Uterus at a Price: Disability Insurance and Hysterectomy

Elliott Fan
Department of Economics, National Taiwan University

Hsienming Lien
Department of Public Finance, National Chengchi University

Ching-to Albert Ma[T]
Department of Economics, Boston University

June 2017

Abstract

Three social disability programs in Taiwan (Government Employee Insurance, Labor Insurance, and Farmer Insurance) provide compensation equal to about five months of salary to female enrollees who become infertile due to hysterectomy and oophorectomy when these treatments happen before enrollees' 45th birthday. These programs create incentives for more and earlier treatments, which we call, respectively, the inducement and timing effects. Using medical records from 1997 to 2011 from Taiwan National Health Insurance Administration, we track of enrollees and the uninsured between their 39th and 50th birthdays. We use difference-in-difference and nonparametric methods to estimate inducement and timing effects of hysterectomy and oophorectomy hazards. Accounting for timing effect, disability insurance results in 11-12% additional hysterectomies for enrollees in Labor Insurance and Government Employee Insurance, but only 3% for enrollees in Farmer Insurance. For oophorectomy, both effects are negligible.

JEL: I00, I10, I12, I18
Keywords: disability insurance, moral hazard, hysterectomy, oophorectomy

---

Emails respectively: elliottfan@ntu.edu.tw, hmlien@nccu.edu.tw, and ma@bu.edu
We thank Stacey Chen, Minchung Hsu, Tor Iversen, Kevin Lang, Tzu-Ting Yang and seminar participants at Boston University, Brandeis University, Institute of Economics at Academic Sinica, Queen's University, University of Bologna, and University of Oslo for their valuable comments and suggestions. We thank Johannes Schmieder for generously providing us with codes that assisted us in the nonparametric estimation, and the Taiwanese National Health Research Administration for providing National Health Insurance Research Data. Supports from National Science Council (MOST 105-2410-H-004-004 to Hsienming Lien) are greatly appreciated. The paper represents the authors' views and does not reflect those of the National Health Insurance Administration or the National Science Council.

1. Introduction

In this paper, we study the effect of disability insurance on the moral hazard of health care use. Disability insurance compensates enrollees for accidental or inherited loss of physical and mental capacities, which adversely affect labor market potentials (Maestas, Mullen and Strand, 2013; Staubli, 2011; French. and Song, 2014). Taiwan adopts a set of mandatory, employment-based comprehensive disability programs. While the incentive consequences of disability insurance are well known, the Taiwanese programs have an uncommon component: they include coverage for infertility of women due to i) hysterectomy (surgical removal of the uterus), ii) oophorectomy (surgical removal of both ovaries), and iii) chemo-radiotherapy. Enrollees are entitled to a cash benefit, equal to about five months of salary, when they undergo these treatments, but the coverage ends when enrollees become 45 years old.

The research here is on disability insurance's effect on the incidence of hysterectomy and oophorectomy, both of which involve the removal of major organs. We are unaware of the same infertility coverage or age limit in any other social insurance program (In Japan, benefits due to infertility suffered at work are covered, but the Taiwan programs are universal, not limited to injury at work.) Although it may sound sinister to suggest that enrollees undergo serious organ-removal surgeries to qualify for disability benefits, the economic perspective is that the coverage, perhaps unintentionally, has created such an incentive. Furthermore, there is an incentive to undergo surgeries before the benefit expires at the 45th birthday. Does disability insurance lead to excessive treatments that would not have been performed in the absence of the insurance, an inducement effect? Does disability insurance lead to earlier treatments that have been expedited to before the 45th birthday, a timing effect?

Only hysterectomy exhibits significant inducement and timing effects; oophorectomy does not. These are striking results. First, we can interpret this particular infertility coverage as a natural experiment that offers the same benefit on infertility due to two different medical interventions. Although both hysterectomy and oophorectomy will lead to infertility (qualifying enrollees for benefit), oophorectomy has far more serious adverse health consequences than hysterectomy (more thorough discussions in

Subsection 2.1). In the very unlikely scenario that an enrollee had chosen to undergo one of the two procedures solely for the disability benefit, the preferred procedure would have been hysterectomy due to lower disutility. The more likely scenario is that a medical intervention is indicated due to an illness. Our results indicate that the monetary disability benefits would have no effect on oophorectomy, but lead to more hysterectomy, again, likely because hysterectomy is associated with a lower disutility. This natural experiment yields behaviors that are consistent with optimizing behavior---even when the decision involves a major surgery to remove an organ.

Second, our results indicate the way the disability insurance is implemented may lead to serious deadweight loss. About 11-12% of hysterectomies in our sample are unnecessary. The deadweight loss is the cost of this surgical procedure, precisely because the qualification requires it. Linking a disability to a surgical procedure creates double moral hazard, one in the claim of disability, and another in health care use. Although the infertility coverage in Taiwan is unique, our results serve as a warning against using a medical treatment as a qualification for disability insurance benefits.

Third, our results shed some light on issues surrounding monetary incentives and organ donation. In almost all countries, direct financial compensations to donors of transplant organs or tissues do not exist. However, our natural experiment does involve financial compensations for someone whose organ is removed---although uterus or ovaries are not used for transplants. Furthermore, the disability insurance programs are sanctioned by the Taiwanese government, and appear to be well run.

Our evidence suggests that individuals are able to make consistent choices. The lack of inducement and timing effects in oophorectomy perhaps is the strongest indication that the usual benefit-cost analysis commonly practiced in economic activities is in force. In fact, the evidence suggests that a well-organized or regulated market may work as expected even for a human organ. For a given price of an organ, if the disutility of its removal is sufficiently high, individuals reject the offer, as in the case of oophorectomy, but if the disutility is sufficiently low, individuals accept it, as in the case of hysterectomy. In fact, we have found that, in the income-stratified samples, induced-hysterectomy rates are increasing in the benefit level.

To avoid the exploitation of the poor in organ donation, our finding supports a donation payment that is positively related to a donor's income.

We estimate the inducement and timing effects by a difference-in-difference method and by a nonparametric method. Our difference-in-difference design is based on the comparison of enrollees in three disability insurance programs and those uninsured for the period between 1997 and 2011 in Taiwan. The three programs are Labor Insurance, Government Employee Insurance, and Farmer Insurance. The control consists of those who have not belonged to these programs, mostly those inactive in the labor force. Our data are from Taiwan's National Health Insurance, which indicate the type of disability insurance for each individual, and if and when hysterectomy, oophorectomy, or both surgeries have taken place. In addition, the data are merged with variables obtained from the Survey of Family Income and Expenditure (SFIE) to help control for socio-demographic and economic factors.

We follow female enrollees in the three insurance programs and in the control between their 39th and 50th birthdays. In the main analysis, we include only enrollees who have not changed their insurance groups in the sample years. We then separate enrollees by their birth cohorts and insurance groups, and calculate the corresponding hazards of hysterectomy or oophorectomy. Our main variable is the quarterly hazard of surgeries in the number of quarters (a 91-day period) from the 45th birthday. Because we follow enrollees for 11 years, there are 24 such benefit quarters before the 45th birthday, and 20 quarters after. For each birth cohort, we also use the average numbers of children and sons, marital status, and household income as covariates.

Our difference-in-difference design is unconventional because there is not a before-and-after policy regime change. However, the insured lose the infertility insurance benefit at age 45. Obviously, that benefit expiration is irrelevant to the uninsured. Our hypothesis is that when there is a lot of time before the benefit expires, the insured's behaviors are unaffected by the deadline. However, at some later point, the deadline becomes pressing. To operationalize our empirical strategy, we let the deadline become relevant when enrollees turn 40 years old. In other words, we treat the benefit expiration at the 45th birthday as a policy intervention on disability-program enrollees at their 40th birthday. Our approach is conservative; quarters just after the 40th birthday need not exhibit differences. If the relevance of the deadline only appears at some time after the 40th birthday, our difference estimates would simply vanish.

Using the hazards before age 40 as the benchmark, we examine the dynamic, quarter-by-quarter hazard differences between the insured and uninsured, for the 5 years before and 5 years after the 45th birthday (20 quarters before and after the benefit expires). How do the hazards differ when enrollees are approaching their 45th birthday? How do they differ thereafter? Indeed, for hysterectomy, Labor Insurance and Government Employee Insurance enrollees' hazards begin to rise rapidly 8 quarters before expiration, but drop rapidly for 2 quarters after. Enrollees in Farmer Insurance show similar but less pronounced hazard changes. For oophorectomy, these rapid changes are absent in all insurance programs.

From the estimates we calculate inducement and timing effects. The inducement effect is the total number of insured enrollees' extra surgeries between their 40th and 50th birthdays compared to the control. The timing effect consists of the total number of surgeries that the insured would have undergone after the 45th birthday compared to the control. In our sample, enrollees in Labor Insurance have a total of 43,845 hysterectomies, and the inducement effect is 5,076 hysterectomies, or about 11.6%. For Government Employee Insurance, the total is 7,262, and the inducement effect is 789, or 10.9%. For Farmer Insurance, the total is 9,100, and the inducement effect is 347, or 3.8%. (Later, we will provide all the details of these three insurance programs, which partially account for some of the magnitude differences.) No inducement or timing effects have been found for oophorectomy.

Besides the difference-in-difference estimation, we also use a nonparametric method. Here, we use the bunching method (e.g. Saez, 2010; Chetty et al., 2011) and assume that hazards follow a smooth time path if the disability benefit would never expire. Conceptually, we maintain the counterfactual assumption that surgeries do not happen abruptly over time. Operationally, we use hazards in benefit quarters far from the 45th birthday to fit a polynomial. Then we use the fitted polynomial to predict the hazards in benefit quarters near the 45th birthday. Any discrepancy between the predicted and actual hazards will be attributed to benefit expiration. These discrepancies will be used to define the inducement and timing effects as in the difference-in-difference approach. In any case, the nonparametric estimates of the inducement and timing effects for hysterectomy and oophorectomy are similar to those in the difference-in-difference method.

As a check, we estimate inducement and timing effects of two other surgeries: partial oophorectomy (the removal of one ovary), and myomectomy (the removal of the inner lining of the uterus). These procedures are used to alleviate problems in the female reproductive system, but do not qualify for infertility insurance benefit. We have found no inducement or timing effects.

We also consider a number of robustness issues. Our primary sample consists of female enrollees who have not switched between insurance programs. We relax this restriction by including those who have switched between programs. This makes the sample larger. Furthermore, our data include medical records of women aged between 39 and 49 years old during the period 1997 to 2011. Therefore, the panel is unbalanced; early cohorts are subject to left censoring while late cohorts are subject to right censoring. Later we restrict our sample to those with uncensored medical records in the sample period. Our results are robust to various compositions of samples.

The plan of the paper is as followed. In the next Subsection, we review the literature. We present the study background in Section 2. Section 3 describes the data for the study, and the construction of our sample of enrollees who have not changed insurance status throughout the entire sample period. We also present sample statistics. In Section 4, we present the two econometric methods. Subsection 4.1 is on the difference-in-difference method, whereas Subsection 4.2 is on the nonparametric method. In each case, we set up the regression equations, and define the inducement and timing effects in terms of the parameters to be estimated. The two subsections of Section 5 contain the estimation results of both methods. Then in Section 6, we consider a number of robustness checks. We consider a bigger sample that includes individuals who have switched insurance programs. We also consider smaller sample in which individuals are never censored. Then we perform various robustness checks based on more or less restrictive sample definitions. In addition, we stratify the sample of labor enrollees according to five benefit levels, and examine the size of inducement with respect to benefit levels. We draw some conclusions in Section 7.

1.1. Literature review

Insurance benefits that are based on age and time are quite common. Medicare in the United States provides health insurance to individuals over 65 years old. In most countries, unemployment benefits expire after a period of time. The incentives of insurance benefits that are based on time and enrollees' age influence the enrollees' behaviors. For instance, research has shown that patients delay treatment or surgeries until they become eligible for Medicare (McWilliams et al., 2003; McWilliams et al., 2007; Card, Dobkin and Maestas, 2008; Card, Dobkin and Maestas, 2009). Research also has shown that recipients of unemployment insurance delay job search until benefits are about to expire; this delay leads to a spike of the unemployment duration distribution around the expiration point (Caliendo, Tatsiramos, and Uhlendorff, 2013; Farber and Valletta, 2015; and Schmieder, von Wachter, and Bender, 2016). In our study setting, similar incentives exist: the infertility benefits of the Taiwanese disability program expire at age 45. However, our setting has one major difference: benefit qualification requires hysterectomy or oophorectomy, and Taiwan's National Health Insurance covers these surgeries. In other words, an enrollee's incentive to qualify for the benefit implies a second incentive for a surgery.

Our empirical strategy uses a modified form of difference-in-difference regression, and a nonparametric method. Difference-in-difference regression is the standard method for program evaluations and policy assessments (see Imbens and Wooldridge, 2009 for a review). However, the standard method considers only the average policy effect. Here, we go beyond the average effect to study policy effects over time, particularly the periods right before and after benefit expiration. As in Chandra, Gruber, and McKnight (2010), we use quarter-by-quarter estimates for the policy effect over time. Autor, Kerr and Kugler (2007) use a similar year-by-year difference-in-difference model to understand how mandated employment protections may reduce productive efficiency. Hoynes, Miller and Simon (2015) also use the same method to study how earned income tax credit influences infant health outcomes.

Our nonparametric method is similar to the bunching method for assessing discontinuity effects created by policies. For example, taxes can be discontinuously related to reported incomes (Saez, 2010; Chetty et al., 2011; Kleven and Waseem, 2013), tax reliefs may be available to couples only if marriages or child births happen before a certain date (Persson, 2014), or teachers use discretion in their grading of nationwide math tests so that students' test scores bump up over key grade cutoffs (Diamond and Persson, 2016). The standard assumption is that, absent the policy, the variable of interest should vary smoothly over time, so any bunching is due to the policy. We use the same basic methodology. However, our method is more closely related to Diamond and Persson (2016) in that we use an optimality criterion---minimum mean-squared errors---to determine the manipulated regions and then estimate the counterfactual polynomial of surgeries.

2. Background

2.1. Hysterectomy and oophorectomy

Hysterectomy is the surgical removal of a woman's uterus, the organ that holds the fetus. This is the second most common elective surgery among women, after cesarean section for childbirth. Usually performed by a gynecologist or an obstetrician, hysterectomy can be either complete (removal of the uterus and cervix) or partial (without the removal of the cervix). There are three variants of the surgical procedure: abdominal, vaginal, and laparoscopic. Hysterectomy carries a minimal morbidity risk, at a mortality rate below 0.05%. Complications, such as bleeding and dysfunctional uterine parity, are also rare (McPherson et. al., 2004). The length of hospital stay for the procedure ranges from 3 to 5 days.

Hysterectomies are performed mainly for uterine fibroids and malignant tumors in a woman's reproductive system.[1] Common indications are menstrual irregularities, such as heavy bleeding, and serious pain (Department of Health and Human Services, 2011). Alternative treatments for some of these indications are available. Myomectomy---the surgical removal of some uterine lining---may be a remedy for uterine fibroids. Endometrial ablation---surgical removal of endometrium---may be suitable for excessive bleeding. Pain medication, synthetic steroid hormones, and pelvic floor exercises are other alternatives.

The incidence of hysterectomy exhibits an age pattern over a woman's lifetime: the rate rises steadily from ages 30 to 39, reaches a peak between 45 and 49, and then declines steeply (McPherson, Gon

[1] In a random sample of 658 Taiwanese women, the most common indication for hysterectomy was uterine fibroids (at 46.2%), followed by malignancy and pre-malignancy (at 22.2%) (Wu et al., 2005).

and Scott, 2013). Despite an obvious age pattern, incidence rates vary substantially across different countries. According to OECD Statistics, in 2012, the average hysterectomy incidence rate was 179 per 100,000 women. However, in the same year, the hysterectomy incidence rate in Germany was 318 per 100,000, whereas it was only 49 per 100,000 in Denmark (McPherson, Gon and Scott, 2013).

For Asian countries, the incidence rate (per 100,000 women) for South Korea was 198 in 2012 (OECD Health Data, 2016). In Taiwan, with a population of 23 million, an average of 23,000 hysterectomies are performed each year. From 1996 to 2005, the hysterectomy incidence rates per 100,000 in Taiwan varied between 268 and 303 (Wu et. al, 2010). According to National Health Insurance Data, when Taiwanese women become 50 years old, more than 20% of them would have had hysterectomies.

Oophorectomy is the surgical removal of a woman’s ovaries, the organs that hold her eggs. The primary indication for oophorectomy is ovarian cancer. The common technique to perform oophorectomy is abdominal laparoscopy. Oophorectomies and hysterectomies are sometimes performed in the same operation, mainly for preventing cancer. A partial oophorectomy, the removal of only one ovary, is commonly performed for medical indications such as endometriosis, benign tumor, inflammation, etc., so may not be combined with a hysterectomy.

The removal of ovaries causes hormonal changes and the spontaneous onset of menopause. Most patients are advised to start hormonal replacement post operation. The health risks associated with total oophorectomy are more serious---women who have had oophorectomies before 45 years old carry a mortality risk 170% higher than women who have retained their ovaries (Parker et al., 2005). The incidence of complete oophorectomy is relatively low. Between 1997 and 2011, for instance, less than 2,300 completed oophorectomies among women aged between 40 and 49 were performed in Taiwan each year. By contrast, more than 10,000 hysterectomies were performed in the same age group each year during the same time period.

2.2 Disability insurance

In Taiwan, three mandatory social insurance programs provide disability insurance to the working population. Enrollment is only for the individual; there is no family coverage. Labor Insurance is the largest program, covering nearly 9 million workers in the private sector in 2012. When it was first established in 1956, it provided only health insurance, but by 1978 insured enrollees had coverage for disability, maternity, occupational injuries, unemployment, pension, and death. After 1995, health insurance in Labor Insurance was replaced by Taiwan's National Health Insurance.

The second largest social insurance program is Farmer Insurance. In 2012, this program covered 1.5 million farmers. Government Employee Insurance, the third program, is for public employees and teachers in both public and private schools and colleges. In 2012, Government Employee Insurance covered about 0.6 million lives. Similar to Labor Insurance, Farmer Insurance and Government Employee Insurance provide a portfolio of benefits, which include disability insurance.[2]

With few exceptions, disability insurance benefit is paid as a lump sum;[3] the benefit amount varies according to the type and severity of disabilities. In this research, we focus on the disability benefit for a female enrollee's loss of her reproductive function. A woman is eligible for this disability benefit if, due to illnesses, she undergoes any of three medical procedures before turning 45 years old: hysterectomy, complete oophorectomy, and radio-chemo therapy on ovaries. The disability insurance benefit is not meant to compensate for medical expenses because Taiwan's National Health Insurance covers most in-patient medical expenses.[4]

Recipients of insurance benefits are mostly patients who have undergone hysterectomy. In contrast, complete oophorectomy and radiation and chemotherapy are less common. Our hypothesis is that disability insurance creates an incentive that the cash benefit may cause excessive treatments. The effect of disability insurance on hysterectomy is plausible because the adverse consequences of hysterectomy are relatively mild. On the other hand, radio-chemo therapies have very serious consequences and side effects, so the insurance benefit is unlikely to have any effect on such decisions. Oophorectomy also carries

[2] Government Employee Insurance does not provide unemployment insurance. Farmer Insurance does not offer unemployment insurance or pension scheme.
[3] Disability insurance does provide long-term benefits for those who contract chronic illnesses or have accidents that impair their capacity to work.
[4] The co-insurance rate of inpatient services for Taiwan's National Health Insurance is 10%, with spending caps. In 2011, the cap per admission and per year are NT28,000 and NT47,000 respectively (In 2015, the exchange rate was about NT$30 to US$1).

adverse consequences, but we will study the effect of insurance on this treatment. As a check, we will study partial oophorectomy and myomectomy, which do not qualify for insurance benefit.

The disability benefits are calculated according to an enrollee's "insurance salary", to be defined next; they are 6 months of insurance salary in Government Employee Insurance, and 5.3 months in Labor Insurance and Farmer Insurance. For Farmer Insurance, the insurance salary is fixed at NT$10,200 per month, so the reproductive disability benefit is fixed at NT$54,060 (= 10200*5.3). (In 2015, the exchange rate is about NT$30 to US$1.) For Labor Insurance, in 2013, the insurance salary is defined to be the lower of an enrollee's actual monthly salary and NT$43,900. In other words, the benefit is 5.3 months of salary for those with monthly pay below NT$43,900, but will correspond to a smaller number of months of salary for those with higher pays. For Government Employee Insurance, the insurance salary is the lower of an enrollee's base monthly salary and NT$53,900. However, the base salary does not include various stipends (e.g. research stipends for teachers) so an enrollee in Government Employee Insurance typically has actual monthly earnings higher than the base salary.[5] For those with base salary below NT$53,900, the benefit is lower than six months of actual earnings. For those with higher base pays, the benefit will correspond to an even smaller portion of monthly salary.

3. Data and samples

3.1. Data sets

Our sample period spans the years between 1997 and 2011. The subjects in our study are females born between 1948 and 1972, and we study their experiences in the time between their 39th and 50th birthdays during the sample period. We use three data sets. The first is the set of hospital claims of Taiwan National Health Insurance between 1997 and 2011. The claims data include all inpatient admissions in Taiwan because National Health Insurance covers the entire population. Each claim includes a patient's demographics (gender and date of birth), admission date, disease diagnoses, medical reimbursement, and any surgery performed during the admission. Each claim also has scrambled unique identifiers for a

[5] For instance, the base salary and the research stipend for an assistant professor in 2012 was approximately the same, at NT$41,755 and NT$39,555, respectively.

patient, doctors and hospitals. We use the surgical-procedure information to identify those who have undergone hysterectomy, oophorectomy, myomectomy, and partial oophorectomy. We use a patient's date of birth and admission date to check whether hysterectomy and oophorectomy have been performed before the 45th birthday, the disability insurance benefit qualification criterion.

Our second data set is the National Health Insurance enrollment file. The file contains the last entry of each enrollee's insurance program and disability insurance salary at the end of a calendar year.[6] We first use an enrollee's insurance type to infer the disability insurance status. National Health Insurance started in 1994 by merging many private and public insurance programs, and its enrollment file has continued to track enrollees' other social insurance modules. Therefore, from the enrollment file, we classify subjects' disability insurance status into four groups: Government Employee Insurance, Labor Insurance, Farmer Insurance, and otherwise uninsured. However, the current-year insurance program status may be inappropriate if some enrollees change insurance status and programs after undertaking hysterectomy. Later we use an enrollee's disability insurance status in the previous year as a robustness check.

Next, we obtain enrollees' disability benefit information in the National Health Insurance enrollment file. Recall that the disability benefit is determined by an enrollee's salary up to a maximum (for Government Employee Insurance and Labor Insurance) or by a fixed amount (for Farmer Insurance). National Health Insurance charges a premium equal to a percent of an enrollee's monthly salary up to NT$188,000, which is much higher than the salary caps for disability insurance benefits. Therefore, from the National Health Insurance premium, we can infer an enrollee's salary, and, in turn, the benefits. This inference is exact for enrollees in Labor Insurance. Government Employee Insurance uses the base salary, a fraction of the total salary, for benefit calculation, so the enrollee's salary in the National Health Insurance enrollment file will over-estimate the benefit. (For this reason, our analysis in Subsection 6.2

[6] Up until 2002, the National Health Insurance enrollment file contains full disability insurance enrollment records. From 2003 onward, the enrollment file only contains enrollees' last disability insurance record in a calendar year; it no longer tracks an enrollee's disability insurance program changes during the year. For consistency, we use the last disability insurance record even for years before 2003.

will be based on Labor Insurance enrollees.) The disability benefit in Farmer Insurance is fixed, so we do not need to use salary information from National Health Insurance.

Our third data set is from the Survey of Family Income and Expenditure, conducted by Taiwan's Directorate General of Budget, Accounting and Statistics. Each year the survey randomly samples 13,000-16,000 households (or about 52,000–68,000 individuals) and collects information on socio-demographic characteristics of each member of the sampled households. For our sample period 1997-2011, we obtain the following information of female respondents who are in the corresponding age group (39-49 years old): highest education level attained, marital status, number of children by gender, monthly household earnings, and disability insurance type. We then use the insurance type information to merge with the enrollment files to control for demographics of enrolled populations.

3.2. Samples

We define our sample in the following way. First, our research is on the effect of disability insurance benefit. This is only available to female enrollees undergoing hysterectomy and oophorectomy before they turn 45 years old. To investigate the insurance program effects, we follow enrollees' decisions for six years before, and for five years after the expiration of the benefit.

Next, we impose a number of restrictions. We remove those in Labor Insurance whose enrollments were through trade union memberships, because their reported incomes might be unreliable.[7] We also delete a small number of enrollees who were in military or welfare programs, because their access to health services might be different.

Furthermore, enrollees may change social insurance programs through employment changes. Whereas most will not change jobs in order to switch from Labor Insurance to Government Employee Insurance principally because of the small difference in disability-insurance benefits, the possibility cannot be excluded for all. The strategic switch between disability insurance programs creates a selection problem.

[7] Labor law in Taiwan requires private companies with five or more employees to purchase Labor Insurance for all employees. Self-employed or those who work in firms with fewer than 5 employees are not required to participate, or they can participate through the trader unions. Salaries of these workers are often unstable or under-reported. For the comparison between insured salary and earned salary in various insurance groups, see Lien (2011).

To control for this selection, for the main analysis we use a sample of enrollees who have never changed their insurance status within the sample period. We call this the non-switching sample. The general sample refers to all female enrollees regardless of any change in social insurance programs during the sample period.

To gauge the extent of strategic switch, we construct the transition matrix of insurance groups for women underwent hysterectomy in two consecutive years. For this comparison, we include enrollees of disability insurance programs through trade unions. From Table B1 in Appendix, there is less than 15% probability of women transiting from an uninsured status to an insured status, of which about 7% to Labor Insurance and 8% to trader unions. The majority of these uninsured enrollees (about 85%) continue to stay in the uninsured group after having the hysterectomy. For women in other insurance groups, the relatively smaller "off-diagonal" transition probabilities indicates they do not change insurance groups after the surgeries. Therefore, we restrict the main analysis to a nonswitching sample---enrollees who have never changed their disability insurance status within the sample period. However, we will check the robustness of our estimates by employing a general sample that includes the switching individuals.

To illustrate the composition of our sample, Table B2 shows the birth cohorts and their corresponding ages in year 1997 and 2011 respectively in nonswitching sample. Because we have included all females whose partial or complete experiences between their 39th and 50th birthdays would happen in the period from 1991 to 2011, our data are both left and right censored. For instance, in year 1997 sample, the oldest members, those born in 1948, would be 49 years old, so would leave the sample period one year later. Naturally, those females do not experience incentive change (45 years old) during the sample period. Likewise, the youngest cohort, those born in 1958, would be 39 years old, so would be in the last year of the sampler period. In year 2011 sample, the oldest and the youngest cohort would be 1962 and 1972 respectively. Again, none of these cohorts experience the change of incentive benefits in the data.

In either general or non-switching samples, we allow censoring. However, those who were born between 1958 and 1962 would spend all their years between the 39th and 50th birthdays within the sample period. We will also use a balanced sample, so study only females born between 1958 and 1962. Both the general and balanced samples will be used for robustness checks.

3.3 Summary statistics

The first half of Table 1 presents the summary statistics of the nonswitching sample in the years 2000, 2005 and 2010. The number of subjects in the nonswitching sample ranges from 0.93 to 1.23 million in the years 2000, 2005, and 2010. Because each subject in the sample is selected between 39 and 49 years old in that sample year, there is a significant change in the distribution of birth cohorts across the years. In 2000, for instance, subjects born between 1960 and 1964 account for 17.9% of the subjects in that year, but those born between 1950 and 1954 account for 38%, and those born between 1955 and 1959 account for the rest. In 2005, the birth-year distribution shifts forward: none were born between 1950 and 1954, but 33.0% were born between 1955 and 1959, with the largest group (45.7%) being born between 1960 and 1964. For the year 2010, the subjects' birth-year distribution follows the same forward-shift pattern.

In addition to birth year, Table 1 shows the distributions of the enrollees' insurance groups in the nonswitching sample. The percentages of enrollees having Government Employee Insurance seem to be quite stable in the three years, ranging from 9.6% to 10.2% in the sample. Labor Insurance has the largest share of enrollment, about 50%, in each of the three years. However, the share of Farmer Insurance enrollments gradually declines over time; this is likely due to the diminishing and aging farmer population. Finally, the shares of the uninsured females seem to exhibit a slightly downward trend, declining from 31.0% in 2000 to 27.5% in 2010.

The second half shows the corresponding figures of the general sample. In contrast with the nonswitching sample, the general sample has a higher percentage in labor insurance and uninsured group. This is because enrollees in government employee and farmer insurance are unlikely to change insurance groups, so their share becomes smaller when this restriction of same insurance group no longer applies. Likewise, a higher percentage of older cohorts can be observed since those enrollees are more likely to switch between insurance groups (e.g. from the employed to the unemployed).

The lowest part of Table 1 displays the incidence rates of four reproductive-organ related surgeries in the nonswitching sample. Whereas the incidence rates of hysterectomy fall from 1,128 (per 100,000) in 2000 to 695 in 2010, myomectomy incidence rates almost double between 2000 and 2010. This likely indicates that myomectomy has become a more effective treatment for those suffering from uterine fibroids. Myomectomy may also become a more popular substitute for hysterectomy. For oophorectomy, the incidence rates of complete oophorectomy decline over time, but the opposite is true for partial oophorectomy. Better diagnosis and more conservative treatments may have been behind this trend. Similar trends can be also found in the general sample.

4. Econometric methods

Our hypothesis is that an enrollee's decision to undergo hysterectomy or oophorectomy is significantly affected by the disability insurance, a hypothesis that is based on cost and benefit. Nonetheless, our data do not allow us to test this directly at the individual level because we do not have information of individual or household income, marital status, or number of children. However, from SIFE we obtain these variables for birth cohorts. Therefore, we aggregate medical records to construct cohort hazards. In any case, each enrollee in a disability insurance program faces the same incentives. If individual surgery decisions respond to incentives, enrollees as a group also respond similarly. Our data allow us to test our hypothesis at the cohort level reliably.

We define a cohort of enrollees by two discrete time scales: (1) the calendar quarter of her chronological age, and (2) by the number of quarters from her 45th birthday. The first time scale measures her age. A birth cohort is represented by the vector $c \equiv (y, s)$ where y is the birth year and s is a season, or a three-month period of the year. Our sample consists of female enrollees born between 1948 and 1972, so we have 100 (= 25 years $\times$ 4 seasons) birth cohorts, with y taking values of 1948, 1949,…, and 1972, and s taking values of 1, 2, 3, and 4.

The second time scale measures how much time an enrollee has available before, or elapsed after, the expiration of the disability insurance benefit. We call the second time scale an enrollee's benefit quarter, and denote it by the variable q. The 91-day period that begins with the 45th birthday is called quarter 0. The next 91 days is quarter 1, whereas the 91 days before quarter 0 is quarter -1, and so on. Enrollees in our sample are between 39 and 49 years old, so the benefit-quarter variable q takes values -24, -23,…, -1, 0, 1,…,19. By making distinctions between year, season, and benefit quarters we allow for more decision variations.

Clearly, the choice of a 91-day length for a time unit, both for chronological and benefit dimensions, is for convenience and practicality. A shorter time length may imply sharper differences between adjacent cohorts because treatment incidences occur less frequently, whereas a longer time length reduces the number of groups. (We have tried defining the cohort length to six months, and have verified that results are similar.)

For each birth cohort in a given benefit quarter, the hysterectomy hazard is defined to be the ratio of the total number of enrollees undergoing hysterectomy within this benefit quarter to the total number of enrollees who have not undergone hysterectomy at the beginning of the benefit quarter. We define analogously the hazards of total oophorectomy, partial oophorectomy, and myomectomy. All the hazards are calibrated separately for the three insured groups and the uninsured group. In Figure 1, we present four panels. Each panel is for a surgery, and plots the hazards of the three insured groups and the control group. The grey curve plots the hazards of the uninsured; the red curve is hazards of enrollees in Labor Insurance, whereas the blue and green curves are for those in Government Employee Insurance and Farmer Insurance, respectively. Recall that hysterectomy and oophorectomy qualify for disability insurance benefit if they are performed by the 45th birthday, but partial oophorectomy and myomectomy do not qualify.

The four panels show some striking patterns. First, in Panel A, enrollees' hysterectomy hazards in Labor Insurance and Government Employee Insurance exhibit a sharp increase just before the 45th birthday, but drop significantly right after; a similar but less pronounced pattern can also be observed for those in Farmer Insurance. After a few quarters past the 45th birthday, hazards of all insured return to the same smooth trend. However, hazards of uninsured enrollees follow a smooth pattern throughout the entire time.

In Panel B, total oophorectomy hazards follow an increasing trend. However, it is unclear Government Employee Insurance and Labor Insurance enrollees' hazards show an accelerated increase just before the 45th birthday. In Panels C and D, myomectomy and partial oophorectomy hazards do not exhibit any abrupt changes, either for the insured groups or uninsured group.

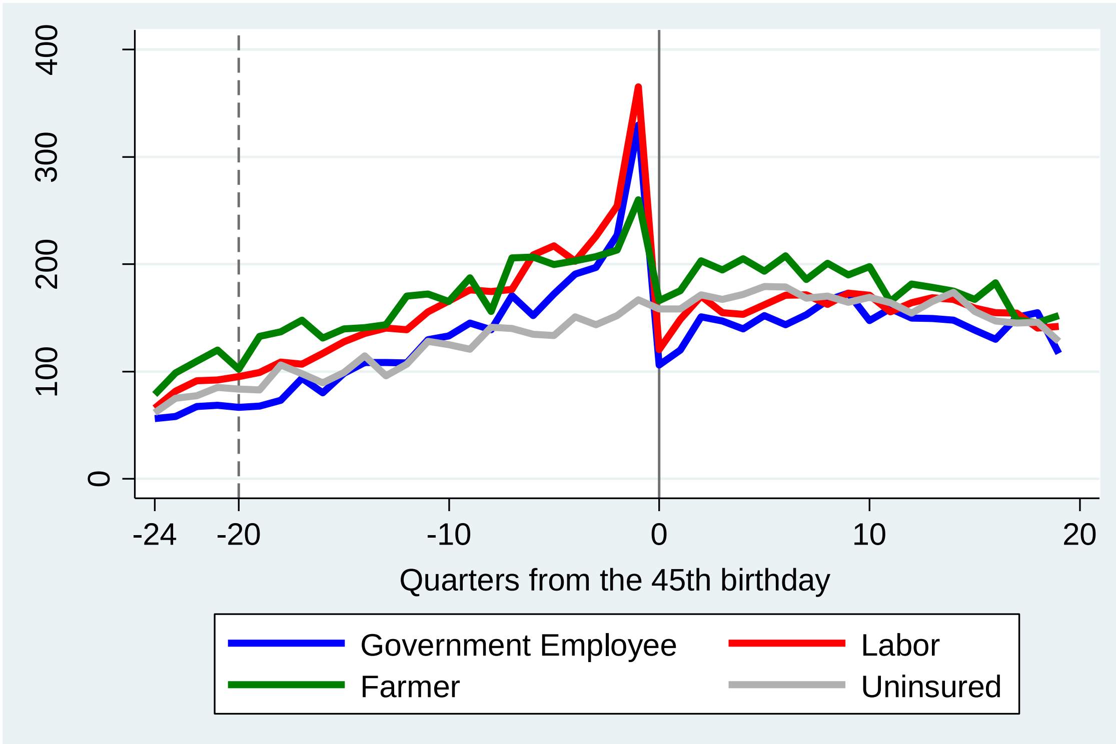
There is no medical literature to support the pattern of hysterectomy among female enrollees in Government Employee, Labor, and Farmer Insurances. Adverse uterine conditions cannot be especially serious in the few quarters before the 45th birthday, but the opposite will happen the few quarters after. Our hypothesis is that such a pattern is caused by the disability-insurance benefit when enrollees turn 45 years old.

The hypothesis is consistent with the lack of any abrupt changes in myomectomy and partial oophorectomy hazards. Total oophorectomy indeed qualifies for the disability benefit when it is performed by the 45th birthday. However, the removal of both ovaries carries much higher short-term and long-term health risks than the removal of the uterus. In other words, because the same benefit is paid to both hysterectomy and oophorectomy, we should expect different responses because these two treatments carry different explicit and implicit health costs.

For hysterectomy, the plots in Panel A suggest a timing manipulation effect. It appears that some hysterectomies might have been moved earlier to qualify for disability benefits, a timing effect. There is, however, a more serious possibility. Some hysterectomies might not have been performed if the insurance benefits were absent, an inducement effect. We estimate these effects by two methods. The first is based on the difference-in-difference method: enrollees in the three insurance programs are the treatment groups, and uninsured females are the control group. Our method allows the treatment group to experience different effects in each benefit quarter, so we estimate dynamic, quarter-by-quarter differences between the insured and the uninsured. The implementation is different from the conventional method, and will be explained in the next subsection. The second is a nonparametric method based on a smoothness hypothesis: there should not be any abrupt changes in enrollees' probability of undergoing hysterectomy at the 45th birthday if there were no disability benefit. The nonparametric method estimates a counterfactual hazard distribution for the insured as if the disability benefits were absent. Kleven and Wassem (2013) introduced this "bunching" method to estimate labor supply elasticity using discontinuity points in income distribution caused by piecewise linear progressive tax rates. Again, we have to adapt that methodology for our sample and the variable of interest, and will explain details later.

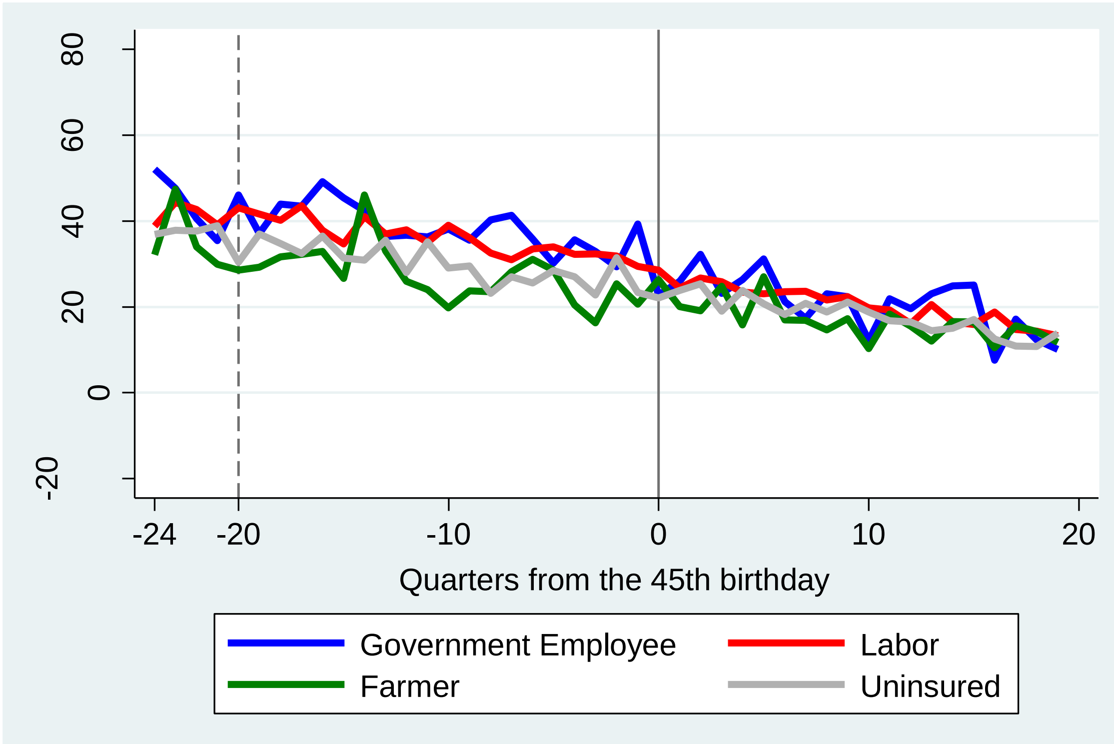
Figure 1

(A) Hysterectomy hazards

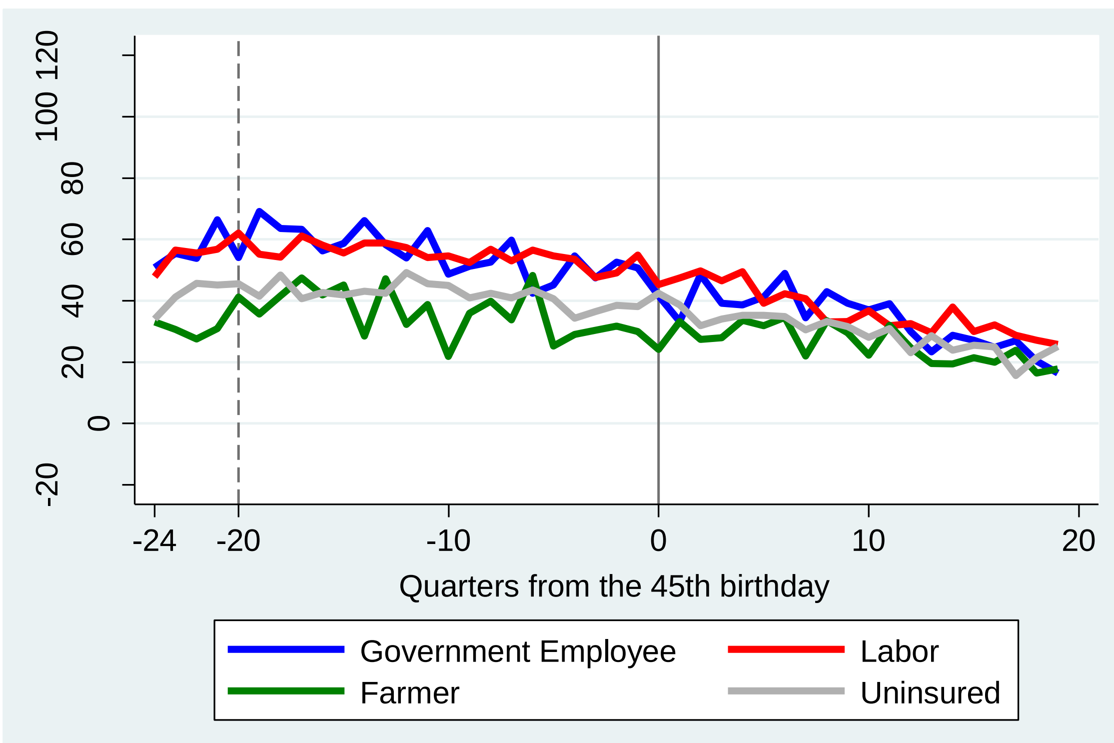


(B) Total oophorectomy hazards

(C) Partial oophorectomy hazards



(D) Myomectomy hazards



4.1 Difference-in-difference by benefit quarters

Obviously, we need to estimate the difference of surgery experiences between the insured and the uninsured, the first difference. However, the disability insurance programs have been in place for the entire sample period, so there are no before-and-after regimes. Our innovation is to exploit the benefit expiration at enrollees' 45th birthday to create a second difference. Clearly, no point in time has special significance to the uninsured's surgery decisions. Our identification assumption is this: at a point in time far before the 45th birthday, the benefit expiration also has no significance to the insured's surgery decision. Notice that we do not claim that the benefit availability has no impact on surgery decisions; we only claim that when the deadline is far into the future, the deadline has no impact on a current decision.

To implement this empirical strategy, we assume that the benefit expiration deadline becomes relevant when enrollees turn 40 years old. It is as if the benefit expiration at the 45th birthday is a policy intervention imposed on insured enrollees starting from their 40th birthday, a pseudo-intervention date, so we have chosen the 4 quarters between the 39th and 40th birthdays as the omitted benefit quarters. This is a conservative way to implement the estimation strategy. From Panels A and B in Figure 1, the hysterectomy and oophorectomy hazard time trends are similar around enrollees' 40th birthday (benefit quarter -20). To guard against an improperly imposed pseudo-intervention date, we actually examine the quarter-by-quarter differences between the insured and the uninsured. If the insured actually thought that the deadline was relevant at a later time, say the 41st birthday, the quarter-by-quarter differences would vanish in the four quarters between the 40th and 41st birthdays in our difference-in-difference setting.

The benefit expiration deadline at 45 years old actually affords us a clear identification of the disability insurance effect. Under a benefit expiration deadline, any deviation from trends around the deadline must be due to the disability insurance, so that the identification is guaranteed. In addition, if the benefit never expired, there would be no timing effect. The presence of expiration enables to differentiate the inducement and timing effect. Furthermore, the expiration deadline allows us to examine how female enrollees respond the quarter-by-quarter dynamic response in face to the insurance benefits.

The three treatment groups (Government Employee Insurance, Labor Insurance, and Farmer Insurance) have different disability benefits, and a separate regression is estimated for each group. For brevity, we present the regression equation for hysterectomy; the regression equations for the other three procedures are set up analogously. The regression equation is:

$$H_{c,q} = \alpha + \beta \times Insured + \sum_{i=-20}^{19} \gamma_i \times 1[i=q] + \sum_{i=-20}^{19} \rho_i \times 1[i=q] \times Insured$$
$$+ \mathbf{X}_{c,q}\boldsymbol{\eta} + \sum_{j=1998}^{2011} k_j \times 1[j = T(c)] + \varepsilon_{c,q} \quad (1)$$

$$c \equiv (y,s) \text{ with } y = 1948,\ 1949,....,1972, \text{ and } s = 1,\ 2,\ 3,\ 4, \text{ and}$$
$$q = -24,\ \text{-}23,\ \text{-}22,....,\ 1,\ 0,\ 1,\ 2,....,\ 19$$

In equation (1), $H_{c,q}$, either for the insured or the control, is the hysterectomy hazard of birth cohort *c* at benefit quarter *q*, a birth cohort *c* being defined by birth year *y* and birth quarter *s.* The variable $Insured$ is the dummy variable for an insured group (Government Employee Insurance, Labor Insurance, or Farmer Insurance). It is set to 1 if the hazard belongs to the insured, and 0 otherwise. The function 1[ ] is an indicator, and set at 1 if the condition inside the square brackets is satisfied; it is set at 0, otherwise. The covariates $\mathbf{X}_{c,q}$ are cohort-cell means of variables of the total number of children, the number of sons, marital status, and log household incomes. The function $T(c) \equiv \text{Int}[y + 45 + s/4]$ is the smallest integer bigger than (y + 45 + s/4), so its range is simply the years between 1997 and 2011; this function is used for the data-year fixed effects. Finally, $\varepsilon_{c,q}$ is the error term. To mitigate the concern of serial correlation in the error term, we allow for the clustered standard error.[8]

In equation (1), parameter γ_q is the mean hysterectomy hazard difference between quarter *q* and the benchmark, which we have taken to be quarters $q = -24$ and $q = -21$, the four quarters in age 39. (Our results change only slightly when the benchmark age quarters are extended to those quarters corresponding to ages from 35 to 39). For $q = -20$,..,19, the parameter ρ_q measures the incremental difference between

[8] For details as to why the standard errors of the coefficient estimates of interest tend to be underestimated in the difference-in-difference model, see Bertrand, Esther, and Sendil (2004) and Donald and Lang, (2007)

the insured and the uninsured, our chief focus. If disability insurance does not affect enrollees' hysterectomy decisions, all estimates of ρ_q should be zero. However, if these estimates are nonzero, we can estimate how disability insurance impacts hysterectomies under the deadline.

In the specification in equation (1), we are able to differentiate the inducement effect and the timing effect. The inducement effect is the total increment of hysterectomies of the insured over the uninsured in the period between the 40th and 50th birthdays. Let $\hat{\rho}_q$ denote the estimate of ρ_q. Let n_q denote the number of enrollees who have not undergone hysterectomy at the beginning of quarter q. The inducement effect on hysterectomy due to disability insurance is $\sum_{q=-20}^{19} \hat{\rho}_q \cdot n_q$. If this measure is zero, we conclude that disability benefit has not increased the total number of hysterectomy among enrollees over their lifetime.

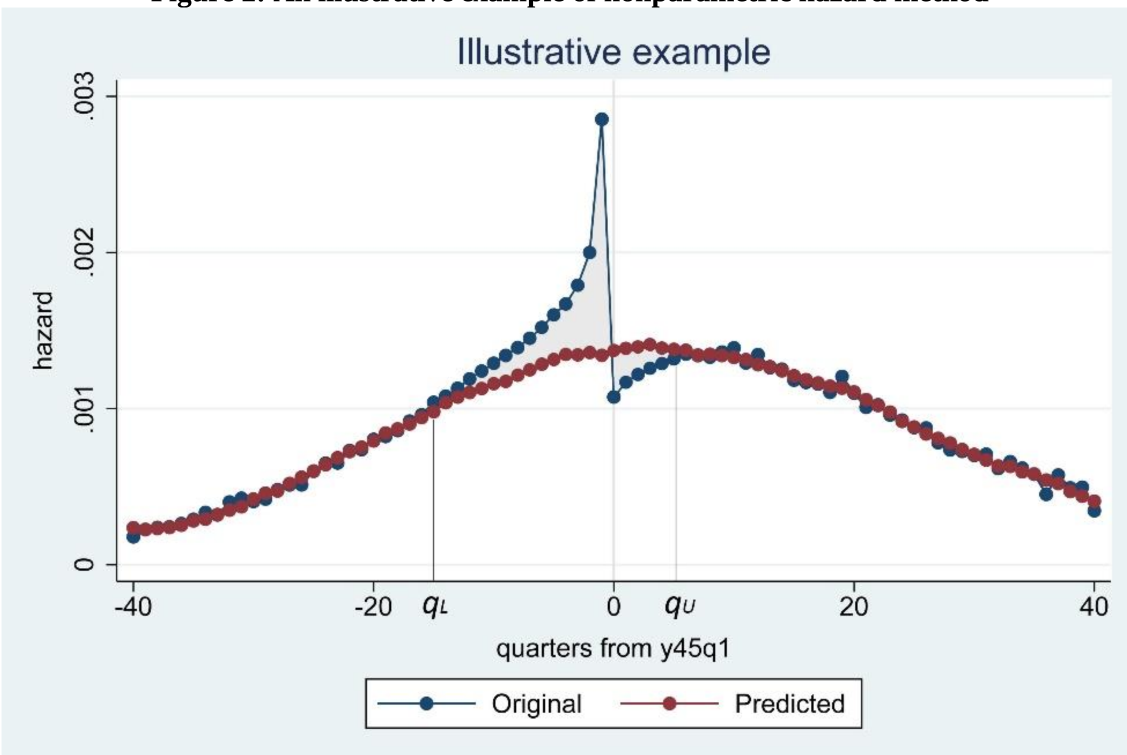
Next, the timing effect is the total number of hysterectomies that the insured would have undergone after the 45th birthday in the absence of the benefit. In other words, disability insurance may have incentivized enrollees to have hysterectomies earlier. As a result, there will be fewer of them after the 45th birthday among insured enrollees. The timing effect on hysterectomy due to disability insurance is $\sum_{q=0}^{19} \hat{\rho}_q \cdot n_q$. If disability insurance has not favored earlier hysterectomies, then this measure will be zero.

Finally, our analysis is at the birth cohort level, and the dependent variable is hysterectomy hazard of enrollees born at a certain time. We thus use the number of enrollees at every birth cohort as weights in the estimation. Given that all the covariates within a birth cohort are constant for each enrollee, estimates obtained from the individual-level regression would be identical to those from the cohort-level regression (Lee and Card, 2008; Lemieux and Milligan, 2008).

4.2 Nonparametric counterfactual estimation

The method above assumes that the hysterectomy hazards of insured groups before the 40th birthday share a common time trend, which is supported by the plots of hazards before age 40 in Figure 1. As an alternative, we use a nonparametric method that is free of any time-trend assumption. Instead, the nonparametric method is based on a "smoothness" assumption: without the expiration of disability benefits, there should not be any abrupt hysterectomy-hazard changes at age 45. We estimate a counterfactual hysterectomy hazard distribution for each insured group, and compare it to the actual hazard distribution.

Figure 2: An illustrative example of nonparametric hazard method



For the nonparametric estimation we have extended the data periods to 40 quarters before and after the 45th birthday, so these hazards are for ages between 35 and 55. We use Figure 2 to illustrate this method. There, the blue curve plots the empirical distribution of hysterectomy hazard of an insured group. It shows the sudden hazard changes around the 45th birthday. To construct a counterfactual distribution, we imagine that the abrupt changes had not existed. We choose a lower quarter threshold and an upper quarter threshold, which are denoted, respectively, by q_L and q_U, with $q_L < 0 < q_U$. We then use hazard data outside of quarters between q_L and q_U to fit an N^{th}-order polynomial. The fitted curve is then used to predict the hazards between quarters q_L and q_U. This is the red curve in Figure 2. The interpretation is that quarter q_L marks the beginning of disability insurance impact on hysterectomy before the 45th birthday, whereas quarter q_U marks the end of the impact after the 45th birthday.

For the nonparametric method, we use more observations outside the q_L and q_U thresholds for a better fit of the polynomial. We thus use a 20-year window of quarter ages, double the time window for the difference-in-difference estimation. To obtain more data points, we use the nonswitching sample employed in the difference-in-difference estimation, and then merge the sample with their enrollment and surgery records between 35 and 39 years old, and between 50 and 54 years old, and construct the corresponding hazards. Nonetheless, for the additional years of data, we only restrict to enrollees belonging to four insurance groups, relaxing the requirement of same insurance group over years

The regression for estimating the counterfactual hysterectomy hazard is this:

$$H_{c,q} = \sum_{n=0}^{N} \alpha_n \times q^n + \sum_{j=q_L}^{q_U} \rho_j \times 1[j = q] + \varepsilon_{c,q}$$

$$c \equiv (y, s) \text{ with } y = 1948, 1949, \ldots, 1972, \text{ and } s = 1, 2, 3, 4, \text{ and } q = -40, -39, -38, \ldots, -1, 0, 1, 2, \ldots, 39, \quad (2)$$

where $H_{c,q}$ is the hysterectomy hazard of birth cohort *c* at benefit quarter *q*; q^n is quarter *q* raised to the power *n* ; q_L and q_U, again, are the lower and upper bounds; and $\varepsilon_{c,q}$ is the error term. Notice that the birth cohorts still range between 1948 and 1972 given that we use the same cohorts in difference-in-difference estimation, but the quarter number now is from -40 to 39 because we incorporate more data points.

Following Kleven and Waseem (2013), we use a fifth order polynomial in the main specification ($N = 5$). For each quarter between q_L and q_U we use a coefficient ρ_j to capture the difference between the empirical and the counterfactual hazards at quarter *q*. If disability insurance has no effects on enrollees' hysterectomy decisions, all estimates of ρ_j should be zero.

For each insured groups, we use a grid search over the ranges of $q_L \in [-18, -9]$ and $q_U \in [2, 12]$ to select a pair of bounds that minimize the root mean squared error (RMSE) of the regression, an optimality criteria widely used in econometric models (Ichimura and Todd, 2007; Imbens and Kalyanaraman, 2012;

Lee and Lemieux, 2010).[9] Because each insured group sets its own lower and upper thresholds, the number of estimated coefficients in each insured group is different.

As in the difference-in-difference estimation, we are able to obtain inducement and timing effects. The inducement effect is the impact of disability insurance on enrollees' lifetime hysterectomy, and is measured by ($\sum_{q=q_L}^{q_U} \hat{\rho}_q \cdot n_q$), where $\hat{\rho}_q$ is the estimate of ρ_q in equation (2), and n_q is the number of enrollees who have not had hysterectomy at the beginning of quarter q. Likewise, the timing effect is the impact of disability insurance on the timing of hysterectomy, and is measured by $\sum_{q=0}^{q_U} \hat{\rho}_q \cdot n_q$. This is the difference in hysterectomy hazard between the counterfactual and the empirical hazard distributions after the 45th birthday but before q_U. More importantly, if our nonparametric estimation approach is valid, the estimates of ρ_j should all be zero for the uninsured group, and this could serve as the validity check in the analysis.

5. Estimation results

5.1. Difference-in-difference estimation

Table 2 presents some of the regression estimates of the difference-in-difference terms $\hat{\rho}_q$ in equation (1). We estimate the regression separately for each treatment group, so the three columns in Table 2 show results of Labor Insurance, Government Employee Insurance, and Farmer Insurance groups, respectively. In each column, the number of observation is 5280, smaller than one would expect from the sample of balanced data (400*11*2). This is because that quite a number of enrollees only appear in a few years in the data; censoring reduces the number of observations.

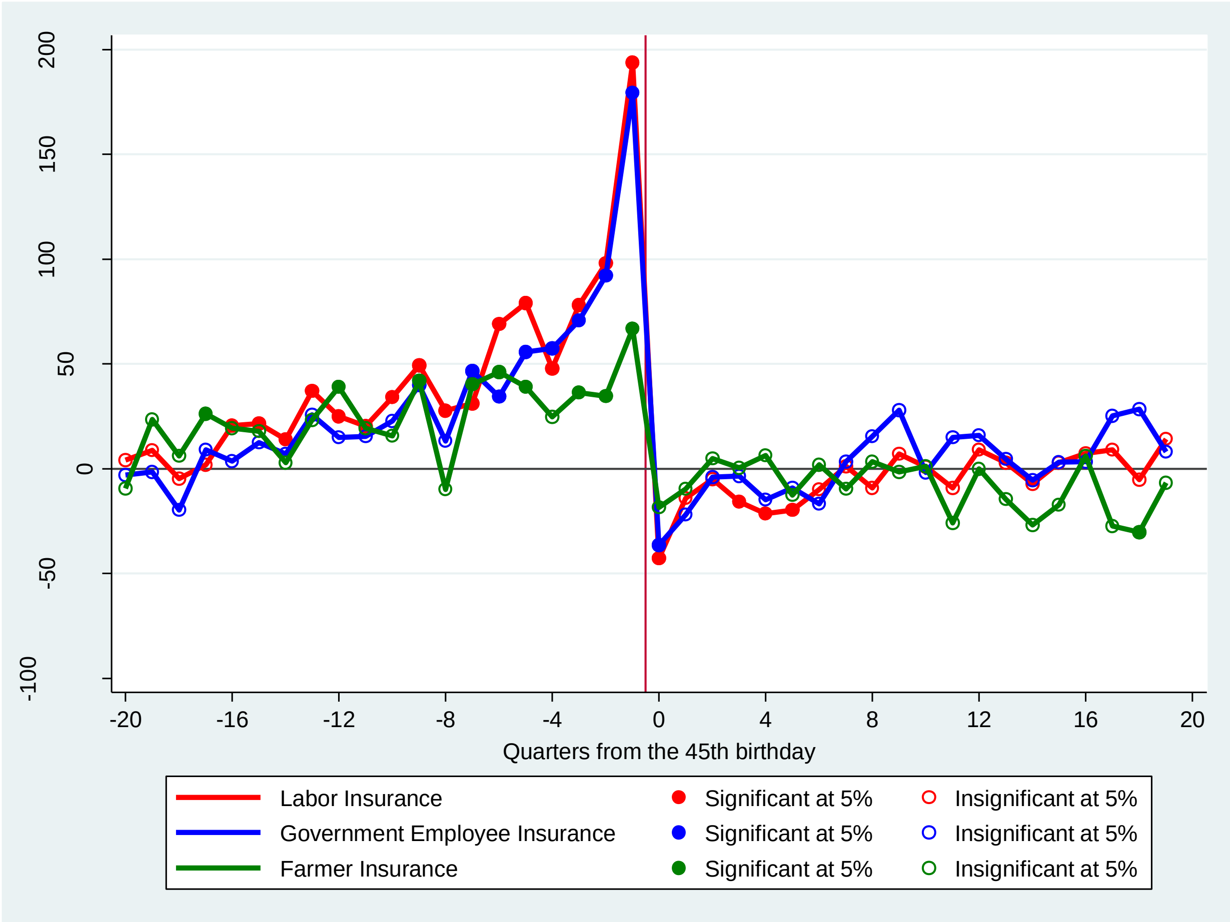
Our focus in Table 2 is the set of interaction terms for the quarterly hazard differences between the treatment group and the uninsured. These are the quarterly parameters $\hat{\rho}_q$, q between -20 and 19. For

[9] RMSE is a common measure for comparing the performance of different econometric models or parameter selection. For example, Ichimura (1993) proposes a semiparametric model and compares its RMSE to those of other models such as the truncated Tobit, binary choice, and duration models. RMSE is also the optimality criterion for selecting the smoothing parameter in nonparametric methods (Ichimura and Todd, 2007); for selecting the bandwidth for regression discontinuity designs (Imbens and Kalyanaraman, 2012); and for selecting the polynomial order of the regression function of regression discontinuity designs (Lee and Lemieux, 2010)

brevity, Table 2 presents only estimates of $\hat{\rho}_q$ for q between -10 and 7; these benefit quarters are around the 45th birthday, at which time the disability benefit expires. In fact, for quarters before -10 and after 7, most estimates of ρ_q are insignificant. However, for an effective illustration, Figure 3 plots the entire set of $\hat{\rho}_q$, $q = -20$ through $q = 19$. The red plots are for enrollees in Labor Insurance, the blue plots and the green plots are for enrollees in Government Employee Insurance and Farmer Insurance, respectively. In addition, significant estimates are plotted with solid dots, whereas insignificant estimates are plotted with hollow dots.

In Figure 3, for Labor Insurance enrollees $\hat{\rho}_q$ starts at almost zero at $q = -20$, gradually increases, and becomes significantly different from zero (solid dots) at $q = -14$. Then $\hat{\rho}_q$ continues to increase as enrollee's age approaches 45, peaks at $q = -1$ (the difference between the two groups is 193.8 cases per 100,000 at $q = -1$) and then sharply declines at $q = 0$. Except for $q = 0$, most estimates after $q = 0$ are small and insignificant. Likewise, the plot of Government Employee Insurance group follows a similar pattern. The plots of Farmer Insurance group also peak at one quarter before age 45, though the magnitude is only half of the other two insurance groups.

Figure 3: Difference-in-difference estimates of $\hat{\rho}_q$ for hysterectomy



In equation (1) the parameter β measures the average difference between the treatment and control groups. The treatment group dummy estimate is also presented in Table 2. The hysterectomy hazard for each treatment group is significantly different from the control group. For both Labor Insurance and Farmer Insurance enrollees, their hysterectomy hazard is higher than the insured, and this is stronger for Farmer Insurance enrollees than Labor Insurance enrollees. For Government Employee Insurance enrollees, this difference turns out to be negative. The identification power is not diminished by the sign differences in $\hat{\beta}$ because the treatment and control groups share the same time trend. This can be seen from the insignificant coefficients in the first few quarters in Figure 3.

Finally, equation (1) includes a number of controls. In all three equations, enrollees with higher household income are less likely to undertake hysterectomy. This is consistent with wealthier households being less responsive towards financial incentives. Enrollees with more children tend to have a smaller hysterectomy hazard, though this effect is insignificant for Government Employee Insurance enrollees.[10] Conditional on the total number of children, the number of sons does not seem to matter. Finally, being married is associated with a higher hazard, but the estimate is only significant for the Labor Insurance enrollees. By and large, estimated coefficients from controls are consistent with common models of health care services.

We now turn to the inducement and the timing effects. Recall that the inducement effect on hysterectomy is $\sum_{q=-20}^{19} \hat{\rho}_q \cdot n_q$, where n_q is the number of female enrollees who have not had hysterectomy at the beginning of quarter q. The inducement effect for enrollees in Labor Insurance is measured at 5,076 hysterectomies. This is about 11.6% of the total 43,845 hysterectomy surgeries undertaken by Labor Insurance enrollees between 40 and 49 years old in the sample period. Recall that the timing effect is measured as $\sum_{q=0}^{19} \hat{\rho}_q \cdot n_q$. As a matter of convention, we change the negative to positive for the ease of presentation. For Labor Insurance enrollees, the timing effect is measured at 1,008 hysterectomies, or at about 20% of the inducement effect. However, the value of $\hat{\rho}_q$ at $q = 0$ is significantly negative for the Labor Insurance enrollees. Although the overall timing effect is not large, we are confident that some hysterectomies have been shifted earlier due to the disability benefit.

In fact, we can visualize both inducement and timing effects from Figure 3. The inducement effect is the weighted sum of the areas under each plot, with weights n_q, for q between -20 and 19. Now the number of enrollees who have not had hysterectomy must be decreasing over time, so n_q is a decreasing sequence. In other words, the inducement effect is area under each plot weighted by this decreasing sequence. Given that the number of hysterectomy is small relative to the total number of survivors, the inducement effect would be slightly smaller than the area between the hazard curve and the x-axis. It is quite obvious that the total weighted area for q before 0 is much bigger than after. The timing effect can be visualized

[10] Studies have shown that the number of pregnancy (or living children) is negatively related to the prevalence of uterine fibroids, one of the major cause for hysterectomy (Ross, et. al, 1986; Chen et. al, 2000).

analogously. This is simply the weighted sum of the area for positive quarters q. It is also obvious that this number is negative, though much smaller than the inducement effect.

Columns (2) and (3) in Table 2 present the estimates of $\hat{\rho}_q$ for Government Employee Insurance and Farmer Insurance enrollees. These two columns exhibit the same pattern as Column (1): a large increase in hazard just before quarter 0, and then vanishing. In Government Employee Insurance, the inducement effect is measured at 789 hysterectomies, or about 11% of the total hysterectomy cases among Government Employee Insurance enrollees between 40 and 49 years old in the sample period while the timing effect is 142. For Farmer Insurance enrollees, the inducement effect is smaller, at 347 cases, or about 3.8% of all hysterectomy surgeries among Farmer Insurance enrollees between 40 and 49 years old in the sample period. The timing effect for Farmer Insurance is measured at 283 cases.

Regression results on hysterectomy hazards are strong evidence that enrollees respond to incentives created by the disability insurance program. The differences in inducement and timing effects in the three treatment groups are consistent with the differences in the three disability insurance programs. Benefits of Labor and Government Employee Insurance are higher than Farmer Insurance.

We now turn to regression results of the other three surgeries: total oophorectomy, partial oophorectomy, and myomectomy. Almost all estimates of regression results for equation (1) for these three surgeries are insignificant. For completeness, we present these in online Tables A1, A2, and A3 at http://goo.gl/z9dZbq (again, only estimates of $\hat{\rho}_q$ for q between -10 and 7). In Figures 4, 5, and 6 we plot the entire set of estimates of $\hat{\rho}_q$ for q between -20 and 19, and we use the same color convention for the three treatment groups. From these estimate plots, it is clear that the disability insurance program has not caused behavioral change.

For partial oophorectomy and myomectomy, these insignificant results are to be expected. These surgeries are not eligible for benefits. The insignificant result for total oophorectomy is actually an important finding. Total oophorectomy and hysterectomy carry the same benefits and qualification. However, it is well known that the health risks and long-term morbidity and harmful consequences of total oophorectomy are much severe than hysterectomy. Our results indicate that the benefits are not enough to change enrollees' behavior. Indeed, ours can be regarded as evidence that enrollees have taken seriously the cost and benefit of each surgery.

Figure 4: Difference-in-difference estimates of $\widehat{\rho}_q$ for oophorectomy

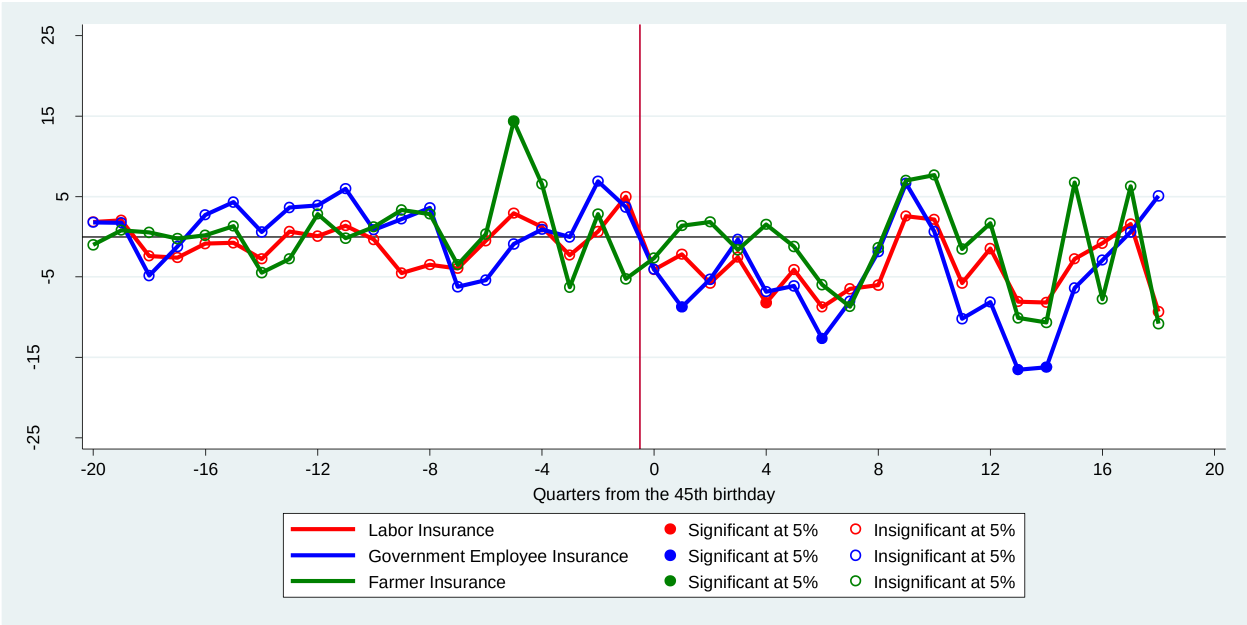


Figure 5: Difference-in-difference estimates of $\widehat{\rho}_q$ for partial oophorectomy

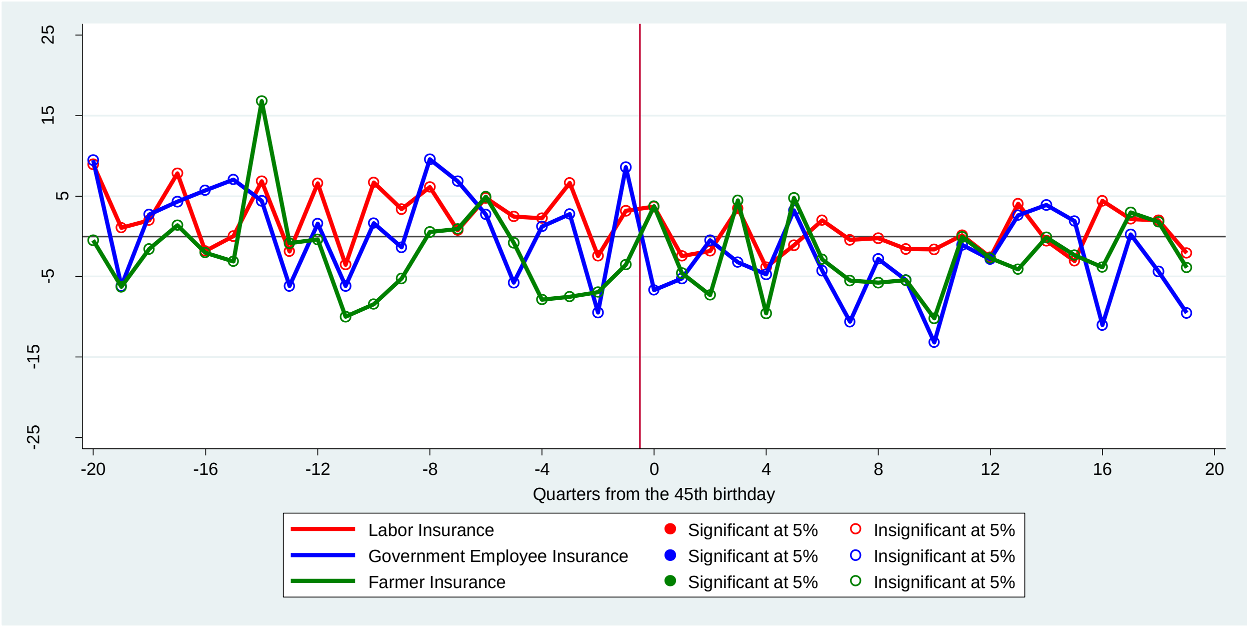
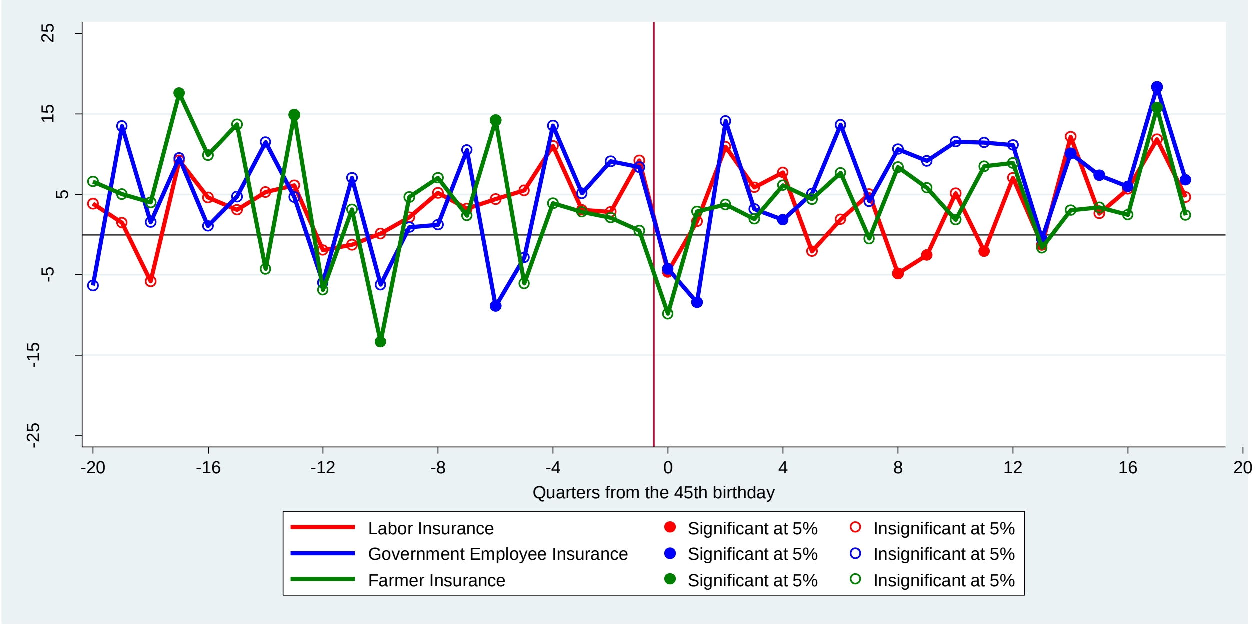


Figure 6: Difference-in-difference estimates of $\hat{\rho}_q$ for myomectomy



5.2 Nonparametric Estimation

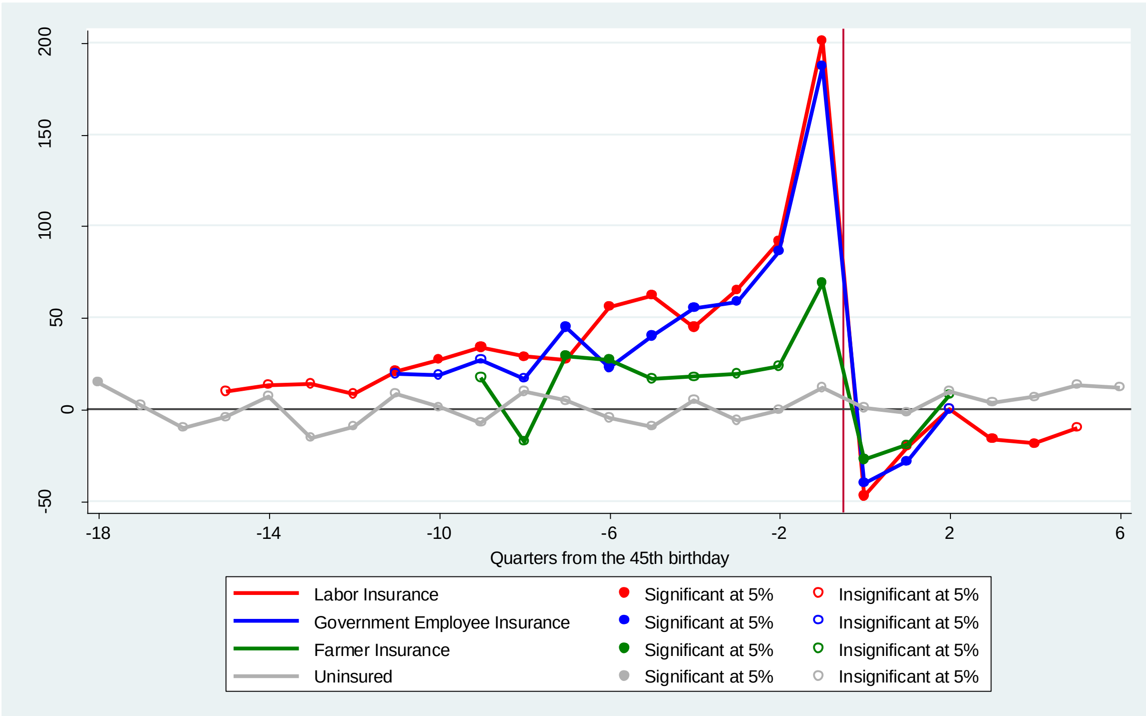
Table 3 presents the estimates $\hat{\rho}_q$ of equation (2). Because the nonparametric approach does not rely on the existence of a control, we have estimated all four insurance groups, including the uninsured one. The number of observation is 4498 for the Labor sample, and 4474 and 4481 for the Government Employee and Farmer sample. The observation number is smaller than that from the balanced data

(400*20) in part because of the censoring, but also because more missing observations when the data are extended to 20 years. As we have mentioned earlier, we apply a grid search over wide ranges of q_L and q_U to select a pair of bounds that minimize the root mean squared error (RMSE) of the regression. In Table 3, the four columns list the estimates of $\hat{\rho}_q$ for the uninsured, and enrollees in Labor Insurance, Government Employee Insurance, and Farmer Insurance, respectively, with the optimal bounds presented at the bottom of each column. Given that each insurance group has its own lower and upper bounds, the number of estimated coefficients in the table differs with respect to the insurance group. As in the previous subsection, we plot the four sets of $\hat{\rho}_q$ estimates, and they are in Figure 7. The gray plots refer to those of the uninsured. The red, blue, and green plots refer to those estimates of enrollees of Labor Insurance, Government Employee Insurance, and Farmer Insurance, respectively.

From Figure 7 the gray line fluctuates minimally along the horizontal axis line; this indicates that the 5th order polynomial fits quite well the uninsureds' hazard rates. In fact, in Table 3 almost all estimates of $\hat{\rho}_q$ of the uninsured are insignificant, and we cannot reject the hypothesis that estimates of $\hat{\rho}_q$ are jointly zero (F statistics = 1.01). This serves to validate our nonparametric approach. We now turn to the three insured groups.[11]

Figure 7: Nonparametric estimates of $\hat{\rho}_q$ for Hysterectomy

[11] We also presents the hazard distributions for three insurance groups (respectively presented in parts (B) to (D) of Figure A1) in online Appendix A at https://goo.gl/DDMYai. Each graph shows the counterfactual hazard distribution (dark gray curve) and the actual hazard distribution (curve in corresponding color).



For enrollees in Labor Insurance, most estimates from q = -11 to -1 are significantly positive, followed by significantly negative estimates from q = 0 to q = 4; see Table 3. The red plots in Figure 7 show the dramatic spike just before the 45th birthday, and then the sharp drop. The pattern is similar to the corresponding difference-in-difference estimates. The estimated number of induced hysterectomies, based on $\sum_{q=q_L}^{q_U} \hat{\rho}_q \cdot n_q$, is 4,842, which accounts for 11 percent of total hysterectomies (43,845) undertaken by Labor Insurance enrollees between 40 and 49 years old. Compared with the corresponding figures in the difference-in-difference method, the percentage is slightly smaller than the percentage (11.6 percent) estimated by the difference-in-difference method. The timing effect, measured by $\sum_{q=0}^{q_U} \hat{\rho}_q \cdot n_q$, is 722 hysterectomies, about 14.9% of the total inducement effect; it is somewhat lower than the corresponding percentage in the difference-in-difference estimates. Nevertheless, the nonparametric approach confirms the timing effect for Labor Insurance enrollees.

From Table 3, estimates of $\hat{\rho}_q$ for Government Employee Insurance enrollees are significantly positive for q between -7 and -1, but significantly negative at q = 0 and q = 1. The pattern can be seen in the blue plots in Figure 7, which also allows a visualization of the inducement and timing effects. We obtain the estimated inducement and timing effects at 756 and 87 hysterectomies, respectively. These estimates are relatively near to the corresponding difference-in-difference estimates (789 and 143 cases).

Finally, in Table 3, for enrollees in Labor Insurance, $\hat{\rho}_q$ is significantly positive at $q = -1$ and significantly negative at $q = 0$. In Figure 7, the green curve plots those estimates $\hat{\rho}_q$ for Farmer Insurance enrollees, suggesting that compared with Table 2, the estimates for Farmer Insurance group (the last column of Table 3) have fewer coefficients significantly different from zero before age 45. The estimated inducement effect is 280 hysterectomies, which is near the difference-in-difference value. However, the estimated timing effect is only 60 hysterectomies, much smaller than the difference-in-difference estimate (283).

So far the results of nonparametric estimation for hysterectomy are by the large similar to those obtained from difference-in-difference estimation. Results from both methods suggest that inducement and timing effects have larger impacts for Labor Insurances and the Government Employee Insurance enrollees, but less so for Farmer Insurance enrollees. However, comparing with results from difference-in-difference model, estimates of nonparametric method find a smaller timing effect in three insurance groups.

We now report results of nonparametric approach for the other three surgeries: total oophorectomy, partial oophorectomy, and myomectomy. For each surgery, we continue to use the uninsured group for our pre-examination to select the polynomial order and validate the nonparametric approach. For two surgeries: partial oophorectomy and myomectomy, the fifth order polynomial achieves a good fit, as most $\hat{\rho}_q$ are insignificant and the corresponding F-test (all coefficients) are insignificant for the uninsured group. For total oophorectomy, however, the polynomial order 5 fails the F test (F statistics = 11.22) and the polynomial order 6 fits the function better and passes the F test (F statistics = 0.92). So here we present the results from estimating the sixth order polynomial function. Again, for completeness, we present the estimates of $\hat{\rho}_q$ for q from optimal q_L to q_U in Tables A4, A5, and A6 (available online at http://goo.gl/xZ8i8N) and plot the entire set of estimates in Figures 8, 9, and 10.

From these three figures, it is obvious that the two estimation methods yield very similar findings. First, for total oophorectomy, which qualifies for insurance benefits, Figure 8 shows that very few $\hat{\rho}_q$ are significantly different from zero in all insurance groups. Again, hysterectomy and total oophorectomy entitle the enrollee to the same benefit, but total oophorectomy typically has worse health consequences than hysterectomy. The evidence is consistent with benefit-cost analysis. Second, almost all the plots in Figures 9 and 10 (for partial oophorectomy and myomectomy, respectively) are insignificant for every insurance group. In addition, the total inducement effect and the timing effect are minimal.

Figure 8: Nonparametric estimates of $\hat{\rho}_q$ for total oophorectomy

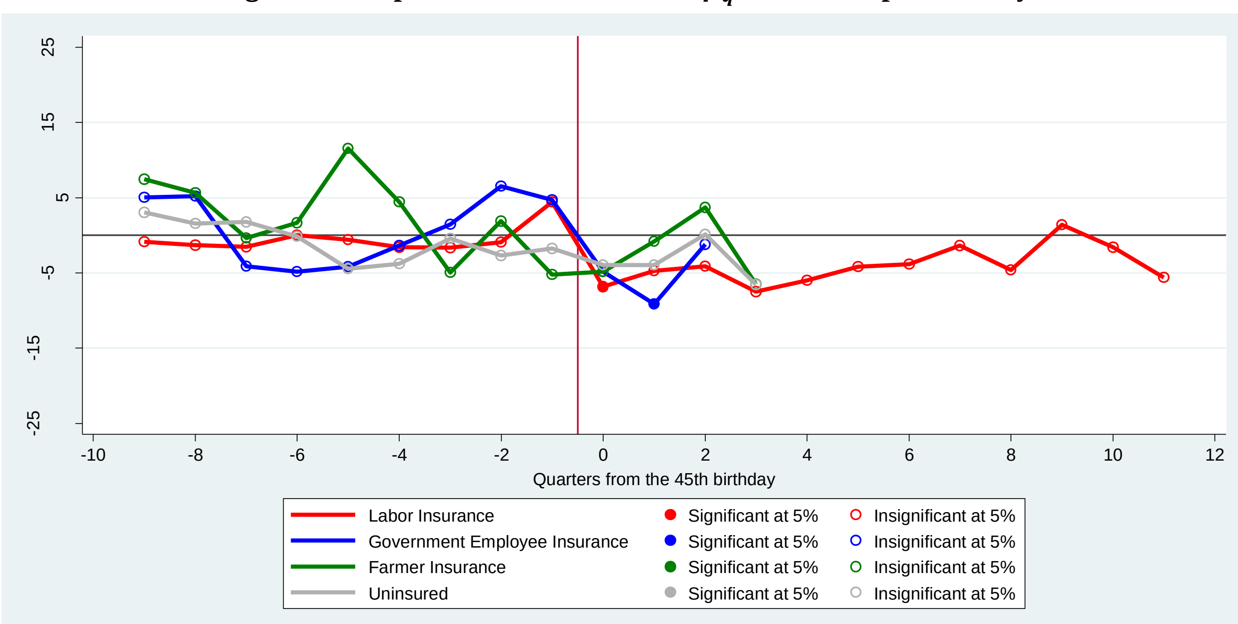


Figure 9: Nonparametric estimates of $\hat{\rho}_q$ for partial oophorectomy

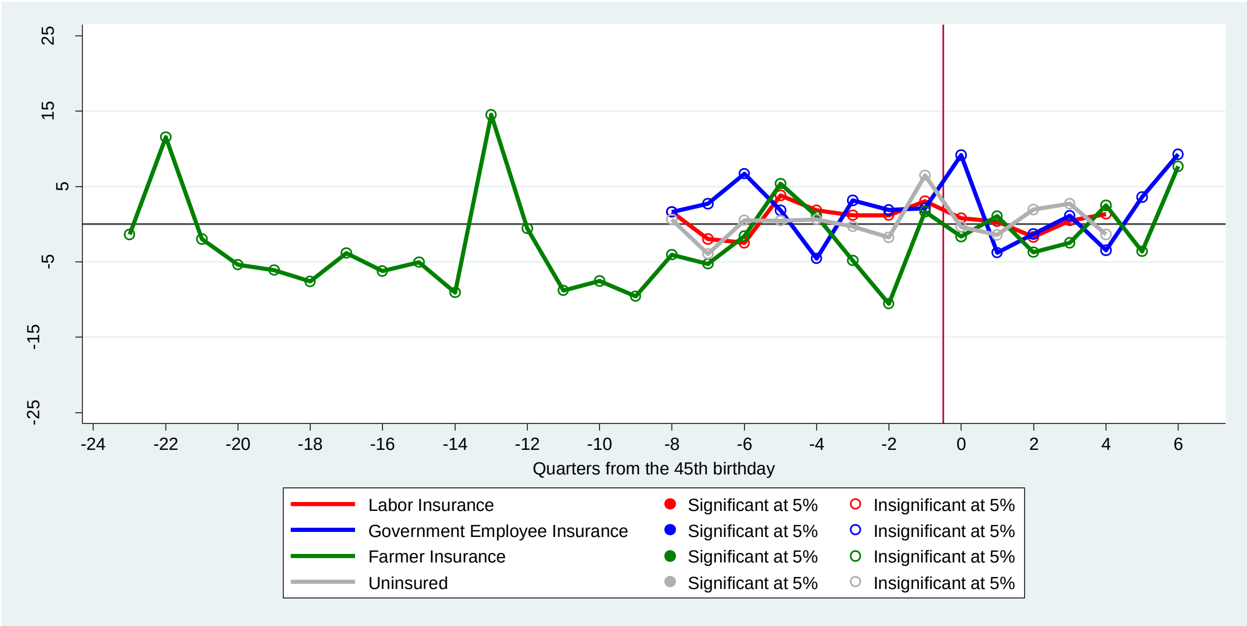
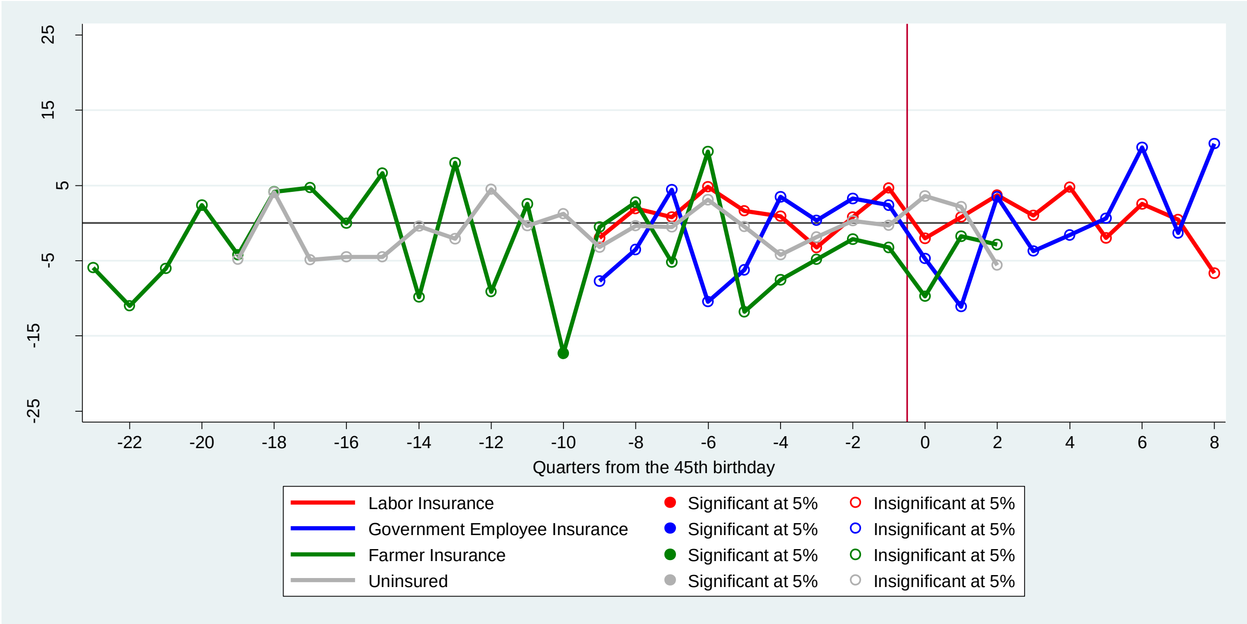


Figure 10: Nonparametric estimates of $\hat{\rho}_q$ for myomectomy



6. Robustness checks and heterogeneous benefit effects

In this section, we further investigate inducement and timing effects in samples that have been chosen by more or less restrictive criteria, and in subsamples determined by different insurance benefit levels.

6.1. Samples without nonswitching restriction and without censoring

In this subsection, we examine the robustness of our results by employing two different samples. First, we use a bigger, "general sample" consisting of all enrollees between the ages of 39 and 49, not only those who have not switched insurance programs. The general sample allows us to detect bias associated with the choice of insurance groups. Our data only allow us to identify an enrollee's insurance status at the end of a calendar year. We do not know the exact time or quarter of an insurance-status change. Therefore, for each individual, for each quarter within a calendar year, we use the end-of-year insurance status for all quarters in the year. Within each calendar year, insurance status of each enrollee remains constant by construction. Then, we calculate the hazard rates for the enrollees in each insurance group for each quarter in the year.

Second, we use a smaller, "balanced sample" consisting of all enrollees born between 1957 and 1962 (see Table B2). The balanced sample consists of those who had complete medical records between 39 and 49 in our sample of 1997-2011. By construction, data censoring does not happen.

To illustrate the differences among various samples, Table 4 lists the total numbers of observations by insurance groups in the general and balanced samples in 2010. For comparison, we also provide the corresponding numbers in the nonswitching sample. In Table 4, in 2010, there are about 1.3 million subjects in the nonswitching sample, but there are more than 1.6 million in the general sample. In other words, the general sample is about 25% larger than the nonswitching sample. By contrast, there are just over 0.25 million in the balanced sample, about 20% of the nonswitching sample.

Among the four insurance groups, the ratios of general sample size to nonswitching sample size is the lowest for Government Employee Insurance, at 1.04. Government employees appear to have higher job stability. By contrast, the corresponding ratios of Labor Insurance and the uninsured are higher, at 1.23 and 1.38, respectively. Enrollees switching in and out of being employed and being unemployed seem to be more common among those in Labor Insurance than those in either Government Employee Insurance or Farmer Insurance. The balanced sample consists of enrollees in the nonswitching sample born between 1957 and 1962, so naturally the corresponding ratios between sample sizes are stable, at about 20% of all insurance groups.[12]

Table 5 presents the summaries of difference-in-difference estimation results of the general and balanced samples for enrollees in Labor Insurance, Government Employee Insurance, and Farmer Insurance. For brevity, we present only total inducement and timing effects for hysterectomy;[13] the effects for all the other three surgeries are negligible.[14] Again, we include results of the nonswitching sample for easy comparison. From the first three rows in Table 5 for Labor Insurance, the inducement effect in the general sample is measured at 7,172 hysterectomies out of a total 61,692; this is about 11.6%. The inducement effect in the balanced sample is measured at 1,855 out of a total of 15,855; this is about 11.7%. The corresponding percentage for the nonswitching sample is 11.6% (5,076/43,845). Thus, the percentages of induced hysterectomy in total hysterectomy remain stable in these three samples. The estimated timing effects in the three samples of Labor Insurance are in column (4) of Table 5. For each timing effect, we also calculate it as a percentage of the inducement effect, and list all of them in column (5). The ratios of timing to inducement effects for the general and balanced samples are 16.8% and 20.1%, respectively; these compare with 19.9% of the nonswitching sample. Overall, the timing effects across the three samples are quite similar.

Estimates of inducement and timing effects of the general sample for Government Employee Insurance are also in Table 5. We do not include the results of the balanced samples for Government

[12] It is possible that an individual contributes to the numerator or denominator for the hazard rate of one insurance group, say, Government Employee Insurance, in one year, but will contribute to that of another insurance group, say, Labor Insurance, in the next year.

[13] We plot the difference-in-difference estimates of $\hat{\rho}_q$ of the general sample for hysterectomy, total oophorectomy, partial oophorectomy, and myomectomy, in Tables A7 to A10, respectively (available online at http://goo.gl/yUQHIh). Except for hysterectomy, most estimates are insignificant, and there are negligible inducement and timing effects.

[14] Tables A11 to A14 (available online at http://goo.gl/qGLTZK), respectively, present the difference-in-difference estimates of $\hat{\rho}_q$ for hysterectomy, total oophorectomy, partial oophorectomy, and myomectomy for the balanced sample. Again, the results are similar to those in Tables A7 to A10.

Employee Insurance and Farmer Insurance because the explanatory variables in these small samples exhibit very limited variations. For Government Insurance enrollees, for the general sample the total inducement effect measured as a percentage of total hysterectomy, and the timing effect measured as a percentage of inducement effect, are 11.2% and 17.6%, respectively, whereas the corresponding percentages in the nonswitching sample are 10.9% and 18.1%. These estimates confirm that results stable in the general and nonswitching samples are stable.

The last two rows in Table 5 are the summaries of the estimations of Farmer Insurance. The total inducement effect in the general sample is small, at 241 hysterectomies, or 2.2% of the total hysterectomy. Although the number is even smaller than the total inducement of 347 in the nonswitching sample, the results may be driven by the imprecise estimates of coefficients after age 45. For enrollees in Farmer Insurance, the inducement effect in the general sample is 2.2% of total hysterectomies, and lower than the 3.8% in the nonswithching sample. For the timing effect, the general sample is about 20 percentage points higher than the 81.6% (of the inducement effect) of the nonswitching sample.

Next, we turn to the nonparametric estimates. We only compare the results of the general and the nonswitching samples. Our nonparameteric method requires a 20-year sample window, but the data for 1997-2011 cannot track any individual for such a long time period, so we cannot construct a balanced sample.

Table 6 summarizes the inducement and timing effects from the nonparametric estimates.[15] For Labor Insurance, the inducement effect is measured at around 11% of total hysterectomy for both nonswitching and general samples. The timing effects are, respectively, 14.9% and 15.6% of the corresponding inducement effects for the general and nonswitching samples. These measures indicate robustness of estimates for the general and nonswitching samples.

For Government Employee Insurance, the inducement effects in the general sample and nonswitching sample are, respectively, 10.4% and 11% of the corresponding total hysterectomy. The

[15] We present the nonparametric estimates of $\hat{\rho}_q$, using the general sample, for hysterectomy, oophorectomy, partial oophorectomy, and myomectomy in Tables A15 to A18, respectively (available online at http://goo.gl/mP79xR).

timing effects in the general and nonswitching samples, as percentages of the corresponding inducement effects, are 11.5% and 15.6%, respectively. Again, these results indicate robustness. Likewise, for Farmer Insurance, the total inducement effects in the general and nonswitching samples are quite similar, measured at 3.1% and 2.9% of the corresponding hysterectomy, with the timing effect being 21.4% and 26.3%, respectively. Again, a smaller timing effect is obtained from the nonparametric estimation, especially for Farmer Insurance.

Results from the general sample allows us to calibrate the social cost associated with disability benefits. Our estimated inducement effects suggest that the provision of disability insurances increased 7,172 hysterectomies for enrollees under Labor Insurance, 885 cases under Government Employee Insurance, and 241 cases under Farmer Insurance. These inducement effects generated at least two kinds of social cost – extra medical expense and claimed disability benefits. Based on the reimbursed payment, the average medical expenditures for hysterectomy are NT$49,756, 50,406, and 50,212 for enrollees under Labor Insurance, Government Employee Insurance, and Farmer Insurance, respectively. Thus, the total medical expenditure caused by the disability insurances is NT$413 million. The average insurance benefits claimed by the enrollees under the three insurances are respectively NT$161,619, 193,452, and 106,684. These imply a total disability benefit of around NT$1.36 billion. Combined, the total social cost is NT$1.77 billion. Notice the current estimate does not include labor enrollees enrolled through trade unions. Therefore, the total social cost here is on the lower end of social cost.

6.2. Inducement and insurance benefit

We now investigate the relationship between benefit amount and the inducement and timing effects. This is an issue pertinent to current policy discussions because the Taiwanese government has been considering reducing fertility disability benefits. Our method is to stratify our sample into five groups of increasing insurance salaries with roughly equal numbers of observations in each group.

We carry out this stratification analysis on enrollees of Labor Insurance for two reasons. First, from Table 1, the sample size of Labor Insurance is at least 5 times larger than Government Employee Insurance, and 4 times larger than Farmer Insurance. The large sample size allows us to obtain more reliable estimates. More importantly, in contrast with other insurance groups, we have more accurate information from labor insurance since the disability-insurance salary is based on the enrollee's full monthly salary. Therefore, we can have a better understanding of the size of disability benefit (equal to 160 days or about 5.3 months of insurance salary, up to the cap of NT$43,900) relative to one's earning.

For each of the 5 groups, we use the difference-in-difference and nonparametric methods to estimate the number of induced hysterectomies and the inducement rate, the ratio of induced hysterectomies to total hysterectomies of enrollees between the ages of 40 and 49.

Table 7 presents the results, with the difference-in-difference and nonparametric method results in the upper and lower panels, respectively. Column 1 lists the average insurance benefits for the 5 groups. The average insurance benefit of group 1 is around NT$84,000, nearly one third of the average benefit of the highest group 5 which has an average of almost NT$220,000. Given that the maximum of disability benefit is approximately NT$232,600 (5.3×NT$43,900), the average insurance benefits of group 5, only a little lower than the maximum disability benefits, suggests that a sizable proportion of enrollees in this group have actual salaries above the cap.

Columns 2 and 3, respectively, present the total number of hysterectomies undertaken by enrollees between ages 40 and 49, and the estimated induced hysterectomies. Total hysterectomies of each group look similar, with the percentage difference between the highest and the lowest group being about 10%. By contrast, the variation of induced hysterectomies is quite large among different benefit groups: induced hysterectomies of group 5 are about twice that of group 1.

The estimated inducement rates are in Column 4. The inducement rate increases with average insurance benefit: the rate ratio of induced hysterectomy increases modestly when moving from group 1 to group 3 (from about 9% to11%), but accelerates when moving from group 3 to group 5 (from about 11% to 15%). In total, in the difference-in-difference estimates, the highest benefit group's inducement rate is about 75% larger than the lowest income group (15.71% versus 8.93%). The corresponding results in the nonparametric estimates are stronger, with group 5's inducement rate being more than twice that of group 1 (14.45% versus 7.11%).

Our stratification analysis shows an inducement effect that is strongly and positively affected by the benefit amount. Results in Table 7 shed some light on the possible impact of a policy change. The current discussion may recommend a reduction of benefit by a half of the highest benefit. If insurance programs were to cut the disability benefits by a half, for group 5 the average benefit would drop from 220k to 110k, falling between the average benefit of group 3 and group 2. A simple projection will therefore put the inducement effect at about 11%, which would be a reduction of more than 4.5 percentage points from the current inducement effect of 15.71%.

Inducement effects are stronger among those enrollees who have a higher insurance salary, and, therefore, are increasing in the benefits. If benefit is paid at a fixed price, say the benefit level being set at the third tier, we predict that inducement effects will be stronger among low-income enrollees, and weaker among high-income enrollees.

Our study here may shed some light on issues related to organ donations. Hysterectomy and oophorectomy do not involve donation or transplant. However, the disability insurance program does work like paying someone a price for having an organ removed. We do recognize that the uterus and ovaries are quite different from the kidney, blood, or the bone marrow. The rationale behind the repugnance constraint (Roth 2007) against financial rewards in organ donation is probably the worry that a price for organ donation adversely affects the poor more than the rich. Inducement effects being positively related to benefits supports this rationale. A fixed price for donation may be more attractive to poor donors than rich donors.

7. Conclusions

We have studied enrollees' response to the infertility coverage in three Taiwanese disability insurance programs. Enrollees having hysterectomy or complete oophorectomy qualify for benefit, but the eligibility at the 45th birthday. This program is probably one can be likened to a natural experiment of putting a price on the removal of a human organ. The results are striking. Compared to the uninsured, enrollees have about 11% more hysterectomy, and about 20% of the induced hysterectomies could be classified as those expedited to beat the deadline. By contract, the disability insurance has not led to any increase in oophorectomy.

Our results are striking but natural. The contrast between the different responses between hysterectomy and oophorectomy is striking. The plausible explanation behind the difference is naturally a cost-benefit calculus. Because organ removal is a discrete choice, economic principle dictates that such an operation is undertaken if and only if the reward is above a threshold. If the goal is an amount of insurance that would not result in induced operations, then our results indicate that in the Taiwanese case, the benefit is above the threshold for hysterectomy, but below for oophorectomy. A simple policy implication, therefore, is that insurance coverage for infertility should depend on whether the infertility is due to hysterectomy or oophorectomy.

The current study serves both as a warning and a validation. Qualifying a disability by a medical treatment may lead to double moral hazard: excessive applications for disability and medical treatments. However, where the benefit is not high relative to the disutility due to treatment, double moral hazard may be negligible.

References

Autor, David H., William R. Kerr, and Adriana D. Kugler. 2007. "Does Employment Protection Reduce Productivity? Evidence from U.S. States." *Economic Journal*, 117 (521): 189-217.

Aigner, Dennis, C. A. Knox Lovell, and Peter Schmidt. 1977. "Formulation and estimation of stochastic frontier production function models." *Journal of Econometrics*, 6 (1): 21-37.

American Congress of Obstetricians and Gynecologists. 2011. *Women's Health STATS & FACTS*, http://www.acog.org/-/media/NewsRoom/MediaKit.pdf.

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

Caliendo, Marco, Konstantinos Tatsiramos, and Arne Uhlendorff. 2013. "Benefit Duration, Unemployment Duration, and Job Match Quality: A Regression Discontinuity Approach." *Journal of Applied Econometrics*, 28 (4): 604-627.

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

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

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight. 2010. "Patient Cost-Sharing and Hospitalization Offsets in the Elderly." *American Economic Review*, 100 (1): 193-213.

Chen, Chao-Ru, Germaine M. Buck, Norman G. Courey, Kimberly M. Perez, and Jean Wactawski-Wende. 2000. "Risk Factors for Uterine Fibroids among Women Undergoing Tubal Sterilization." *American Journal of Epidemiology*, 153 (1): 20-26.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." *Quarterly Journal of Economics*, 126 (2): 749-804.

Diamond, Rebecca, and Petra Persson. 2016. "The Long-term Consequences of Teacher Discretion in Grading of High-stakes Tests." NBER Working Papers 22207.

Donald, Stephen G., and Kevin Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *Review of Economics and Statistics*, 89 (2): 221-233.

Department of Health and Human Services. 2011. "Hysterectomy Frequently Asked Questions." Office of Women's Health: Washington, DC.

Farber, Henry S., and Robert G. Valletta. 2015. "Do Extended Unemployment Benefits Lengthen Unemployment Spells? Evidence from Recent Cycles in the U.S. Labor Market." *Journal of Human Resources*, 50 (4): 873-909.

Finkelstein, Amy. 2007. "The Aggregate Effects of Health Insurance: Evidence from the Introduction of Medicare." *Quarterly Journal of Economics*, 122 (3): 1-37.

French, Eric, and Jae Song. 2014. "The Effect of Disability Insurance Receipt on Labor Supply." *American Economic Journal: Economic Policy,* 6 (2): 291-337.

Hsiao, Cheng, H. Steve Ching, and Shui Ki Wan. 2012. "A Panel Data Approach for Program Evaluation: Measuring the Benefits of Political and Economic Integration of Hong Kong with Mainland China." *Journal of Applied Econometrics*, 27(5): 705-740.

Hoynes, Hilary, Doug Miller, and David Simon. 2015. "Income, the Earned Income Tax Credit, and Infant Health." *American Economic Journal: Economic Policy*, 7(1): 172-211.

Ichimura, Hidehiko. 1993. "Semiparametric least squares (SLS) and weighted SLS estimation of single-index models." *Journal of Econometrics*, 58 (1-2): 71-120.

Ichimura, Hidehiko, and Petra E. Todd. 2007. "Implementing Semiparametric Estimators." *Handbook of Econometrics,* Volume 6B: 5369-5468.

Imbens, Guido M., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature,* 47 (1): 5-86.

Imbens, Guido M., and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*, 79 (3): 933-959.

Kleven, Henrik J., and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics*, 128 (2): 669-723.

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

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

Lien, Hsien-Ming. 2011. "How to Construct Social-Economic Variables from National Health Insurance Data." *Journal of Humanities and Social Sciences*, 23 (3): 371-398.

Lemieux, Thomas, and Kevin Milligan. 2008. "Incentive Effects of Social Assistance: A Regression Discontinuity Approach." *Journal of Econometrics*, 142 (2): 807-828.

Maestas, Nicole, Kathleen J. Mullen, and Alexander Strand. 2013. "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt." *American Economic Review*, 103 (5): 1797-1829.

McPherson, K., M. A. Metcalfe, A. Herbert, M. Maresh, A. Casbard, J. Hargreaves, S. Bridgman, and A. Clarke. 2004. "Severe Complications of Hysterectomy: the VALUE Study." *BJOG: An International Journal of Obstetrics & Gynaecology*, 111 (7): 688-694.

McPherson, Klim, Giorgia Gon, and Maggie Scott. 2013. "International Variations in Selected Number of Surgical Procedures." OECD Health Working Paper No. 61.

McWilliams, J. Michael, Alan M. Zaslavsky, Ellen Meara, and John Z. Ayanian. 2003. "Impact of Medicare Coverage on Basic Clinical Services for Previously Uninsured Adults." *Journal of the American Medical Association*, 290 (6): 757-764.

McWilliams, J. Michael, Ellen Meara, Alan M. Zaslavsky, and John Z. Ayanian. 2007. "Use of Health Services by Previously Uninsured Medicare Beneficiaries." *New England Journal of Medicine*, 357 (2): 143-153.

OECD Health Statistics. 2016. http://stats.oecd.org/index.aspx?DataSetCode=HEALTH_STAT.

Parker, William H., Michael S. Broder, Zhimei Liu, Donna Shoupe, Cindy Farquhar, and Jonathan S. Berek. 2005. "Ovarian Conservation at the Time of Hysterectomy for Benign Disease." *Obstetrics and Gynecology*, 106 (2): 219-226.

Persson, Petra. 2015. "Social Insurance and the Marriage Market." Institute for Evaluation of Labour Market and Education Policy (IFAU) Working Paper Series 2015:6.

Picard, Richard R., and R. Dennis Cook. 1984. "Cross-Validation of Regression Models." *Journal of the American Statistical Association*, 79 (387): 575-583.

Ross, Ron K., Malcolm C. Pike, Martin P. Vessey, Diana Bull, David Yeates, and John T. Casagrande. 1986. "Risk factors for Uterine Fibroids: Reduced Risk Associated with Oral Contraceptives." *British Medical Journal*, 293: 359-362.

Roth, Alvin. 2007. "Repugnance as a Constraint on Markets." *Journal of Economic Perspectives*, 21 (3): 37-58.

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

Schmieder, Johannes F., Till von Wachter, and Stefan Bender. 2016. "The Effect of Unemployment Benefits and Nonemployment Durations on Wages." *American Economic Review*, 106 (3): 739-777.

Staubli, Stefan. 2011. "The Impact of Stricter Criteria for Disability Insurance on Labor Force Participation." *Journal of Public Economics*, 95 (9-10): 1223-1235.

Wu, Ming-Ping, Kuan-Hui Huang, Cheng-Yu Long, Eing-Mei Tsai, and Chao-Hsiun Tang. 2010. "Trends in Various Types of Surgery for Hysterectomy and Distribution by Patient Age, Surgeon Age, and Hospital Accreditation: 10-year Population-Based Study in Taiwan." *Journal of Minimal Invasive Genecology*, 17 (5): 612-619.

Wu, Shu-Mei, Yu-Mei Chao Yu, Cheng-Fang Yang, and Hui-Lian Che. 2005. "Decision-Making Tree for Women Considering Hysterectomy." *Journal of Advanced Nursing*, 51 (4): 361-368.

Table 1: Summary Statistics of Women Aged Between 39-49 in the Sample*

	Nonswitching Sample			General Sample		
	2000	2005	2010	2000	2005	2010
Birth Year						
1950-54	37.7%	0.0%	0.0%	32.3%	0.0%	0.0%
1955-59	44.3%	33.0%	0.0%	46.9%	32.4%	0.0%
1960-64	17.9%	45.7%	29.9%	20.7%	47.2%	34.0%
1965-69	0.0%	21.4%	47.3%	0.0%	20.4%	46.4%
1970-74	0.0%	0.0%	22.8%	0.0%	0.0%	19.6%
Insurance Type						
Government Employee	9.6%	10.2%	9.6%	7.3%	7.3%	8.1%
Labor	47.8%	52.7%	56.5%	49.1%	51.3%	55.6%
Farmer	11.7%	9.0%	6.3%	9.7%	7.5%	6.2%
Uninsured	31.0%	28.2%	27.5%	33.9%	33.8%	30.1%
Surgery Incidence Rate (/100000)						
Hysterectomies	766.5	582.8	526.7	702.1	556.2	516.0
Myomectomies	124.2	198.5	273.1	109.9	173.6	234.1
Total oophorectomies	146.2	67.4	55.2	127.4	63.6	57.5
Partial oophorectomies	93.7	246.9	267.3	81.6	220.7	236.3
N	931,614	991,982	1,232,855	1348399	1477955	1523738

Standard errors are in parentheses.

*The number of obseravtions is calculated by the number of female enrollees who we
between 39 and 49 at the the end of that year.

Table 2: Difference-in-Difference Estimates of ρ_q for Hysterectomy (Nonswitching Sample)

	(1) Labor Insurance	(2) Government employee Insurance	(3) Farmer insurance
Quarter to 45th brithday x Treatmnt Dummy (ρ_i)			
-10	34.18**	22.65	15.70
	(7.950)	(12.06)	(15.01)
-9	49.37**	39.66**	41.90*
	(8.231)	(14.67)	(16.79)
-8	27.68**	13.19	-10.03
	(9.025)	(12.78)	(13.61)
-7	31.19**	46.67**	40.20*
	(7.053)	(16.19)	(15.74)
-6	69.09**	34.25**	46.10**
	(9.824)	(11.08)	(15.98)
-5	79.12**	55.79**	38.98*
	(8.240)	(15.62)	(15.46)
-4	47.85**	57.38**	24.64
	(8.316)	(16.96)	(15.52)
-3	78.11**	70.66**	36.32**
	(10.81)	(14.46)	(13.09)
-2	97.91**	92.00**	34.47*
	(11.05)	(15.91)	(15.24)
-1	193.8**	179.2**	66.70**
	(14.95)	(19.70)	(15.91)
0	-42.88**	-36.57**	-18.56
	(9.071)	(13.37)	(13.83)
1	-14.10	-21.98	-9.601
	(10.70)	(13.94)	(16.11)
2	-4.997	-4.015	4.841
	(8.992)	(12.62)	(16.31)
3	-15.78*	-3.447	0.480
	(7.703)	(13.86)	(15.61)
4	-21.51*	-14.79	6.320
	(8.839)	(14.65)	(14.86)
5	-19.64*	-9.188	-12.71
	(9.068)	(14.54)	(15.06)
6	-10.01	-16.81	1.950
	(10.21)	(12.54)	(15.08)
7	1.204	3.337	-9.745
	(7.572)	(15.31)	(14.17)
Treatment group dummy	8.764**	-17.23**	33.22**
	(3.289)	(4.228)	(5.825)
Log household income	-12.10*	-8.788*	-25.33**
	(5.502)	(4.044)	(5.940)
Total number of children	-20.34*	-11.82	-18.43*
	(8.306)	(8.496)	(7.395)
Number of sons	4.045	0.544	-1.099
	(10.98)	(10.57)	(10.32)
Married	48.33*	38.68	25.93
	(18.66)	(21.29)	(23.07)
Estimated inducement effect	5,076	789	347
Estimated timing effect	1,008	143	283
Observations	5,280	5,280	5,280

Notes: The dependent variable is quarterly hazard of hystereomy. Uninsured women are used as the control group in each regression. Other covariates are the full sets of quarter dummies and data year fixed-effects. Robust errors clustered at the birth cohort level are in parentheses. ** p<0.01, * p<0.05.

Table 3: Nonparametric Estimates of ρ_i for Hysterectomy (Nonswitching Sample)

Quarters to age 45	(1) Uninsured	(2) Labor insurance	(3) Government employee Insurance	(4) Farmer insurance
-18	14.61*			
	(6.95)			
-17	2.22			
	(7.32)			
-16	-10.15			
	(7.40)			
-15	-4.43	9.67		
	(7.73)	(7.33)		
-14	6.90	13.16		
	(7.40)	(7.00)		
-13	-15.66	13.75		
	(8.70)	(8.18)		
-12	-9.19	8.27		
	(8.50)	(8.14)		
-11	8.40	20.73*	19.20	
	(9.64)	(8.72)	(11.55)	
-10	1.27	27.13**	18.82	
	(8.82)	(9.36)	(13.07)	
-9	-7.15	33.64**	27.08	17.44
	(10.26)	(8.70)	(15.16)	(12.73)
-8	9.67	28.55**	16.82	-17.75
	(9.30)	(8.99)	(12.73)	(12.68)
-7	4.55	26.96**	44.99**	29.01*
	(7.95)	(8.24)	(16.03)	(12.63)
-6	-4.76	55.91**	22.92*	26.97*
	(9.29)	(10.25)	(10.78)	(12.59)
-5	-9.24	61.88**	39.84*	16.66
	(9.11)	(10.45)	(16.09)	(12.55)
-4	4.75	44.50**	55.15**	17.69
	(9.56)	(9.46)	(17.69)	(12.48)
-3	-6.23	65.06**	58.50**	19.13
	(10.70)	(10.77)	(15.43)	(12.41)
-2	-0.39	91.44**	86.19**	23.41
	(9.72)	(10.60)	(18.09)	(12.34)
-1	11.72	201.20**	187.31**	68.64**
	(10.25)	(17.53)	(22.06)	(12.29)
0	0.69	-47.24**	-40.39**	-27.49*
	(10.68)	(7.99)	(11.31)	(12.20)
1	-1.65	-20.55*	-28.25*	-19.43
	(10.44)	(9.11)	(12.27)	(12.12)
2	9.94	0.42	0.20	7.78
	(11.07)	(8.79)	(11.63)	(12.04)
3	3.71	-16.17*		
	(9.55)	(6.98)		
4	6.80	-18.57*		
	(10.86)	(8.66)		
5	13.15	-10.05		
	(10.45)	(8.11)		
6	11.91			
	(10.50)			
Optimal bounds (q_L, q_U)	(-18, 6)	(-15, 5)	(-11, 2)	(-9, 2)
Estimated inducement effect	--	4,842	756	280
Estimated timing effect	--	722	87	60
Observations	4,483	4,498	4,474	4,481

Notes: The dependent variable is quarterly hazard of hystereomy. The covariates are quarterly age with polynomial order

Table 4: Comparisons of Nonswitching, General, and Balanced Samples (2010)

	(1)	(2)	(3)	(4)	(5)
Insurance types	**Nonswitching sample**	**General sample**	**(2)/(1)**	**Balanced sample**	**(4)/(1)**
Government Employee	126,608	132,256	1.04	25,972	0.21
Labor	738,386	908,708	1.23	132,141	0.18
Farmer	85,917	104,808	1.22	23,691	0.28
Uninsured	361,614	498,157	1.38	68,664	0.19
	1,312,525	1,643,929	1.25	250,468	0.19

*The number of obserations is calculated by the number of enrollees at the end of that year.

Table 5: Comparisons of Difference-in-Difference Results for Hysterectomy using Various Samples

		(1)	(2)	(3)	(4)	(5)
Sample	**Insurance types**	Total hysterectomies undertaken from 40 to 49	Total Inducement Effect	(2)/(1)	The Timing Effect	(4)/(2)
Labor Insurance	Nonswitching Sample	43,845	5,076	**11.6%**	1,008	**19.9%**
	General Sample	61,692	7,172	**11.6%**	1,202	**16.8%**
	Balanced Sample	15,885	1,855	**11.7%**	373	**20.1%**
Government employee Insurance	Nonswitching Sample	7,262	789	**10.9%**	143	**18.1%**
	General Sample	7,888	885	**11.2%**	156	**17.6%**
Farmer insurance	Nonswitching Sample	9,100	347	**3.8%**	283	**81.6%**
	General Sample	10,987	241	**2.2%**	241	**100.0%**

Table 6: Nonparametric Method Results for Hysterectomy Based on Different Samples (Optimal Bounds)

Sample	Insurance types	(1) Total hysterectomies undertaken from 40 to 49	(2) Total Inducement Effect	(3) (2)/(1)	(4) The Timing Effect	(5) (4)/(2)
Labor Insurance	Nonswitching Sample	43,845	4,842	**11.0%**	722	**14.9%**
	General Sample	61,692	6,736	**10.9%**	1,053	**15.6%**
Government employee Insurance	Nonswitching Sample	7,262	756	**10.4%**	87	**11.5%**
	General Sample	7,888	864	**11.0%**	135	**15.6%**
Farmer Insurance	Nonswitching Sample	9,100	280	**3.1%**	60	**21.4%**
	General Sample	10,987	319	**2.9%**	84	**26.3%**

Table 7: Difference-in-difference results for hysterectomy (non-switching sample)

Subgroup	Average insurance salary (NT$)	Total hysterectomies undertaken from 40 to 49	Induced hysterectomies	Percent inducement (3)/(2)*100
Difference-in-Difference method				
1	84,592	6,158	739	12.0
2	100,425	6,873	1,060	15.4
3	133,435	6,485	949	14.6
4	185,197	6,811	1,319	19.4
5	219,500	6,488	1,356	20.9
Non-parametric approach				
1	84,592	6,158	589	9.6
2	100,425	6,873	903	13.1
3	133,435	6,485	848	13.1
4	185,197	6,811	1,520	22.3
5	219,500	6,488	1,247	19.2

Table 8: Difference-in-difference results for oophorectomy (non-switching sample)

Subgroup	Average insurance salary (NT$)	Total hysterectomies undertaken from 40 to 49	Induced hysterectomies	Percent inducement (3)/(2)*100
Difference-in-Difference method				
1	84,592	639	-84	-13.1
2	100,425	657	35	5.3
3	133,435	609	-54	-8.9
4	185,197	668	0	0.0
5	219,500	612	-58	-9.5
Non-parametric approach				
1	84,592	639	-59	-9.2
2	100,425	657	0	0.0
3	133,435	609	-80	-13.2
4	185,197	668	0	0.0
5	219,500	612	0	0.0

Table B1: Transition Matrix of Insurance Types for Women Having Hysterectomy in the General Sample						
	Insurance type at the time of having hysterectomy					
	(1)	(2)	(3)	(4)	(5)	(6)
Insurance type a year before having hysterectomy	Labor Insurance	Government Employee Insurance	Farmer Insurance	Trade Union	Uninsured	Total
Percent changes						
(1) Labor Insurance (%)	91.94	0.03	0.18	3.54	4.3	100
(2) Government Employee Ins. (%)	0.95	97.55	0.06	0.19	1.26	100
(3) Farmer Insurance (%)	1.66	0	96.54	0.79	1.01	100
(4) Trade union (%)	2.19	0.01	0.03	96.93	0.83	100
(5) Uninsured (%)	5.81	0.1	0.64	6.84	86.61	100
Total (%)	37.62	4.96	6.5	29.95	20.97	100
Frequency changes						
(1) Labor Insurance (obs.)	64,051	24	125	2,469	2,994	69,663
(2) Government Employee Ins. (obs.)	86	8,851	5	17	114	9,073
(3) Farmer Insurance (obs.)	194	0	11,294	93	118	11,699
(4) Trade union (obs.)	1,101	7	17	48,638	414	50,177
(5) Uninsured (obs.)	2,287	38	252	2,694	34,108	39,379
Total (obs.)	67,719	8,920	11,693	53,911	37,748	179,991

Table B2: The Construction of Nonswitching, General and Balanced Sample

Cohorts	Age at 1997	Age at 2011		years in sample	data record experience in incentive change
1948	49	63	left-censored	1	no
.	.	.	.	.	.
1952	45	59	left-censored	5	no
1953	44	58	left-censored	6	yes
.	.	.	.	.	.
1956	41	55	left-censored	9	yes
1957	40	54	left-censored	10	yes
1958	39	53	Balanced	11	yes
1959	38	52	Balanced	11	yes
1960	37	51	Balanced	11	yes
1961	36	50	Balanced	11	yes
1962	35	49	Balanced	11	yes
1963	34	48	right-censored	10	yes
1964	33	47	right-censored	9	yes
.	.	.	.	.	.
1966	31	45	right-censored	6	yes
1967	30	44	right-censored	5	no
.	.	.	.	.	.
1972	25	39	right-censored	1	no