure 2-(b): this number was also smooth during the night, and the number of on-scene arrivals, which is recorded by the ambulance crews, showed no bunching around midnight. This indicates that the ambulance crews did not exhibit any reporting bias around midnight.

Despite the smoothness of the numbers of emergency calls and on-scene arrivals, our results, as shown in Figure 3-(a), clearly demonstrated a distinct bunching of hospital arrivals around midnight. We observed irregular surges in hospital arrivals in the period from 23:55 to 23:59, followed by a noticeable drop in numbers from 0:00 to 0:04. This is consistent with our prediction that hospitals may be manipulating the timing of emergency arrivals. A minute-by-minute plot provides more in-depth insight into the manipulation of hospital arrival numbers (Figure 3-(b)). The number was observed to consistently spike every 5 minutes because the ambulance crew would generally report arrival times that are rounded to the closest 5-minute mark. For example, they may report 23:55 as the official arrival time for patients who actually arrived at 23:56 or 23:54. Despite these regular spikes in arrival times, the number of arrivals at 0:00 was distinctly low. Instead, we observed many arrivals at 23:58 and 23:59. This suggests that the recorded arrival times of emergency patients who had arrived just after midnight were frequently shifted forward.

6.2 Statistical Analysis of Arrival Time Manipulation

In order to derive the CIs, we first explored the determinants of prehopital transport time, as explained in the previous section. As the determinants of prehospital transport time are also informative, we present the detailed results in Online Appendix B. Using the coefficients obtained from the regression analysis, we calculated the predicted values of prehospital transport times that included a random error component. Based on the distribution of the error term, we derived the alternative counterfactual value of hospital arrival time for each emergency episode. Hospital arrival time was then calculated as the sum of the emergency call time and the counterfactual prehospital transport time. After repeating this procedure 1,000 times, the 95% CIs were derived.

The results are presented in Figure 4 and Table 1. In addition to the number of hospital arrivals in each 5-minute interval, Figure 4 plots the 95% CIs and the counterfactual numbers. The data in Figure 4 are summarized in Table 1. Furthermore, Table 1 also presents the ratios of the actual number of hospital arrivals at each 5-minute interval to its corresponding counterfactual number, which was derived as the mean value of 1,000 simulations. From Figure 4, we can confirm the irregular patterns of arrival density in the 5-minute intervals from 23:55 to 23:59 and from 0:00 to 0:04; in these 5-minute intervals, the deviations from the counterfactual numbers amounted to 2.5% and -2.6%, respectively. If we evaluate the size of manipulation from this deviation in the time interval from 0:00 to 0:04, we can conclude that 2.6% of the patients who should have arrived in this interval had been shifted forward to the preceding time interval. The estimated number of shifted patients was 801[14] during the 4 years from January 2008 to January 2011.

6.3 Patient Severity and Arrival Time Manipulation

If the financial incentive to increase the number of bed days triggers the manipulation of arrival time, we would expect to find bunching in hospital arrivals only when this manipulation is truly beneficial to the hospital. The payment system in Japan allows us to conduct a falsification test to examine this hypothesis. We posit that hospitals under the FFS and DPC/PDPS system do not have a large incentive to manipulate arrival time for patients with serious symptoms because reimbursements under the Japanese system decrease for hospitalizations that extend over 14 days. Specifically, the daily reimbursement for hospital stay is reduced by 2,580 JPY when hospitalization exceeds 14 bed days. Furthermore, daily reimbursement is reduced further by 1,920 JPY when hospitalizations exceed 30 days. It should also be noted that the manipulation does not increase hospital reimbursements if the patient had died before arrival. However, this manipulation may be beneficial to the hospital if applied to patients with mild or moderate severity. Although patients with mild symptoms are defined as those who do not require hospitalization (based on the evaluation of the first physician), a substantial number of these cases are eventually hospitalized after examination by an emergency care specialist. For example, Hosoda et al (2005) conducted a follow-up study of emergency patients in Tottori prefecture, located in southern Japan, and reported that 34.8% of "mild severity" cases were actually hospitalized.[15]

We present these predictions graphically in Figure 5. The 4 graphs in the figure clearly demonstrate that

[14] 31,276-30,475=801

[15] Although there are no comprehensive studies on the probability of hospitalization among "mild severity" patients, the probability may be higher during the night because patients who do not require hospitalization may have difficulties finding transportation to return home. Based on interviews with several physicians, we found that elderly patients who do not require hospitalization are unlikely to go home at night, and hospital staff may sometimes admit them for a few days.

the extent of bunching varies across different levels of patient severity. If the patients are dead on arrival or in severe condition, the number of hospital arrivals in the 5-minute interval from 0:00 to 0:04 is within the CIs, indicating that there was no manipulation for these patients.[16] In contrast, we observed significant bunching among patients with mild and moderate symptoms.

The size of manipulation was found to be larger among patients with mild symptoms. As shown in Table 2, the estimated proportion of manipulated patients who are thought to have actually arrived from 0:00 to 0:04 was 3.6%, which was larger than in patients with moderate symptoms (2.3%). These results rule out the possibility that reporting bias accounted for the observed bunching, and indicate that the financial incentive to prolong hospitalization is a key driver of arrival time manipulation. This manipulation has no health benefits for patients with mild and moderate symptoms, despite the higher reimbursement for hospitals. Based on Labelle et al (1994)'s conceptual framework[17], the manipulation of hospital arrival time may be regarded as a clear example of socially-wasteful inducement of care.[18]

6.4 Geographical Variations in Arrival Time Manipulation

This section discusses the geographical locations where manipulation occurs. Our main hypothesis here is that private hospitals are more likely to engage in manipulation in order to increase revenue, whereas publicly-owned hospitals are less likely to do so because they receive more financial support from the government. Hence, we predict that manipulation would occur mainly in prefectures where the majority of emergency patients are treated at a private hospital. Previous studies have also underscored the importance of hospital ownership in financial motivations.[19] For example, Silverman and Skinner (2004) reported that upcoding in Medicare was prevalent in for-profit hospitals, but not in their not-for-profit counterparts. Lindrooth and Weisbrod (2007) examined financial incentives in hospices, and found that for-profit hospices were significantly less likely to admit patients with shorter predicted lengths of stay due to their lower profitability. In addition, Bayindir (2012) reported a significant difference in hospital treatment choice for unprofitable patients, such as the uninsured and Medicaid patients, suggesting that not-for-profit hospitals care more about less profitable patients than for-profit hospitals.

This section explores the relationship between hospital ownership and arrival time manipulation. To discern whether the arrival times of emergency patients are manipulated at the prefectural level, we first implemented the same procedure described in Section 5 for each prefecture, and then created a binary variable that takes a value of 1 if significant manipulation was observed in a prefecture.[20] This binary

[16]The width of the CIs appears to depend on the number of observations in each bin, but the point estimates also exhibited no irregular reductions in the number of hospital arrivals in the time interval from 0:00 to 0:04.

[17]Labelle et al (1994) reviewed the literature on SID and classified the several types of demand inducement based on clinical effectiveness and the effectiveness of agency between physicians and patients; they emphasized the importance of clarifying the impact of this inducement on patient health in order to facilitate normative evaluation.

[18]The manipulation of hospital arrival time may be categorized into "Cell IV" in Figure 1 of Labelle et al (1994).

[19]There is also a growing body of literature addressing the association between hospital ownership and quality of care. See the review published by Perotin et al (2013). Here, we contribute to the literature on the association between hospital ownership and the extent of financial motivations.

[20]As in the total sample results, the existence of manipulation was determined from an irregular deviation of the number of hospital arrivals in the 5-minute interval from 0:00 to 0:04. The results are graphically presented in Figure C1 in Online

variable was then regressed on various prefecture-level characteristics such as population, population density and per capita income. A variable of interest was the share of emergency patients who were transported to a private hospital, which was calculated using data from an annual survey on emergency transportation and rescue operations in 2010 (FDMA, 2010). If the coefficient of this variable was not significantly different from zero, we can conclude that private hospitals are more likely to engage in manipulation of hospital arrival time than their publicly-owned and not-for-profit counterparts.

We first identified the prefectures where manipulation was observed, according to prefecture-level hospital ownership. Figure 6 plots the proportion of emergency patients who were transported to private hospitals in each of the 46 prefectures. Although private hospitals are generally concentrated in urban areas, we observed considerable variations in the share of private hospitals at the prefectural level, ranging from 23% in Toyama prefecture to 78% in Osaka prefecture. However, the observation of significant manipulation (depicted in blue) was disproportionately higher in prefectures with a large share of private hospitals. The extent of manipulation was particularly high in the following areas: Gunma prefecture (4.5%), Saitama prefecture (3.6%), Chiba prefecture (5.6%) and Osaka prefecture (4.7%).

Finally, we address the results of the statistical analysis on the cross-sectional data of the 46 prefectures.[21] The regression results are summarized in Table 4 and the descriptive statistics are presented in Table 3. In Table 4, we implemented various specification checks with and without the covariates. First, Column (1) shows a significantly positive correlation between the dependent variable and the share of private hospitals. These results were largely unchanged when we controlled for other covariates, such as population, population density, elderly population ratio, the number of emergency hospitals per 1,000 population, and the natural log of per capita taxable income. Even if we incorporated all these covariates at once, the coefficient of the share of private hospitals was significant with the 90% CIs in Column (7). Taken together, these results are consistent with our hypothesis that private hospitals are more responsive to this financial incentive than their publicly-owned and not-for-profit counterparts.

6.5 Extent of Arrival Time Manipulation

In summary, manipulated patients accounted for 2.6% of the entire sample, but this share increased to 3.6% among patients with mild symptoms; we also found that manipulation was more prevalent in prefectures where EMS are provided mainly by private hospitals.[22] These results suggest that the extent of manipulation is modest, but not negligible in some subpopulations where the financial incentives have a stronger influence. When compared to the size of other manipulative behaviors reported in previous studies, the level of manipulation observed in our study sample was substantially lower. For example, studies on upcoding

Appendix C.

[21]The database includes Fukushima and Miyagi prefectures, which suffered immense damage during the 2011 Great East Japan Earthquake. As described in Online Appendix D, we examined the manipulation at the monthly level and found no trend breaks before and after the disaster.

[22]Since our study period included the period after the Great East Japan Earthquake in March 2011, we confirmed the robustness of the results by analyzing monthly data. The results are summarized in Online Appendix D, and showed no systematic changes throughout the study period.

have reported a larger magnitude of manipulative behavior than was observed in this study. Shigeoka and Fushimi (2014) noted that approximately 15% of newborns with birth weights that just exceeded 1,500 g were actually recorded as being below 1,500 g, as this would allow longer stays in neonatal intensive care units. Furthermore, Jürges and Köberlein (2015) reported that 92% of children with birth weights between 1,480 g and 1,519 g were recorded as being below 1,500 g in Germany.[23] Silverman and Skinner (2004) and Dafny (2005) also revealed that physicians have made substantial alterations to DRG codes to increase reimbursements in the U.S. In particular, Silverman and Skinner (2004) reported that the share of the most lucrative DRGs for pneumonia and respiratory infections rose by 23 percentage points among for-profit hospitals in the mid-90s.

There are several reasons that may explain our smaller degree of manipulation than those reported in previous studies, although it is difficult to evaluate their relative importance. First, in Japan, increasing reimbursements by manipulation of arrival times leads to a greater financial burden for patients, as patients must pay a fixed percentage of total health care costs.[24] This system may prevent manipulation if altruistic physicians care about patients' net benefits, as was discussed earlier. In contrast, manipulation would not harm patients directly if the copayment is fixed. This may partly explain the reasons why Jürges and Köberlein (2015) found prevalent upcoding in neonatal care in Germany, whereas our study observed less intensive manipulation. Second, the costs to ensure that patients arrive at hospitals before midnight may be higher than the costs of upcoding. While upcoding can be achieved with almost no costs for physicians, accepting patients earlier may be costly for EMS providers. Third, the information asymmetry between providers and patients may be less severe in this case, as patients can understand the reimbursement system without any specific knowledge on medical technologies. A conventional theoretical model of SID also points out that physicians induce less demand as patients ' diagnostic skills increase (Dranove, 1988).

7 Concluding Remarks

In many countries, reimbursement for hospital care is linked to the number of patient bed days, which are calculated by the number of midnights where a patient is hospitalized. This midnight-to-midnight method to count bed days may encourage hospitals to earn an extra bed day for reimbursements by ensuring that more patients arrive before midnight. Given that patients who are admitted just before midnight bring larger revenues for hospitals than patients who arrive after midnight, it is a rational response for health care providers to move patient arrival times forward. We tested this hypothesis using administrative records of emergency transportation in Japan, which contained 2.1 million records of patients whose emergency call times had been made within 180 minutes before and after midnight.

The results supported our prediction: we observed significant bunching in the number of hospital ar-

[23] Jürges and Köberlein (2015) also found that the upcoding was selective as physicians were more likely to upcode DRGs when treating newborns with poor health conditions and high expected health care costs.

[24] In general, the coinsurance rate in Japan is 30% for patients aged from school age to 69 years and 10% for those aged 70 years and older. In addition, the total amount of copayment per month exceeds monthly stop-loss limits.

rivals around midnight, where the number surged a few minutes before midnight but dropped just after midnight. Since the density of emergency calls was distributed smoothly throughout the night, the observed bunching in hospital arrivals suggests that EMS providers shifted arrival times forward in order to increase reimbursements. In addition, we observed significant bunching only in patients with mild and moderate symptoms. This is likely because the financial incentive to earn an extra bed day is stronger when applied to patients with shorter hospitalizations under the Japanese reimbursement system. Since a slightly earlier arrival at a hospital would not detrimentally affect the health status of patients with mild symptoms, the incremental reimbursement from arrival time manipulation can be regarded as socially wasteful. Finally, significant manipulation was more prevalent in prefectures where EMS are provided mainly by private hospitals, suggesting that hospital ownership is associated with engagement in revenue-raising behavior that has absolutely no benefits for patient health. With regard to the impact of arrival time manipulation on health care costs, the size of manipulation was found to be modest but not negligible.

In closing, the authors would like to assert the following two points. First, our paper documents a clear example of cheating among EMS providers, regardless of the highly specific context. Welfare implications from our empirical results are also clear, as they show that health care providers are sometimes completely dishonest agents for the patients. At the same time, we suggest that the extent of this wasteful behavior is relatively restrained. Since there is currently no consensus on the magnitude and economic importance of SID, it is important that this paper adds new quantitative evidence to the literature addressing SID. Second, we found significant manipulation regardless of the fact that the salary of health care providers is not tied to practice load in Japan (Ikegami and Campbell, 1995). Although we argue that EMS providers probably engage in manipulation to benefit their hospital of employment, further investigations into the organizational behavior of health care providers may be beneficial for future research.

References

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

Bayindir EE (2012) Hospital ownership type and treatment choices. Journal of Health Economics 31(2):359 – 370

Chetty R, Friedman JN, Olsen T, Pistaferri L (2011) Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. The Quarterly Journal of Economics 126(2):749–804

Dafny LS (2005) How do hospitals respond to price changes? American Economic Review 95(5):1525–1547

Delattre E, Dormont B (2003) Fixed fees and physician-induced demand: A panel data study on french physicians. Health Economics 12(9):741–754

Devlin RA, Sarma S (2008) Do physician remuneration schemes matter? the case of canadian family physicians. Journal of Health Economics 27(5):1168 – 1181 van Dijk CE, van den Berg B, Verheij RA, Spreeuwenberg P, Groenewegen PP, de Bakker DH (2013) Moral hazard and suppier-induced demand: Empirical evidence in general practice. Health Economics 22(3):340–352

Dranove D (1988) Demand inducement and the physicians/patient relationship. Economic Inquiry 26(2):281–298

Echevin D, Fortin B (2014) Physician payment mechanisms, hospital length of stay and risk of readmission: Evidence from a natural experiment. Journal of Health Economics 36(C):112–124

Einav L, Finkelstein A, Schrimpf P (2013) The Response of Drug Expenditures to Non-Linear Contract Design: Evidence from Medicare Part D. NBER Working Papers 19393, National Bureau of Economic Research, Inc, URL http://ideas.repec.org/p/nbr/nberwo/19393.html

Ellis RP, McGuire TG (1986) Provider behavior under prospective reimbursement: Cost sharing and supply. Journal of Health Economics 5(2):129 – 151

FDMA (2010) Annual surveys on emergency transportation and rescue operations in 2010

Fukuoka Y, Dracup K, Rankin SH, Froelicher ES, Kobayashi F, Hirayama H, Ohno M, Matsumoto D (2005) Prehospital delay and independent/interdependent construal of self among japanese patients with acute myocardial infarction. Social Science and Medicine 60(9):2025 – 2034

Godager G, Wiesen D (2013) Profit or patients'health benefit? exploring the heterogeneity in physician altruism. Journal of Health Economics 32(6):1105 – 1116

Grytten J, Sorensen R (2001) Type of contract and supplier-induced demand for primary physicians in norway. Journal of Health Economics 20(3):379–393

Hadley J, Holahan J, Scanlon W (1979) Can fee-for-service reimbursement coexist with demand creation? Inquiry 16(3):pp. 247–258

He D, Mellor JM (2012) Hospital volume responses to medicare's outpatient prospective payment system: Evidence from florida. Journal of Health Economics 31(5):730 – 743

Ministry of Health L, Welfare (2014) Materials for central social insurance medical council (in japanese; cyuui kyou shiryou), URL http://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000041708.pdf, in Japanese

Hennig-Schmidt H, Selten R, Wiesen D (2011) How payment systems affect physicians'provision behaviour–an experimental investigation. Journal of Health Economics 30(4):637 – 646

Hosoda T, Fujita T, Tanihata T (2005) Follow-up study on the emergency patients. Journal of Health and Welfare Statistics 52(6):27–34, in Japanese

Ikegami N, Campbell JC (1995) Medical care in japan. New England Journal of Medicine 333(19):1295–1299

Jacobson M, Earle CC, Price M, Newhouse JP (2010) How medicare's payment cuts for cancer chemotherapy drugs changed patterns of treatment. Health Affairs 27(9):1391–1399

Jürges H, Köberlein J (2015) What explains drg upcoding in neonatology? the roles of financial incentives and infant health. Journal of Health Economics

Kitamura T, Iwami T, Kawamura T, Nagao K, Tanaka H, Hiraide A (2010) Nationwide public-access defibrillation in japan. New England Journal of Medicine 362(11):994–1004

Kitamura T, Iwami T, Kawamura T, Nishiyama C, Sakai T, Tanigawa-Sugihara K, Sasaki M, Kajino K, Irisawa T, Hayashida S, Nishiuchi T, Hiraide A (2014) Ambulance calls and prehospital transportation time of emergency patients with cardiovascular events in osaka city. Acute Medicine and Surgery 1(3):135–144

Kleindorfer DO, Lindsell CJ, Broderick JP, Flaherty ML, Woo D, Ewing I, Schmit P, Moomaw C, Alwell K, Pancioli A, Jauch E, Khoury J, Miller R, Schneider A, Kissela BM (2006) Community socioeconomic status and prehospital times in acute stroke and transient ischemic attack: Do poorer patients have longer delays from 911 call to the emergency department? Stroke 37(6):1508–1513

Labelle R, Stoddart G, Rice T (1994) A re-examination of the meaning and importance of supplier-induced demand. Journal of Health Economics 13(3):347–368

Lindrooth RC, Weisbrod BA (2007) Do religious nonprofit and for-profit organizations respond differently to financial incentives? the hospice industry. Journal of Health Economics 26(2):342 – 357

Mooney M, McKee G, Fealy G, Brien FO, O'Donnell S, Moser D (2014) A randomized controlled trial to reduce prehospital delay time in patients with acute coronary syndrome (acs). The Journal of Emergency Medicine 46(4):495 – 506

Perotin V, Zamora B, Reeves R, Bartlett W, Allen P (2013) Does hospital ownership affect patient experience? an investigation into public-private sector differences in england. Journal of Health Economics 32(3):633 – 646

Shafrin J (2010) Operating on commission: analyzing how physician financial incentives affect surgery rates. Health Economics 19(5)

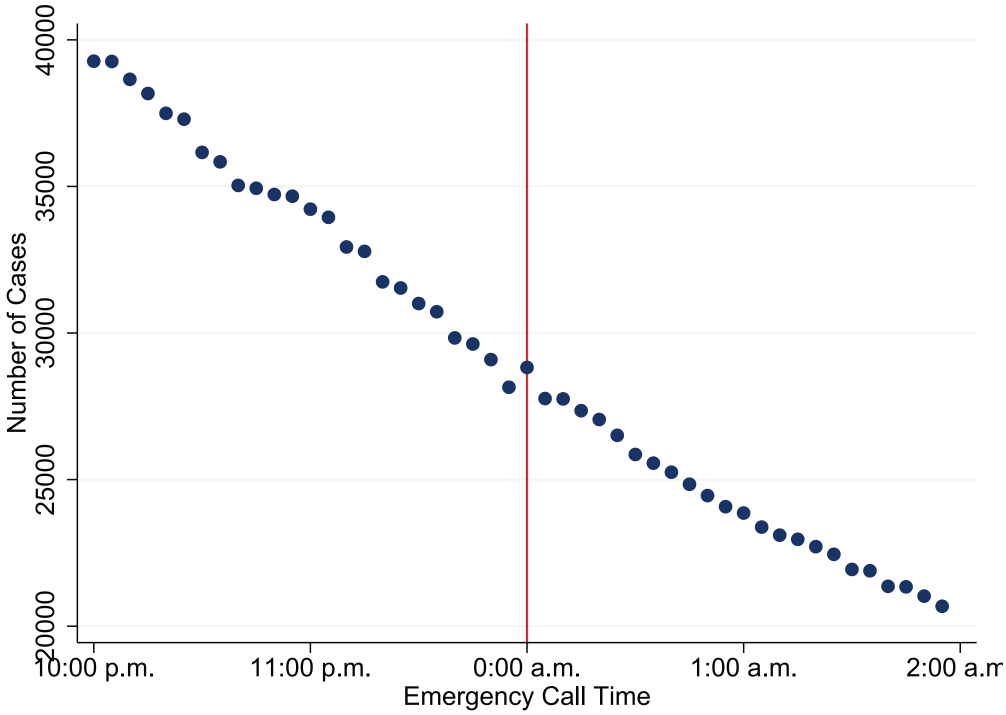
Sheehy A, Graf B, Gangireddy S, Formisano R, Jacobs E (2013) "observation status" for hospitalized patients: Implications of a proposed medicare rules change. JAMA Internal Medicine 173(21):2004–2006

Shigeoka H, Fushimi K (2014) Supplier-induced demand for newborn treatment: Evidence from japan. Journal of Health Economics 35(0):162 – 178

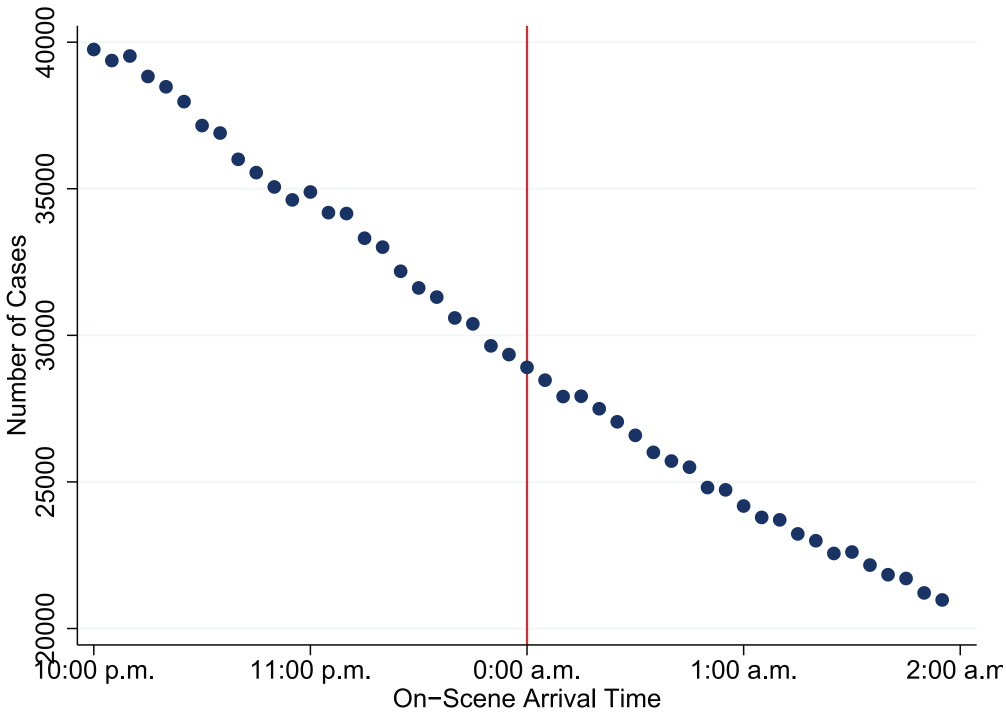
Silverman E, Skinner J (2004) Medicare upcoding and hospital ownership. Journal of Health Economics 23(2):369–389

Smolderen K, Spertus J, Nallamothu B, et al (2010) Health care insurance, financial concerns in accessing care, and delays to hospital presentation in acute myocardial infarction. JAMA 303(14):1392–1400

Yip WC (1998) Physician response to medicare fee reductions: changes in the volume of coronary artery bypass graft (cabg) surgeries in the medicare and private sectors. Journal of Health Economics 17(6):675 – 699



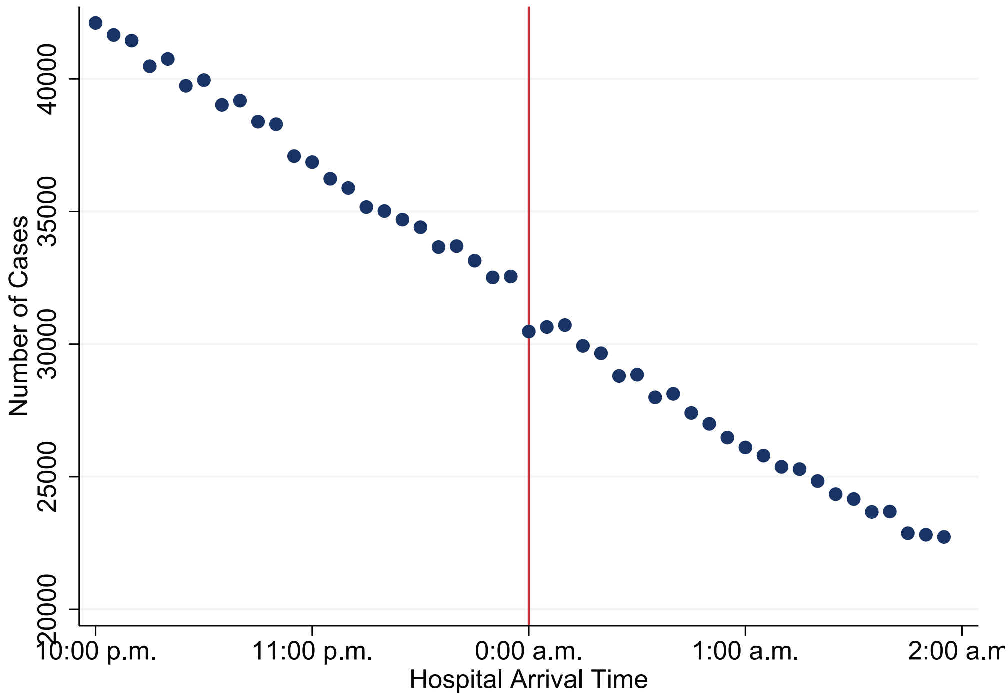
(a) Number of Emergency Calls



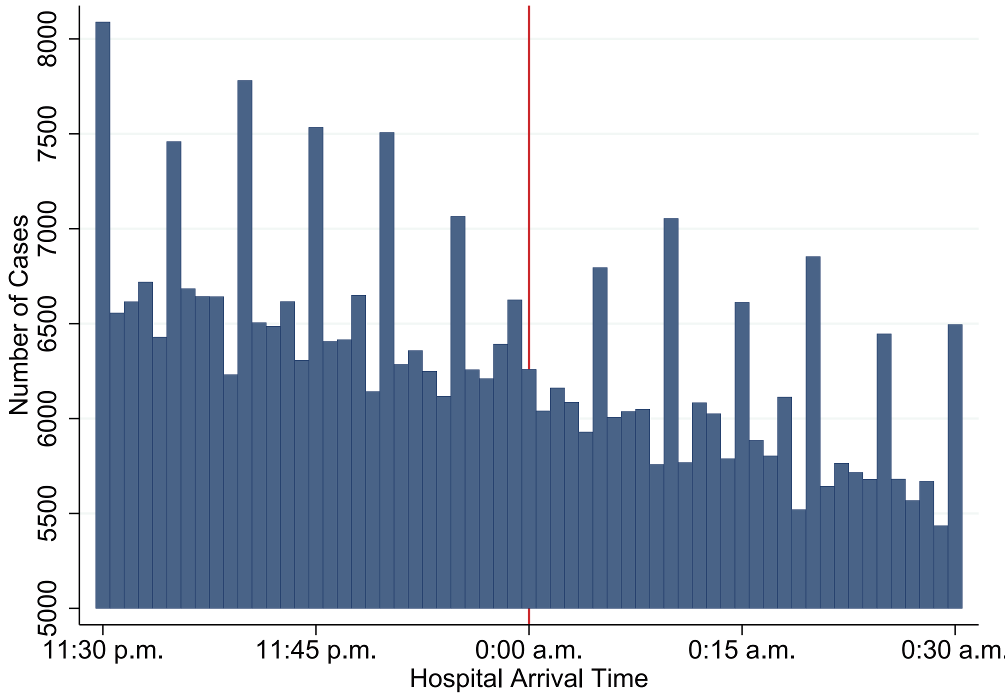
(b) Number of On-Scene Arrivals

Note: This figure plots the number of emergency calls and on-scene arrivals in 5-minute time intervals in the 120 minutes before and after midnight. Data include all emergency calls from 2008 to 2011 throughout Japan except for Tokyo prefecture.

Figure 2: Number of Calls and On-Scene Arrivals



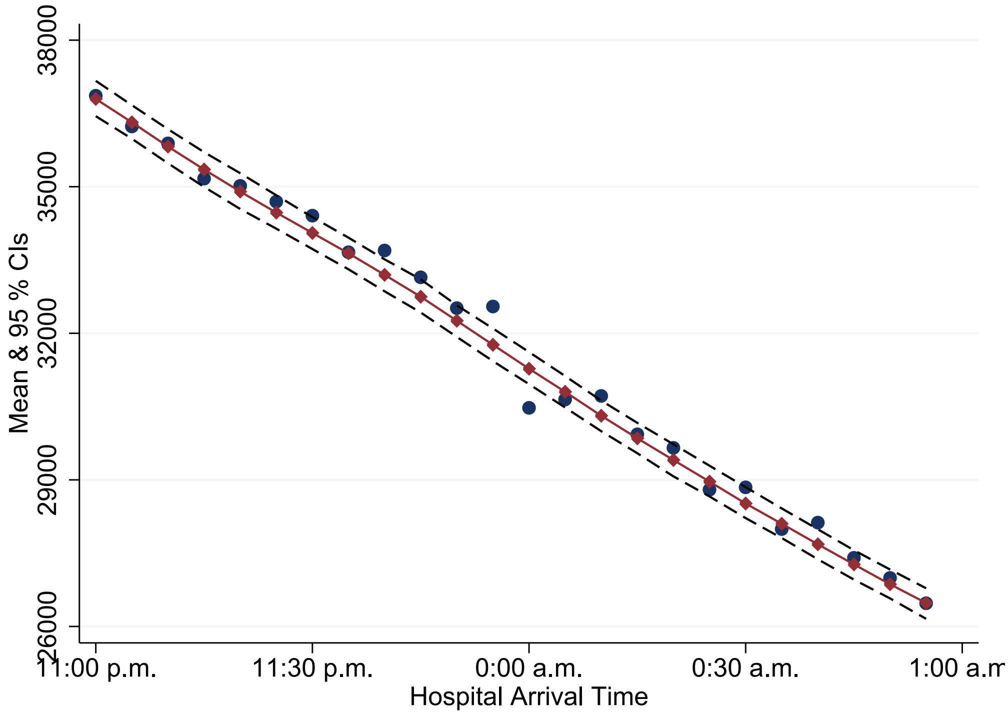
(a) In 5-minute intervals



(b) In 1-minute intervals

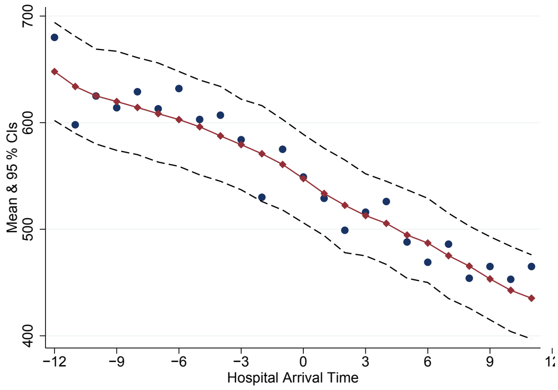
Note: Figure (a) plots the number of hospital arrivals in 5-minute time intervals in the 120 minutes before and after midnight. Figure (b) shows the minute-by-minute plot during the 30 minutes before and after midnight. Data include all emergency hospital arrivals from 2008 to 2011 throughout Japan except for Tokyo prefecture.

Figure 3: Number of Hospital Arrivals

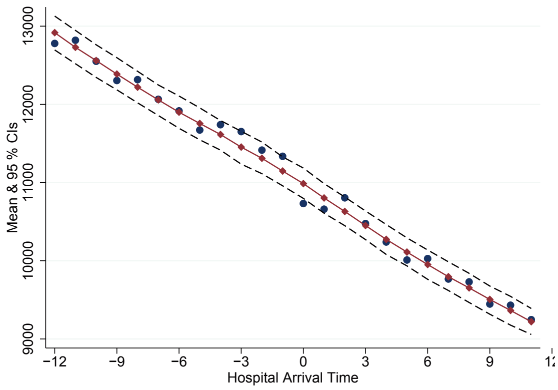


Note: Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals that would have been obtained if the arrival times were not manipulated. Dashed lines denote 95% CIs.

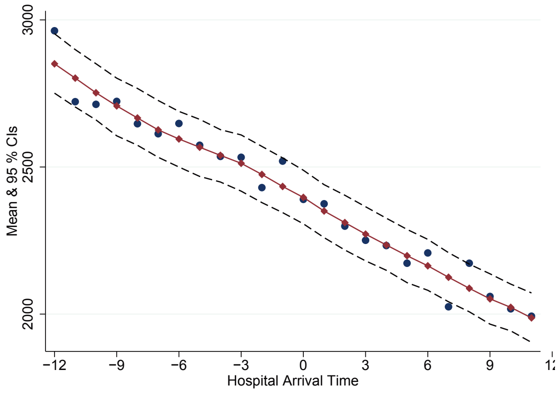
Figure 4: Number of Hospital Arrivals and Confidence Intervals



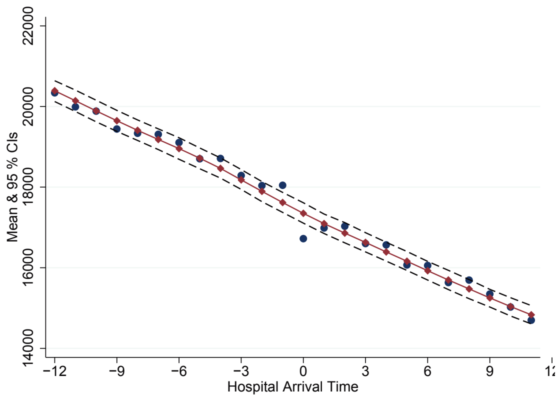
(a) Dead



(c) Moderate



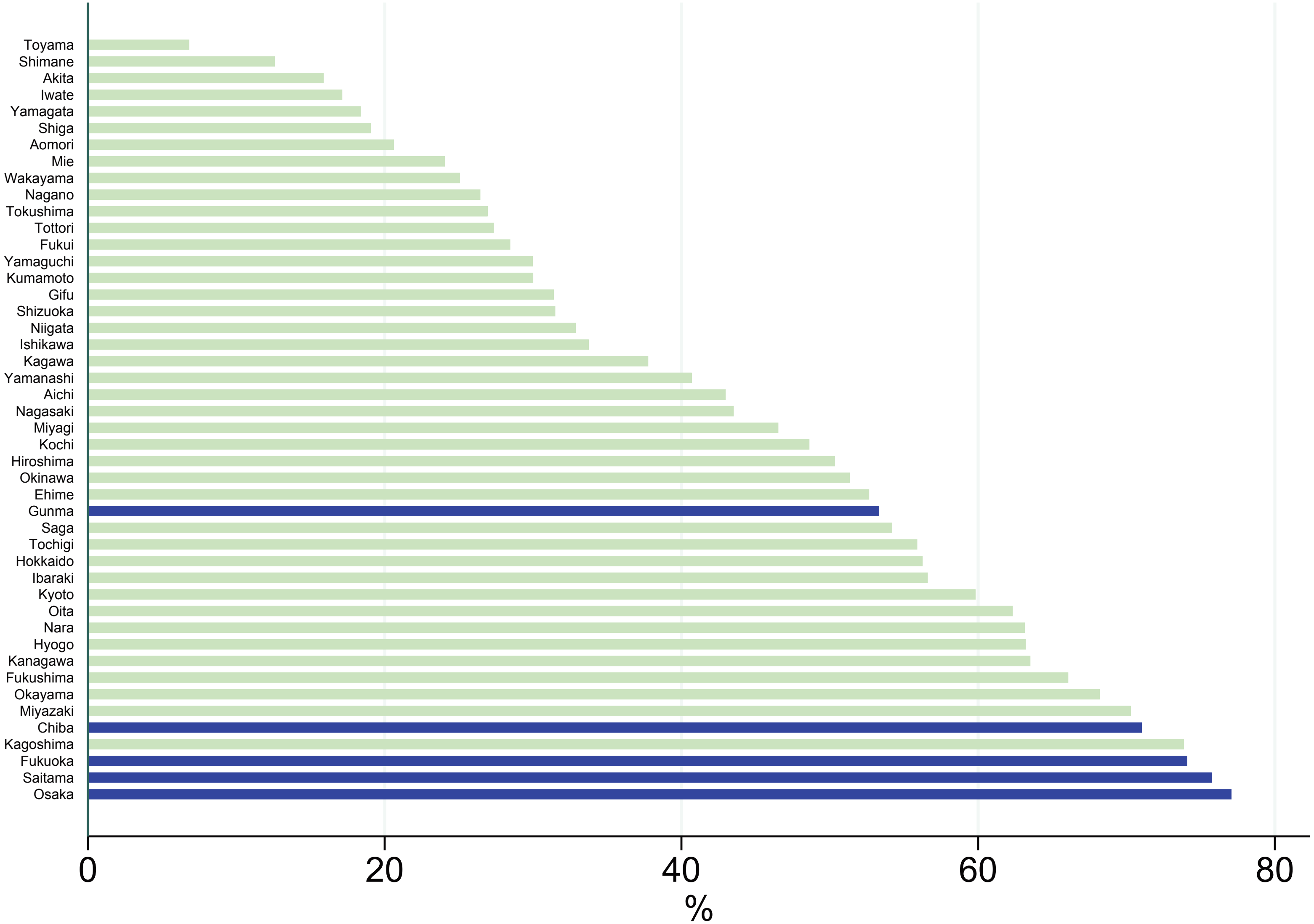
(b) Severe



(d) Mild

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals that would have been obtained if the arrival times were not manipulated. Dashed lines denote 95% CIs.

Figure 5: Number of Hospital Arrivals by Patient Severity



Note: The bars represent the proportions of emergency patients transported to private hospitals for each prefecture in 2010. Tokyo prefecture is excluded. The prefectures with significant manipulation of arrival times are depicted in blue.

Figure 6: Proportion of Patients Transported to Private Hospitals by Prefecture

Table 1: Test of Manipulation

	Actual (1)	Mean (2)	95% CI (3)	Ratio (4)
11:00 p.m.	36,862	36,797	[36444-37167]	0.2%
11:05 p.m.	36,234	36,318	[35982-36669]	-0.2%
11:10 p.m.	35,887	35,821	[35480-36178]	0.2%
11:15 p.m.	35,167	35,353	[34996-35707]	-0.5%
11:20 p.m.	35,018	34,901	[34544-35277]	0.3%
11:25 p.m.	34,695	34,471	[34142-34825]	0.7%
11:30 p.m.	34,408	34,054	[33725-34378]	1.0%
11:35 p.m.	33,659	33,637	[33310-33958]	0.1%
11:40 p.m.	33,695	33,197	[32860-33514]	1.5%
11:45 p.m.	33,146	32,748	[32426-33105]	1.2%
11:50 p.m.	32,516	32,258	[31927-32573]	0.8%
11:55 p.m.	32,549	31,765	[31431-32100]	2.5%
0:00 a.m.	30,475	31,276	[30959-31617]	-2.6%
0:05 a.m.	30,646	30,803	[30482-31125]	-0.5%
0:10 a.m.	30,718	30,313	[29997-30623]	1.3%
0:15 a.m.	29,933	29,852	[29546-30167]	0.3%
0:20 a.m.	29,656	29,407	[29073-29730]	0.8%
0:25 a.m.	28,799	28,965	[28662-29291]	-0.6%
0:30 a.m.	28,848	28,517	[28219-28852]	1.2%
0:35 a.m.	27,994	28,102	[27811-28416]	-0.4%
0:40 a.m.	28,126	27,683	[27376-27990]	1.6%
0:45 a.m.	27,405	27,271	[26960-27561]	0.5%
0:50 a.m.	26,994	26,864	[26575-27166]	0.5%
0:55 a.m.	26,475	26,480	[26157-26782]	0.0%

Note: The “Actual” column represents the actual number of hospital arrivals in each 5-minute time interval. In order to estimate the counterfactual numbers of hospital arrivals, we calculated 1,000 potential hospital arrival times for each emergency call. The potential counts of hospital arrivals were calculated for each time interval. The “Mean” column represents the mean of these potential counts. The 95% CIs were derived from the 2.5th and 97.5th percentile points of the 1,000 simulations. The ratios of excessive bunching are presented in the rightmost column, and were calculated by comparing the “Actual” and “Mean” values. The shaded row indicates the time interval from 0:00 to 0:04.

Table 2: Test of Manipulation Based on Patient Severity

	Dead		Severe		Moderate		Mild	
	Actual	Ratio	Actual	Ratio	Actual	Ratio	Actual	Ratio
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
11:00 p.m.	680	4.9%	2,963	3.9%	12,779	-1.1%	20,339	-0.3%
	[602-694]		[2751-2953]		[12696-13129]		[20123-20639]	
11:05 p.m.	598	-5.7%	2,722	-2.9%	12,820	0.7%	19,991	-0.8%
	[590-681]		[2704-2899]		[12520-12946]		[19876-20406]	
11:10 p.m.	625	0.0%	2,713	-1.4%	12,551	-0.1%	19,885	0.0%
	[580-669]		[2660-2851]		[12345-12762]		[19620-20156]	
11:15 p.m.	614	-0.9%	2,723	0.6%	12,305	-0.7%	19,442	-1.0%
	[574-667]		[2606-2802]		[12186-12597]		[19383-19904]	
11:20 p.m.	629	2.4%	2,647	-0.7%	12,316	0.8%	19,336	-0.4%
	[570-661]		[2574-2767]		[12020-12423]		[19158-19667]	
11:25 p.m.	613	0.7%	2,613	-0.5%	12,065	0.1%	19,310	0.7%
	[563-656]		[2533-2726]		[11858-12248]		[18932-19445]	
11:30 p.m.	632	4.8%	2,648	2.0%	11,917	0.1%	19,106	0.8%
	[559-648]		[2500-2689]		[11694-12107]		[18691-19225]	
11:35 p.m.	603	1.2%	2,574	0.3%	11,672	-0.7%	18,703	-0.1%
	[551-640]		[2468-2662]		[11547-11955]		[18456-18971]	
11:40 p.m.	607	3.3%	2,536	-0.1%	11,740	1.1%	18,711	1.3%
	[545-634]		[2449-2627]		[11415-11794]		[18224-18725]	
11:45 p.m.	584	0.8%	2,533	0.8%	11,652	1.7%	18,289	0.6%
	[537-622]		[2418-2609]		[11236-11656]		[17944-18436]	
11:50 p.m.	530	-7.2%	2,430	-1.8%	11,415	0.9%	18,037	0.8%
	[526-616]		[2379-2570]		[11114-11518]		[17640-18138]	
11:55 p.m.	575	2.5%	2,520	3.5%	11,336	1.7%	18,046	2.4%
	[518-603]		[2345-2531]		[10959-11329]		[17373-17869]	
0:00 a.m.	549	0.3%	2,390	-0.3%	10,732	-2.3%	16,722	-3.6%
	[506-589]		[2307-2489]		[10798-11186]		[17107-17612]	
0:05 a.m.	529	-0.9%	2,375	1.0%	10,660	-1.3%	16,986	-0.6%
	[494-576]		[2260-2440]		[10609-10987]		[16849-17332]	
0:10 a.m.	499	-4.5%	2,299	-0.5%	10,805	1.6%	17,030	1.0%
	[478-565]		[2218-2404]		[10449-10815]		[16618-17120]	
0:15 a.m.	516	0.6%	2,251	-0.9%	10,476	0.2%	16,600	-0.2%
	[475-552]		[2181-2365]		[10271-10637]		[16393-16870]	
0:20 a.m.	526	4.1%	2,233	-0.1%	10,240	-0.3%	16,567	1.1%
	[467-545]		[2149-2326]		[10081-10464]		[16165-16632]	
0:25 a.m.	488	-1.3%	2,173	-1.2%	10,010	-1.0%	16,069	-0.6%
	[454-537]		[2107-2288]		[9935-10293]		[15933-16403]	
0:30 a.m.	469	-3.7%	2,208	2.0%	10,028	0.8%	16,056	0.8%
	[450-529]		[2081-2254]		[9763-10138]		[15695-16158]	
0:35 a.m.	486	2.3%	2,025	-4.7%	9,767	-0.3%	15,633	-0.4%
	[435-515]		[2042-2209]		[9620-9987]		[15459-15936]	
0:40 a.m.	454	-2.4%	2,173	4.1%	9,730	0.8%	15,694	1.4%
	[426-503]		[2008-2170]		[9467-9840]		[15233-15707]	
0:45 a.m.	465	2.6%	2,060	0.4%	9,447	-0.6%	15,348	0.6%
	[415-493]		[1966-2137]		[9317-9677]		[15026-15464]	
0:50 a.m.	453	2.3%	2,018	-0.2%	9,432	0.7%	15,025	-0.1%
	[404-484]		[1943-2102]		[9177-9545]		[14804-15257]	
0:55 a.m.	465	6.8%	1,993	0.3%	9,247	0.3%	14,699	-0.9%
	[397-476]		[1904-2072]		[9058-9390]		[14607-15062]	

The "Actual" columns represent the actual number of hospital arrivals in each 5-minute time interval. The "Ratio" columns represent the ratio of excessive bunching calculated by comparing the actual numbers and the counterfactual numbers of hospital arrivals. The 95% CIs are reported in brackets. The shaded row indicates the time interval from 0:00 to 0:04.

Table 3: Summary Statistics

	Mean	SD	Min	Max
Manipulation	0.11	0.31	0	1
Proportion of Private Hospitals	44.28	19.72	6.81	77.06
Ln Population	14.45	0.72	13.29	16.02
Population Density	539.13	874.37	70.00	4,670.00
Elderly Population Ratio	24.65	2.59	17.40	29.60
Number of Emergency Care Hospitals	3.58	0.98	1.90	5.50
Ln Per Capita Taxable Income	4.77	0.15	4.44	5.15

Note: The number of observations was 46 prefectures, excluding Tokyo prefecture. Number of emergency care hospitals is per 1,000 population.

Table 4: Determination of Arrival Time Manipulation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Proportion of Private Hospitals	0.0074***	0.0043**	0.0047**	0.0061**	0.0072***	0.0066***	0.0046*
	[0.003]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Ln Population		0.1570**					0.0614
		[0.074]					[0.071]
Population Density			0.0001*				0.0001
			[0.000]				[0.000]
Elderly Population Ratio				-0.0213			0.0090
				[0.016]			[0.013]
Number of Emergency Care Hospitals					-0.0795*		-0.0339
					[0.039]		[0.042]
Ln Per Capita Taxable Income						0.4431	-0.0125
						[0.278]	[0.316]
Const.	-0.2198**	-2.3476**	-0.1743**	0.3638	0.0750	-2.2996*	-1.0787
	[0.086]	[1.042]	[0.067]	[0.414]	[0.118]	[1.342]	[1.497]
R Squared	0.22	0.31	0.34	0.24	0.28	0.26	0.36
Observations	46	46	46	46	46	46	46

Note: Results are based on ordinary least squares regression. Robust standard errors are in brackets. Number of emergency care hospitals is per 1,000 population.

Online Appendix

A Further Theoretical Discussion

In our model, the impact of financial incentives on the degree of SID was derived from Equation (2), and was calculated as follows:

$$\frac{dI}{d\theta} = -\frac{U_{\pi}(1-\rho)}{U_{\pi\pi}\theta^2 + U_{BB}(B_I)^2 + U_B B_{II}}, \tag{4}$$

where ρ is a parameter of relative risk aversion in hospital revenue ($\rho = -\frac{\pi U_{\pi\pi}}{U_{\pi}}$). Since the utility function is concave for B and π, the denominator is always negative with a reasonable assumption of $B_{II} \leq 0$. The numerator takes a positive value when ρ is less than 1. This assumption is plausible because EMS providers may be almost risk neutral with respect to the revenue of their hospital, implying that ρ takes a very low value. Finally, we have $\frac{dI}{d\theta} > 0$, which suggests that greater financial incentives trigger the manipulation of hospital arrival time. On the other hand, the value of $\frac{dI}{d\theta}$ may be very small if the treatment is harmful because it decreases when B_I takes a large absolute value. This indicates that financial incentives *per se* are not the fundamental cause of SID. Rather, the amount of harmful behavior based on SID is closely related to each physician's subjective evaluation on patient net benefit. This point echoes the findings from an experimental study by Hennig-Schmidt et al (2011).

B Determinants of Prehospital Transport Time

Many epidemiological studies have been conducted on the determination of prehospital transport time, but most focus on the time from the onset of symptoms rather than from the emergency call. This is because the time from the onset of symptoms to placement of the emergency call can account for a large portion of total prehospital delay, depending on the disease. For example, Fukuoka et al (2005) reported that the median prehospital delay time was 3 h 34 min for patients with acute myocardial infarction (AMI) in Japan, which was seven times longer than the average prehospital transport time (call time to hospital arrival time). In addition, we found no studies that investigated prehospital emergency time in general emergency episodes. Most studies on this issue have focused on specific diseases, such as AMI (Smolderen et al, 2010) and acute coronary syndrome (Mooney et al, 2014). Hence, the regression analysis presented here was designed to derive counterfactual values of hospital arrival times in general emergency patients using the procedure outlined in Section 5. Although our focus was limited in this study, the determinants of prehospital transport time in general emergency patients are of particular interest in Japan due to recent increases in prehospital transport time, which has become an issue of public debate.

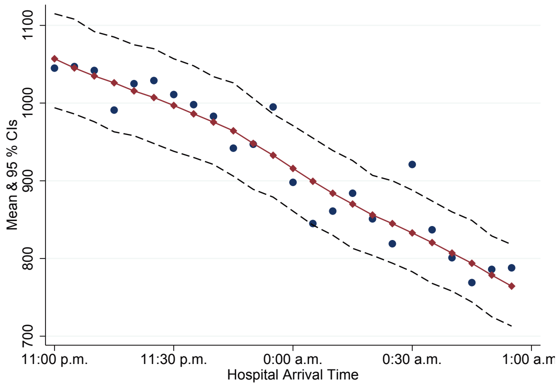
For the regression analysis, we included 2,146,498 emergency episodes. We included all emergency episodes where the 119 call was recorded as being placed between 21:00 and 3:00. Place of emergency occurrence, patient severity and causes of the emergency were controlled in the regression model. In addition, prefecture fixed effects, monthly dummy variables, and emergency call time fixed effects were also incorporated into the regression model. The summary statistics and estimated coefficients are reported in Table B1. The results showed that the place of emergency occurrence was associated with prehospital transport time. Compared with patients whose emergencies occurred at home, emergency calls from public facilities, workplaces and the road were more likely to have prompt transportations. When analyzed according to the severity of symptoms, our results showed that prehospital transport time for dead patients was the shortest because all hospitals can accept these patients. On the other hand, we found that prehospital transport time for living patients with more complicated and severe symptoms was likely to be delayed as these patients cannot be treated in primary EMS providers. The length of delay in patients with severe symptoms was approximately 3 minutes (0.9607+2.0438) longer than those with mild symptoms. In addition, emergency calls due to accidents and injuries had longer transport times when compared to emergency calls due to a disease. Transfers between hospitals took shorter times because these calls are likely to be promptly dealt with by hospital staff.

Table B1: Determinants of Prehospital Transport Time

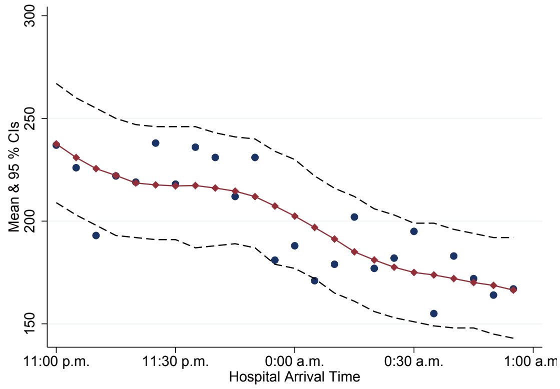
Variables	Definition	Mean [S.D.] (2)	Coef [S.E] (3)
Women	Woman = 1	0.54	-0.1356***
		(0.50)	[0.028]
Place of Emergency Occurrence			
Home	In the patient ’ s home = 1	0.69	
		(0.46)	
Public Facility	In a public facility = 1	0.17	-0.9274***
		(0.38)	[0.044]
Workplace	In the patient ’ s workplace = 1	0.01	-1.4145***
		(0.10)	[0.151]
Road	In a public road =1	0.11	-0.7954***
		(0.32)	[0.077]
Others	In other locations = 1	0.01	1.4445***
		(0.11)	[0.144]
Severity			
Dead	Dead patient, = 1	0.02	-3.3685***
		(0.13)	[0.103]
Severe	Patient with severe symptoms = 1	0.07	0.9607***
		(0.26)	[0.064]
Moderate	Patient with moderate symptoms = 1	0.35	
		(0.48)	
Mild	Patient with mild symptoms = 1	0.56	-2.0438***
		(0.50)	[0.034]
Cause			
Disease	Disease = 1	0.72	
		(0.45)	
Fire	Fire accident = 1	0.00	12.1525***
		(0.04)	[0.568]
Natural Disaster	Accident from natural disaster = 1	0.00	6.0893***
		(0.01)	[2.307]
Water	Water accident = 1	0.00	8.0572***
		(0.01)	[1.154]
Traffic	Traffic accident = 1	0.08	2.4737***
		(0.27)	[0.112]
Accident in Workplace	Accident in workplace = 1	0.00	2.2869***
		(0.05)	[0.287]
Sports	Accident due to sports and recreational activities = 1	0.00	0.6138*
		(0.04)	[0.322]
General Accident	Other causes	0.12	2.0052***
		(0.33)	[0.050]
Violence	Injured by someone else =1	0.02	3.7888***
		(0.12)	[0.137]
Attempted Suicide	Attempted suicide by the patient =1	0.02	6.7184***
		(0.12)	[0.139]
Hospital Transfer	Transfer from one hospital to another =1	0.05	-0.9184***
Prefecture Fixed Effects			X
Monthly Level Fixed Effects			X
Emergency Call Time Fixed Effects			X
Age Group Fixed Effects			X
Root MSE			14.1785
R Squared			0.101
Number of Observations			2,146,498

Note: Results are based on ordinary least squares regression. Standard errors are clustered around the emergency call time. Age groups were also controlled by 5-year dummy variables from 0 to 100 years of age. The definitions of patient severity are described in Section 4.1.

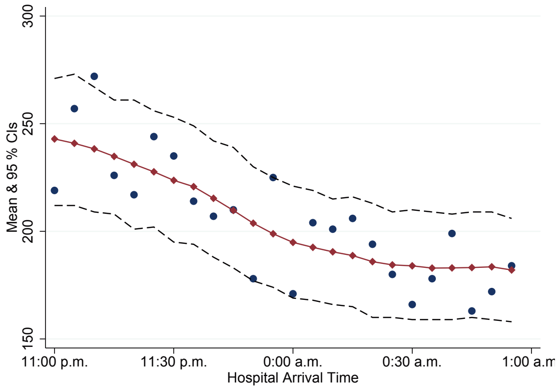
C Heterogeneous Effects by Prefecture



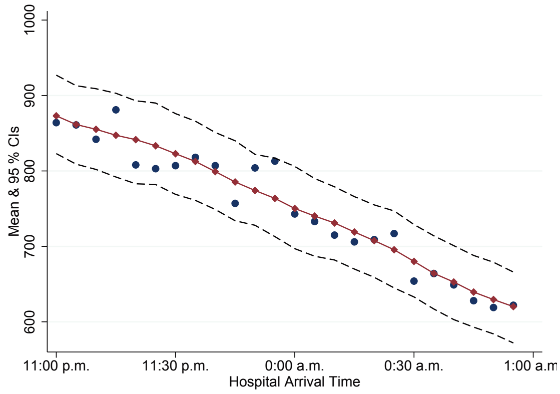
(a) Hokkaido



(c) Iwate



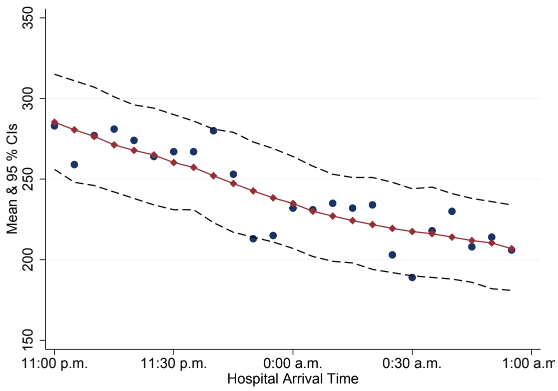
(b) Aomori



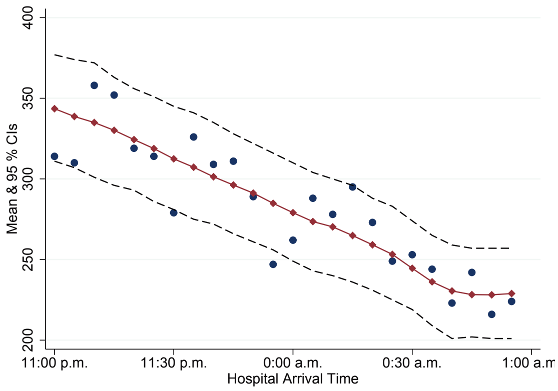
(d) Miyagi

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

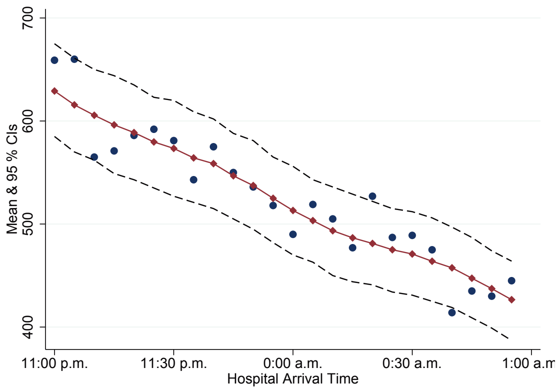
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



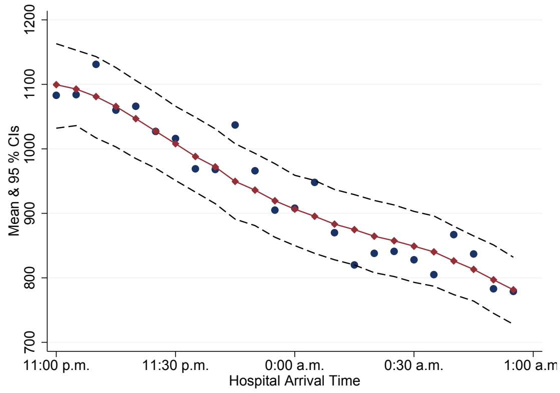
(a) Akita



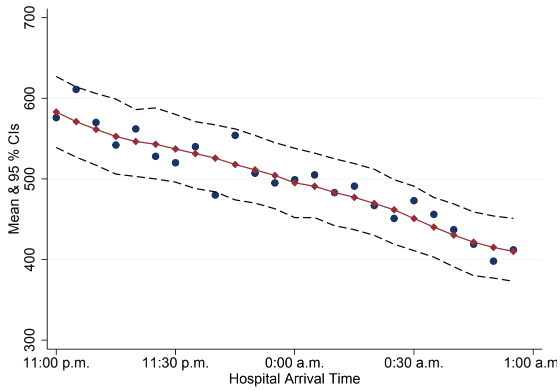
(b) Yamagata



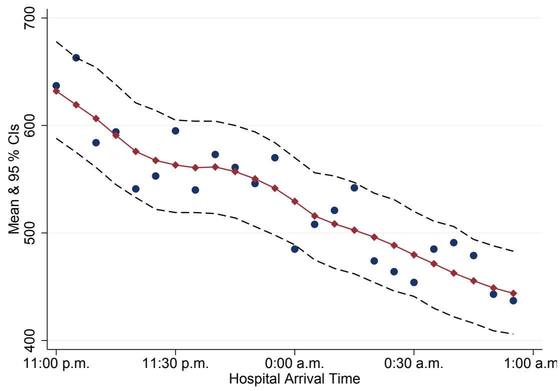
(c) Fukushima



(d) Ibaraki



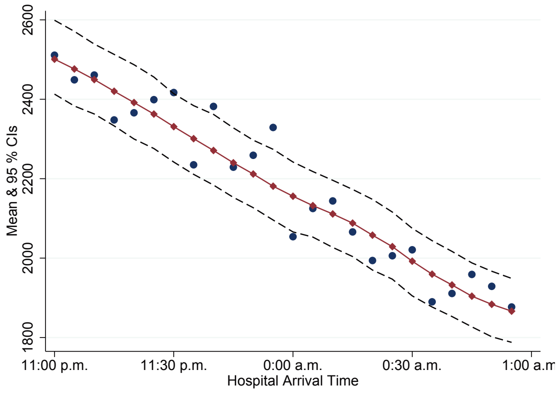
(e) Tochigi



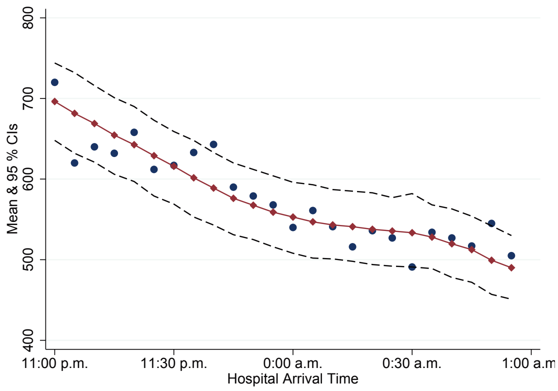
(f) Gunma

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

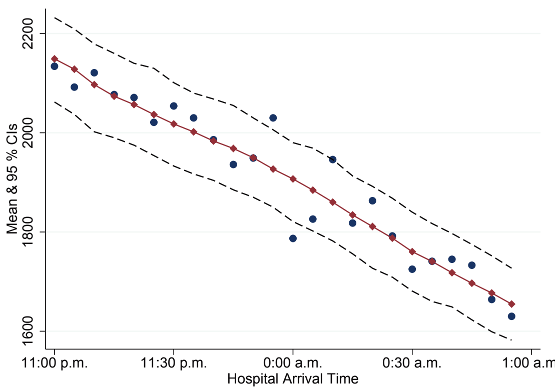
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



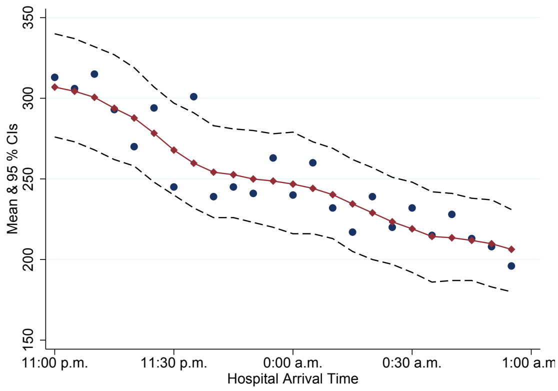
(a) Saitama



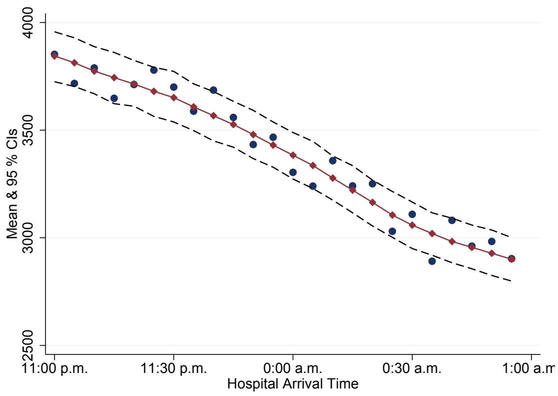
(d) Niigata



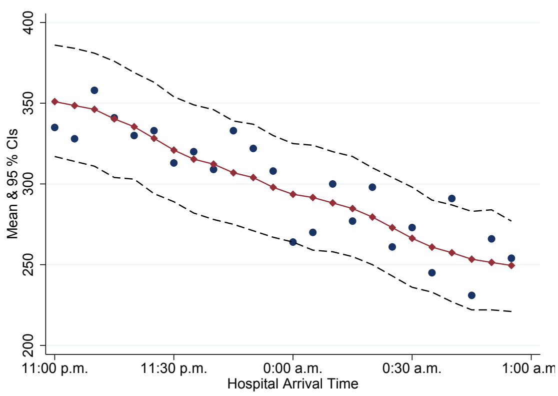
(b) Chiba



(e) Toyama



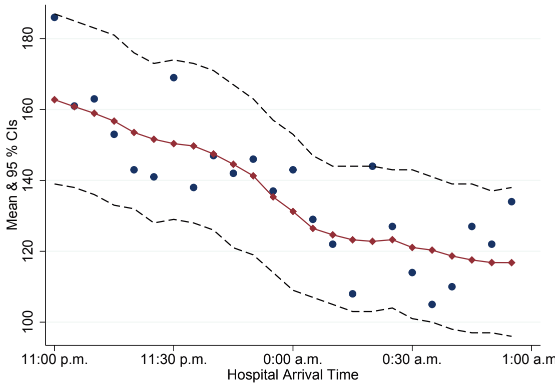
(c) Kanagawa



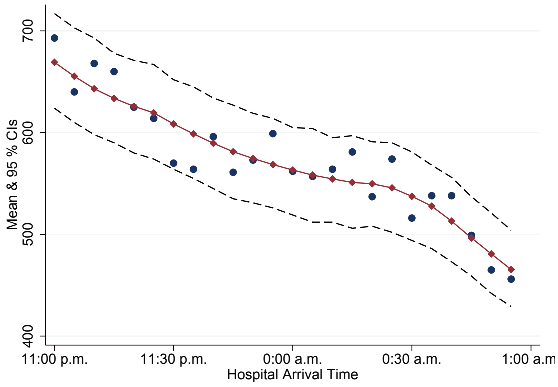
(f) Ishikawa

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

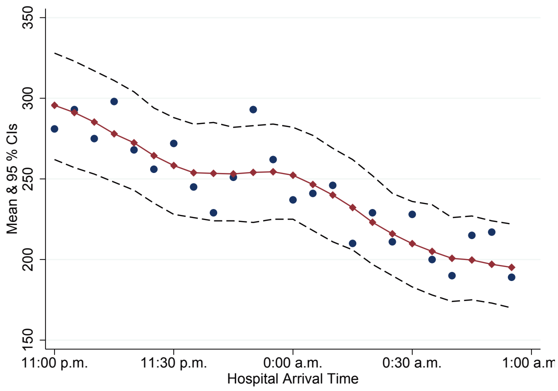
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



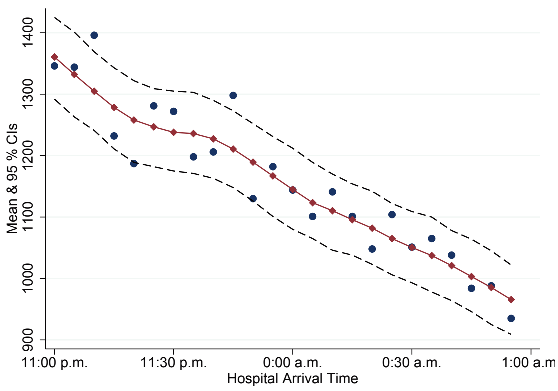
(a) Fukui



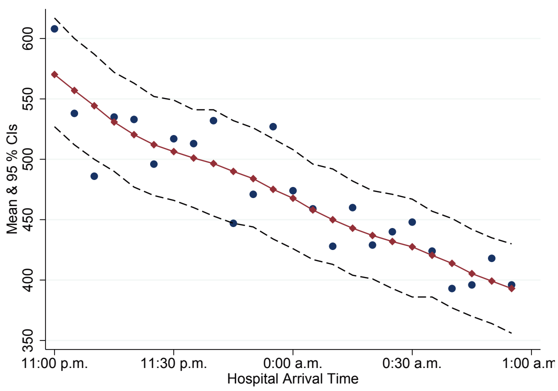
(d) Gifu



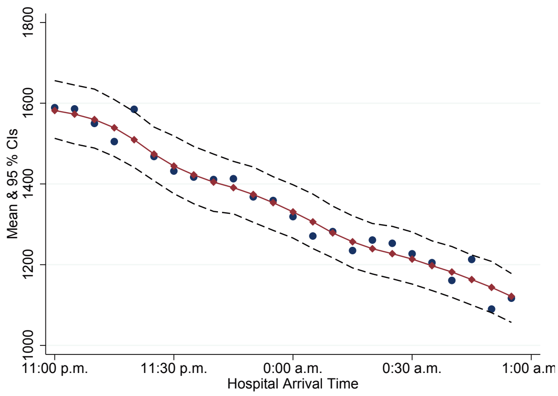
(b) Yamanashi



(e) Shizuoka



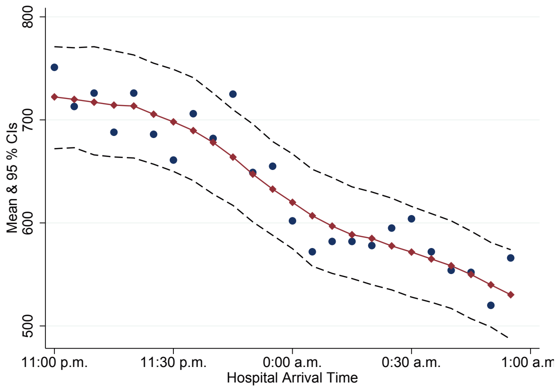
(c) Nagano



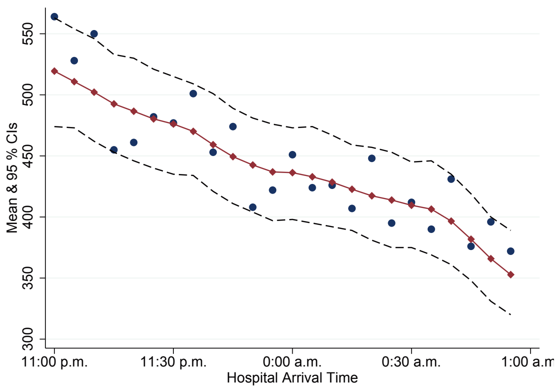
(f) Aichi

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

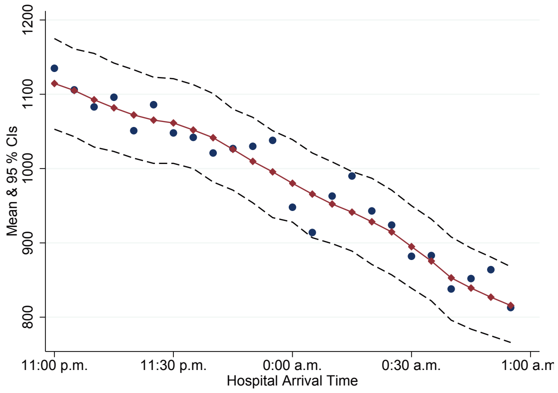
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



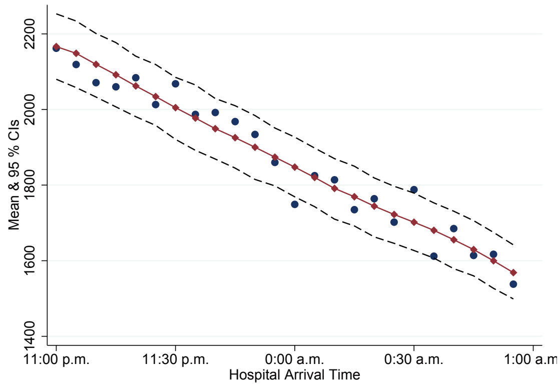
(a) Mie



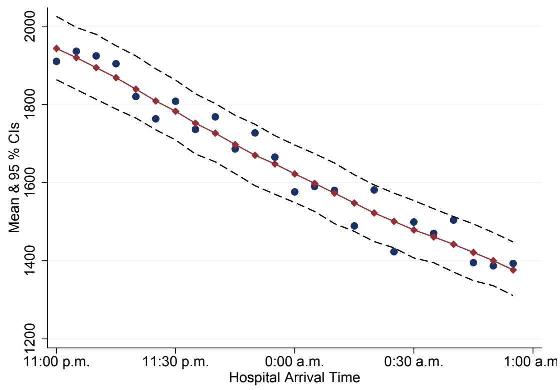
(b) Shiga



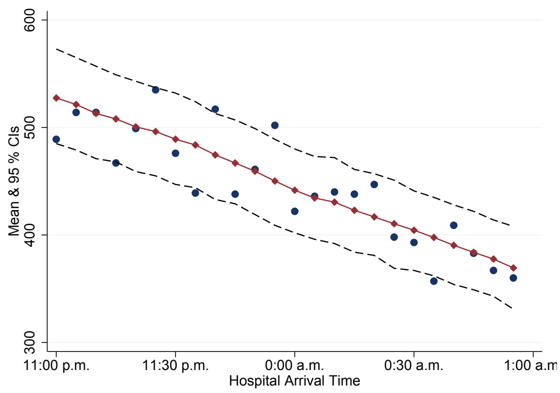
(c) Kyoto



(d) Osaka



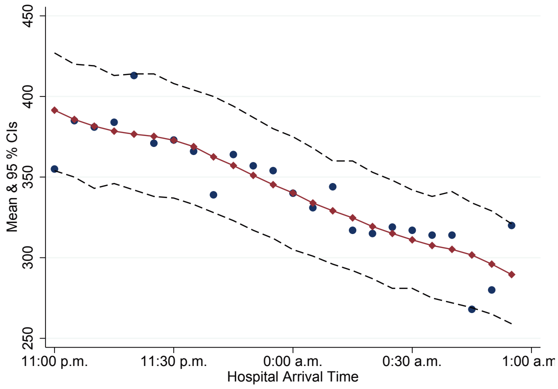
(e) Hyogo



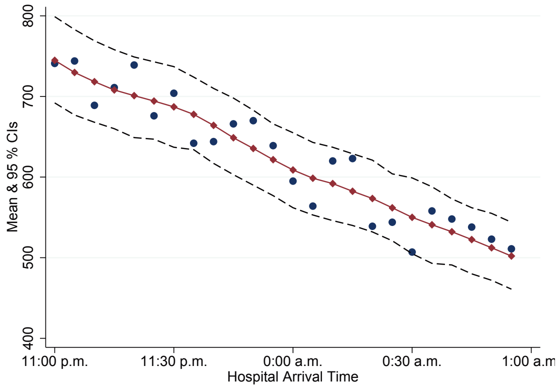
(f) Nara

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

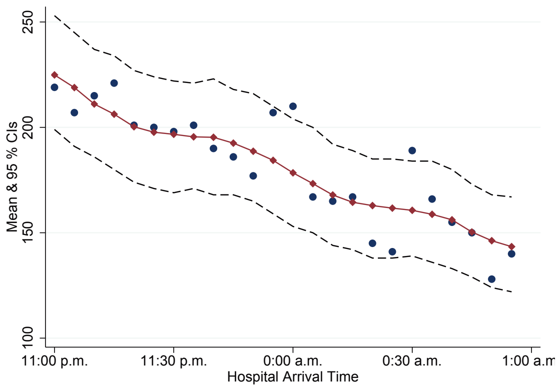
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



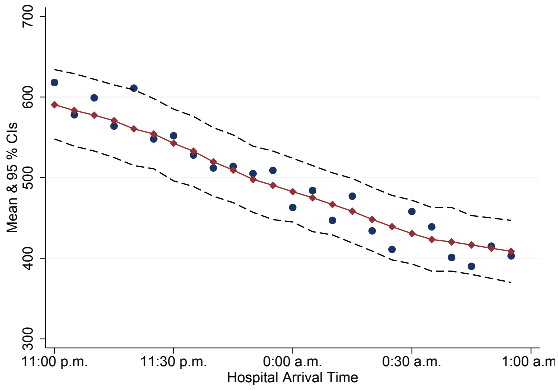
(a) Wakayama



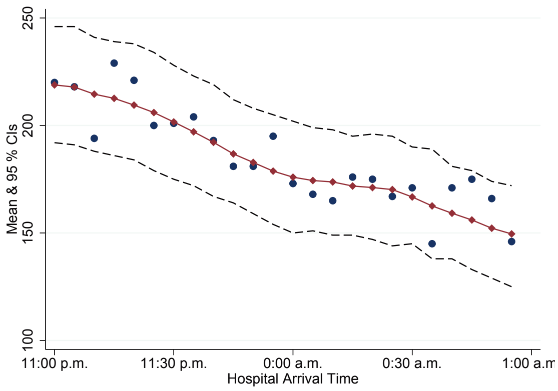
(d) Okayama



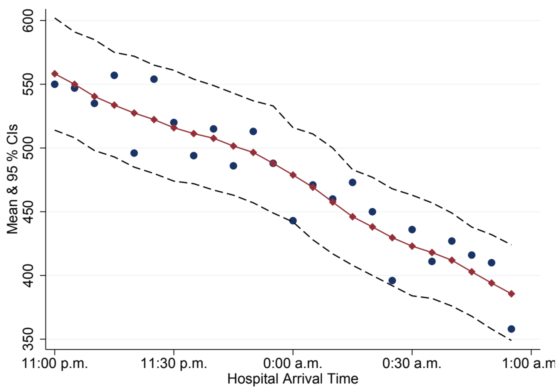
(b) Tottori



(e) Hiroshima



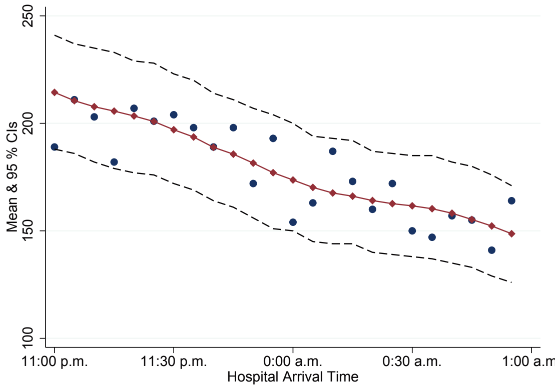
(c) Shimane



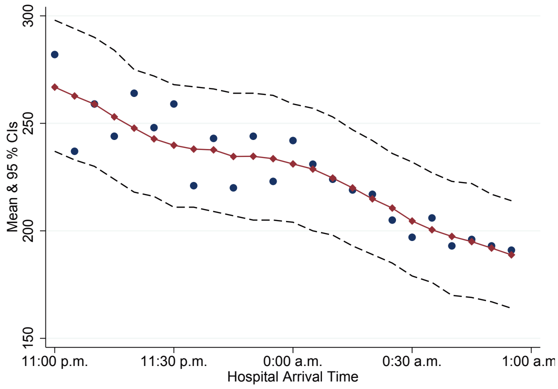
(f) Yamaguchi

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

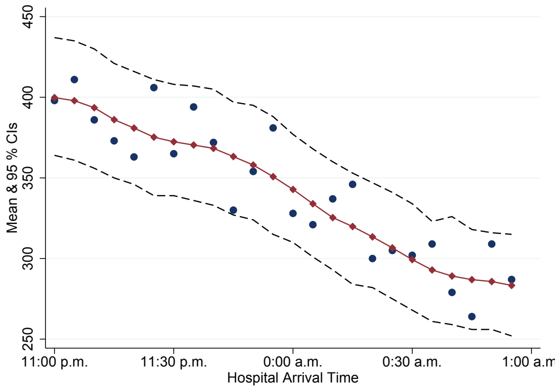
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



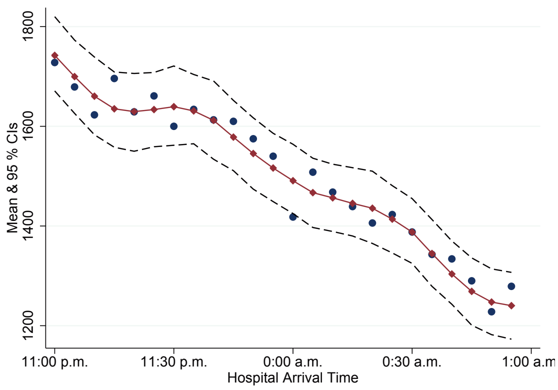
(a) Tokushima



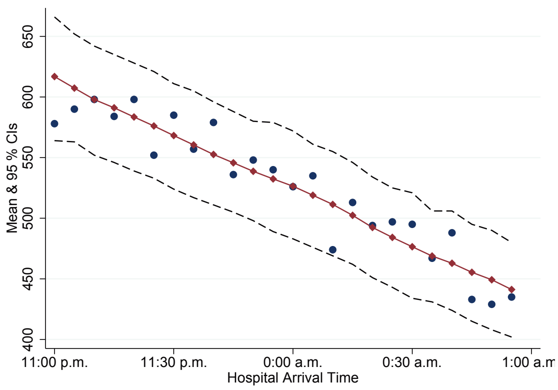
(d) Kochi



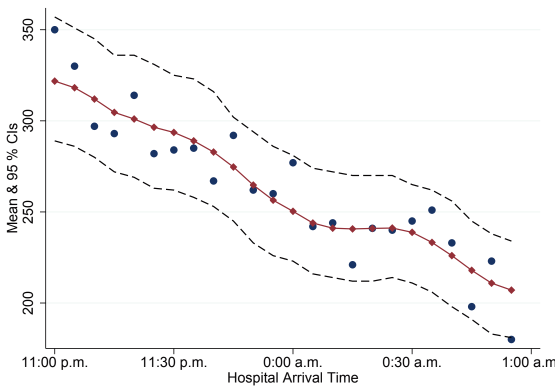
(b) Kagawa



(e) Fukuoka



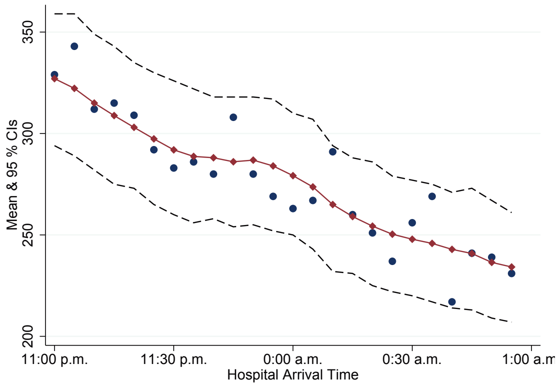
(c) Ehime



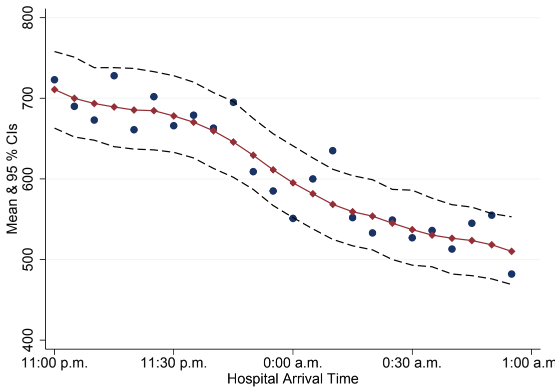
(f) Saga

Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

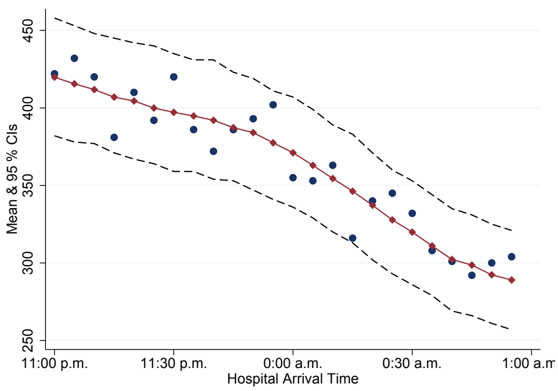
Figure C1: Number of Emergency Hospital Arrivals by Prefecture



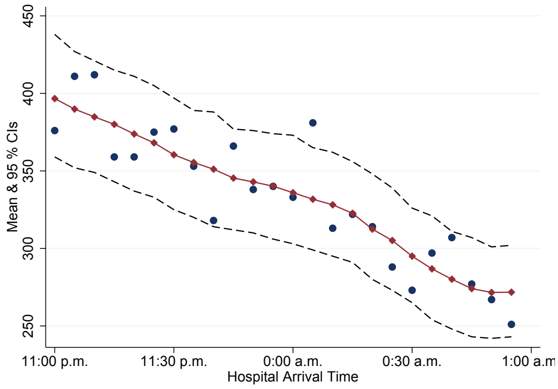
(a) Nagasaki



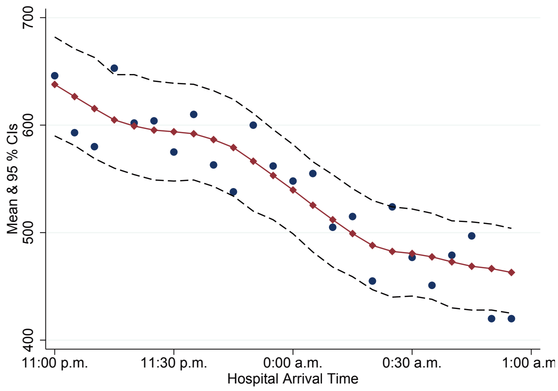
(b) Kumamoto



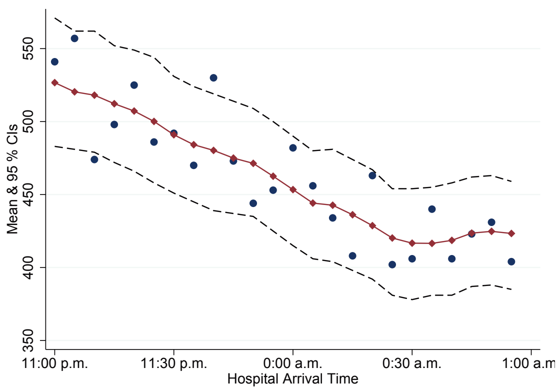
(c) Oita



(d) Miyazaki



(e) Kagoshima



(f) Okinawa

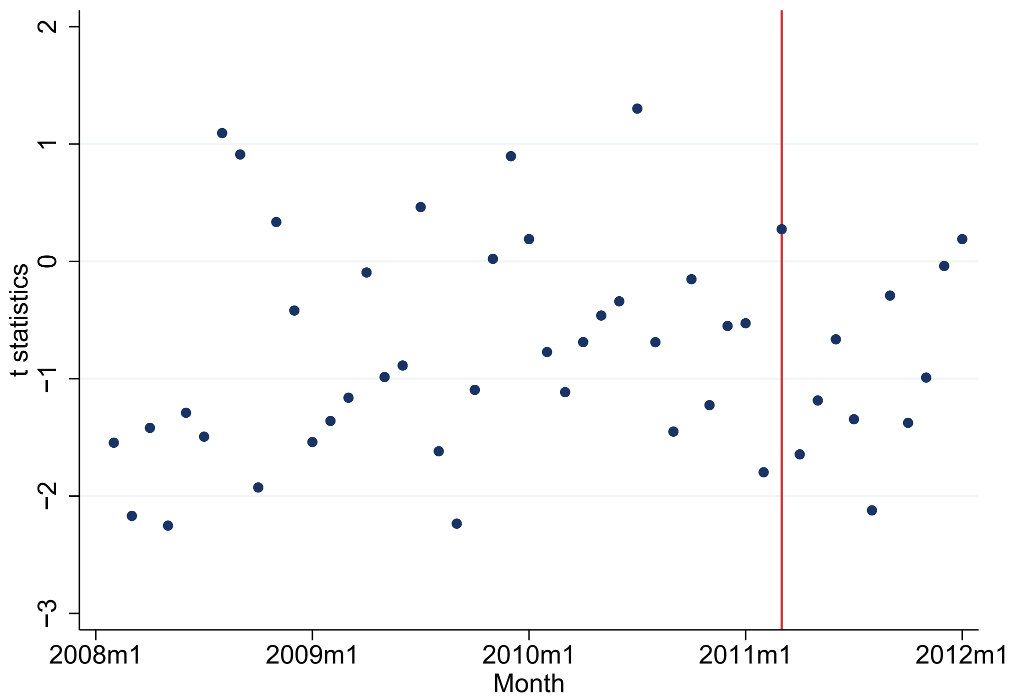
Circles represent the actual numbers of hospital arrivals. Diamonds with a solid red line represent the mean values of counterfactual numbers of hospital arrivals. Dashed lines denote 95% CIs.

Figure C1: Number of Emergency Hospital Arrivals by Prefecture

D Subsample Analysis by Month

In this Appendix, we describe the subsample analysis by month. Since our data covered all periods from January 2008 to December 2011, our sample comprised 48 (4*12) month-based subsamples. Following the same procedure for the main analysis, we derived the counterfactual hospital arrival times in the period from 0:00 to 0:04. In the interest of brevity, we do not present the results for all 48 subsamples here. Instead, we calculate the t statistics to evaluate how the actual count for each subsample deviates from the mean of 1,000 counterfactual counts. If the t statistic was negative and lower than -1.96, this indicated that some of the potential arrivals in the time period were shifted.

The results are summarized in Figure D1. In this figure, some of the circles (representing t statistics) are below -2, suggesting that the actual number of hospital arrivals in the period from 0:00 to 0:04 was significantly lower than the corresponding counterfactual number. In addition, although the t statistics were not significant for many of the months, the signs were consistently negative. The stability of the results over different months suggests that our primary results are not driven by irregular behavior in specific months or years. Importantly, we found no trend breaks before and after the Great East Japan Earthquake in March 2011.



Note: Circles represent t statistics that evaluate how the actual count for each subsample deviates from the mean of 1,000 counterfactual counts in the period from 0:00 to 0:04. The vertical line represents March 2011, when the Great East Japan Earthquake occurred.

Figure D1: Bunching Estimations by Month

Acknowledgement

The road to my PhD was never straight. For the first 3 years in my doctoral course, from 2009 to 2012, I was a full-time researcher in Japan Center of Economic Research (JCER), as well as a student of doctoral course in Keio University. Although the experience of full-time work contributed to my basic skills to deal with empirical data, my responsibility in the company was macroeconomic forecast which was not directly associated with my interest. With heavy load of tasks in JCER, I could not image that I would manage to finish PhD dissertation. In addition, I could not agree with the quality of my papers in this period, which deal with the issues in health economics. Although 3 papers written in Japanese are published in peer-reviewed journal, they are all far from completely plausible even for me because these papers lack clear research design. After discarding the approaches I adopted in these 3 papers, I began to study statistical causal inference from the beginning. Fortunately, flexible working condition in the Institute for Health Economics and Policy (IHEP), where I have joined since April 2012, provided sufficient time to rethink such a methodological issue. In addition, I am grateful for Michihito Ando, who was a PhD candidate in Uppsala University, for helping me to study these approaches. Without many discussions with him, I could not finish this dissertation.

The 5 chapters in this dissertation were written after joining IHEP. In the first year in IHEP, I spent a lot of time studying orthodox approaches to program evaluation such as instrumental variable, difference-in-differences and regression discontinuity design. In 2013, I started to apply these methodologies to health care in Japan with a grant-in-aid for scientific research provided by Japan Society for the Promotion of Science. Thanks to the eligibility for this grand-in-aid, I could utilize the questionnaire information of the Comprehensive Survey for Living Conditions. In addition, Nobuyuki Izumida from the National Institute of Population and Social Security Research generously shared the insurance claims of the National Health Insurance with me. Given that the availability of those data are very restricted for young health economist in Japan, I am very fortunate.

My supervisor, Yoshikazu Kenjoh, always provides many critical comments for all of my manuscript, while these comments are sometimes so incisive that I cannot reply sufficiently. It is one of my lifetime goal to answer all of his comments. In addition, I would like to thank many researchers. Atsuhiro Yamada, Isamu Yamamoto and Shun-ichiro Bessho have already listened my presentations on the many chapters in this dissertation, and provided helpful suggestions. During the course of my studies, I have had the opportunity to work with some great researchers. In particular, discussion with Naoki Ikegami and Shuzo Nishimura have all been important sources of inspiration. I also thank Rei Goto, Junya Hamaaki, Haruko Noguchi, Sayaka Nakamura and Yoko Ibuka who are all lead economist in the field of health economics in Japan.

Finally, many research projects in this dissertation have come from my life, rather than from reading economic journals. Since September 2012, happy future for my daughter has become the strongest motivation to complete my studies. In addition, I owe a great deal of support to my nearest family····my wife and my parent····all to whom the dissertation is dedicated.