

Does expanding health insurance coverage lead to higher birth rates?

The Effect of Health Insurance on Fertility

by Olivia Ondeck, Tasheen Rahimtoola and Amy Qiu

INTRODUCTION

This paper examines the relationship between health insurance and birth rates with the goal of answering the following question: What is the effect of health insurance on fertility?

This topic is interesting because family planning is an issue everyone faces at some point in his or her life. We want to know how important a factor health insurance is for people deciding how many children to have. *The Milbank Quarterly* states that between 2001 and 2009, US households faced an average increase in health care related financial burdens of 21.9 percent. They show convincingly that health care is becoming less affordable over time for almost everyone. Therefore it is valuable to study the implications of health insurance.¹

We are investigating the relationship between health insurance coverage and the probability of giving birth in a given year, both of which are binary variables. We will first examine the effect of health insurance coverage on the probability of giving birth using a simple regression analysis. We will then run multiple regressions controlling for more variables. First we will control for age and marital status and then in later regressions we will add economic variables including food stamps, public housing status,

¹ Dawson, Milly. "Cost of Health Care a Burden for Most U.S. Households". *Health Behavior News Service*. 13 March, 2014. Web.

hours worked and total family income. We will then control for education. Finally, we will include both year and state fixed effects.

In addition to our main regression model, we will analyze the effect of a specific policy implemented in Massachusetts in 2006, which is the Mandatory Health insurance law passed in 2006 in 48 states. We could not include Alaska and Hawaii, as the data was not available. This policy provides us with a natural experiment in using difference in difference regression analysis. We will examine the effect of the health insurance policy in Massachusetts.

LITERATURE

Before we discuss our model, we want to address previous research on our topic. Shin-Yi Chou, Michael Grossman, and Jin Tan Li published a paper “The National Health Insurance on Birth Outcomes: A natural experiment in Taiwan” in 2011². This paper discusses the impact of the introduction of National Health Insurance (NHI) in Taiwan in 1995, which covered all households. We see that the introduction of NHI led to reductions in birth rate for infants born in farm households but not for those born in private sector households. Farm families have lower levels of education and income than private sector families and premature and low-weight infants are in worse health than other infants. Thus, taken as a set, their findings suggest that health insurance improves infant health outcomes of population subgroups characterized by low levels of education, income, and health. For the former group, the mortality rate fell by 0.5 deaths per

² Chou, Shin Yi, Michael Grossman, Jin-Tan Liu. “The Impact of National Health Insurance on Birth Outcomes: A Natural Experiment in Taiwan.” *National Bureau of Economic Research*. (2011). Web.

thousand survivors or by 13 percent relative to the mean in the pre-NHI period of 4 deaths per thousand survivors. The percentage of the population with health insurance rose from approximately 54 percent in the month prior to implementation to approximately 92 percent. The Taiwanese study thus finds that better health insurance would understandably lead to a decrease in infant mortality, which would increase the number of children per household. In other words, it suggests a positive correlation between health insurance and number of children in a household (Chou, Shin Yi et. al, 2011).

In 1990, Arleen Leibowitz pushed a paper “The Response of Birth changes to Health Care Cost”.³ This paper examines whether an exogenous short-term change in the cost of medical care affects fertility in a cross section of women. Giving birth to a child leads both time and monetary costs, but this article focuses on monetary cost (especially in areas of medical care). Some women were randomly assigned to receive free medical care for three to five years and the results showed that these women had a birthrate increase of 29 percent compared to those who were assigned to the insurance plans that were not free. This response to changes in health insurance suggests that loss of insurance coverage during recessions may increase the effect of lower price in increasing the birth rates during economic downturns. The results confirms that birth rates respond to short term changes in the cost of inputs to children as well as race, age, marital status etc. For example: birth rates increase when there is a reduction in cost of one element of a good input such as health care (Leibowitz 1990).

However, a paper published in 2011, “Estimating the Fertility Effect of

³ Leibowitz, Arleen. “The Response of Births to Changes in Health Care Costs.” *The Journal of Human Resources*. 1990: 25(4): pp 197-711. Web.

Expansions of Publicly Funded Family Planning Services in California” (Foster, Briggs, et al, 2011).⁴ suggests the opposite result for the effect of health insurance coverage on fertility rates. The third paper found that in California, Medicaid’s expanded coverage of contraception and family planning under the family PACT program prevented an estimated 286,700 unintended pregnancies including 133,000 unintended births in 2007. This study suggests that expanding health care access, especially to poor women, could have a negative impact on the fertility rate by expanding access to contraception, which decreases the number of unplanned pregnancies and births.

ECONOMIC MODEL

In order to address our question, we used both simple and multiple regression models. We used panel data for all 50 states and the District of Columbia. The data covered the years 2000 to 2013. Since we are testing the effect of having health insurance coverage on the probability of giving birth in a given year, we used probit regression models. Furthermore, in 2006, Massachusetts passed a mandatory health insurance law, requiring that everyone be covered by health insurance. This policy provided us with a natural experiment in which we could test the effect of having health insurance on the probability of giving birth in a given year by comparing Massachusetts with another state that didn’t pass such a policy.

⁴ Foster, Diana G, M. Antonia Biggs, Daria Rostovtseva, Heike Thiel de Bocanegra, Phillip D. Darney, Claire D. Brindis. “Estimating the Fertility Effect of Expansions of Publicly Funded Family Planning Services in California.” *Women’s Health Issues Journal*. 21-6(2011): pp 418-424. Web.

First, we used a simple regression model and tested how much the probability of having a child would be expected to change for those who are covered by health insurance relative to those who are not covered by health insurance. The dependent variable specifies whether or not the person gave birth in the given state in the given year, and the variable for health insurance coverage indicates whether or not the person is covered by health insurance.

$$1. \text{ Birth}_i = B_0 + B_1 \text{healthinsurancecoverage}_i + E_i$$

Then we controlled for age since that we would expect age to have a significant effect on the probability of having a child. We predicted that age would have a negative effect on birth since older women are less fertile.

$$2. \text{ Birth}_i = B_0 + B_1 \text{healthinsurancecoverage}_i + B_2 \text{age}_i + E_i$$

After controlling for age, we controlled for four economic variables: food stamp coverage, whether or not the person was living in public housing, how many hours the individual worked last week, and total family income. We added them one at a time to assess each variable's relevance in the model. We expected that people with food stamps and public housing would be more likely to give birth since their costs of living are subsidized. We also expected that those with a higher family income would have a higher probability of giving birth since they

can afford more children. We were unsure about how hours worked last week would affect the probability of giving birth.

$$3. \text{ Birth}_i = B_0 + B_1 \text{healthinsurancecoverage}_i + B_2 \text{age}_i + B_3 \text{publichousing}_i + B_4 \text{foodstamps}_i + B_5 \text{totalfamilyincome}_i + B_6 \text{hoursworked}_i + E_i$$

Then we wanted to control for education. We used categorical variables where “high school” indicated that the individual only completed high school, “college degree” indicated a college degree (Bachelor’s) only, and “doctorate” indicated a doctorate degree. We omitted the group of individuals who did not complete high school. We expected that those who completed high school were more likely to give birth than those who didn’t, but we were unsure how higher education would affect birth, if at all.

$$4. \text{ Birth}_i = B_0 + B_1 \text{healthinsurance coverage}_i + B_2 \text{age}_i + B_3 \text{publichousing}_i + B_4 \text{foodstamps}_i + B_5 \text{totalfamilyincome}_i + B_6 \text{hoursworked}_i + B_7 \text{highschool}_i + B_8 \text{collegedegree}_i + B_9 \text{doctorate}_i + E_i$$

Then we wanted to finally control for marital status, for which we used three categorical variables: married, single, and widowed. Our joint omitted group included those who are divorced or separated:

$$5. \text{ Birth}_i = B_0 + B_1 \text{healthinsurance coverage}_i + B_2 \text{age}_i + B_3 \text{publichousing}_i + \\ B_4 \text{foodstamps}_i + B_5 \text{totalfamilyincome}_i + B_6 \text{hoursworked}_i + B_7 \text{highschool}_i + \\ B_8 \text{collegedegree}_i + B_9 \text{doctorate}_i + B_{10} \text{married}_i + B_{11} \text{single}_i + B_{12} \text{widow}_i + E_i$$

Since we are using panel data, we also added both year and state fixed effects. We added year fixed effects in order to account for changes that affected all the states, i.e. national changes. This is important because in 2010, President Barack Obama passed a national healthcare law called the Affordable Care Act. Furthermore, the recession in 2008 also affected all of the states financially and is an important factor to control for. State fixed effects allow us to account for unobserved heterogeneity. These are all of the differences between states, which don't change over time, like the state's general culture for example.

DATA AND DESCRIPTIVE STATISTICS

Our data all came from IPUMS CPS, which is a website operated by the University of Minnesota. IPUMS CPS data comes from the Current Population Survey, which is given each year in March. It is conducted by the US Census Bureau and the US Bureau of Labor Statistics, and so it contains labor force statistics.

Key Dependent Variable

We created our key dependent variable, whether or not an individual gave birth in a given year, using a variable describing the age of an individual's youngest child and filtered our data so that only the head of each household remained. This is

a dummy variable with a value of either 0 or 1, where those with a child younger than 1 in a given year (meaning they had given birth that year) had a value of 1 and those without a child younger than 1 had a value of 0. Since we have observations for many people in many states over about 10 years, we start out with 1,043,366 observations. The mean of this variable is 0.073, meaning that 7.3% of these 1,043,366 people had a child who was less than 1 years old any time from 2000 to 2013. (This is about 76, 165 people).

Independent Variables

Our key independent variable was also a dummy variable, indicating whether or not a given individual was covered by any type of health insurance. Therefore, the minimum value is 0 and the maximum value is 1. A value of 0 indicates no health insurance coverage and a value of 1 indicates health insurance coverage. The mean, 0.864, indicates that 86.4% of the observations have a value of 1, or are covered by health insurance. We also included categorical variables for marital status, which (as mentioned before) are “married”, “single” and “widow”, with those who are separated and divorced being the omitted group. According to the means of the marital status variables, 53.4% of those observed are married, 18.4% are single, and 9% are widows. (This means that 19.2% are divorced or separated). There are three continuous variables: age, hours worked last week, and total family income. The mean age is 48 years old, which makes sense because the data was filtered so that all the individuals in the dataset were the head of the household. The mean household income is \$62,439.98, with a standard deviation of \$68,455.16. The maximum income is over 2.5 million dollars, so it makes sense that the standard

deviation would be larger than the mean considering the large variation. The average number of hours worked last week is about 25, with a standard deviation of about 22, which also makes sense since the maximum number of hours worked in the last week is 99 hours.

EMPIRICAL RESULTS

To estimate the most accurate model possible, we first ran our probit regression models. A probit model estimates the probability of the dependent variable having a value of 1, and so in our case, these models estimate the additional change in probability of giving birth in a given year for those covered by health insurance relative to those not covered by health insurance. One issue with dummy dependent models is that there is a danger of heteroskedasticity, which means that the variance in the model is not constant. To correct for heteroskedasticity, we estimated robust standard errors in all of our regressions. Our regression results can be seen in Table 2.

Regression 1 is a simple regression testing the effect of having health insurance on the probability of giving birth in a given year. According to this regression, those with health insurance coverage are in expectation 2.4 percent less likely than those without health insurance coverage to give birth in a given year. A negative result could be plausible since we would expect those with health insurance to have more access to services such as contraceptive surgeries, but we suspect that omitted variables are causing negative bias on our health insurance coverage coefficient. Furthermore, our pseudo R^2 (the R^2 value for logit and probit

regression models) is only 0.0018, meaning that our model only explains 0.18 percent of the variation in Y. So we added age to our model, and sure enough, R^2 increased to 0.1531. Our coefficient on health insurance coverage also changed from negative to positive, meaning that the omission of age must have caused negative bias. According to the model, age has a negative effect on the probability of giving birth in a given year, which makes sense. Therefore, age must have a positive relationship with health insurance coverage, which also makes sense considering the fact that as people get older, they are probably more likely to go get health insurance coverage. (The older people get, the more chances they have to get sick).

In Regressions 3, 4, 5, and 6, we added each economic variable one by one in order to see the effect on our model. After we controlled for food stamps, and then hours worked last week in the next regressions, the coefficient on health insurance coverage increased to 0.010, meaning that in expectation those with health insurance are 1% more likely to give birth in a given year, and it is statistically significant. R^2 increased as well. After we controlled for total family income, the coefficient on health insurance coverage decreased by 0.003 and R^2 increased. The coefficient also remained statistically significant. However, after we controlled for whether the individual lives in public housing or not, R^2 sharply decreased and the health insurance coverage variable became negative again. The z value of the health insurance coverage coefficient also decreased, making the coefficient statistically insignificant. This suggests that public housing is an irrelevant variable since adding irrelevant variables causes the R^2 to decrease and standard errors to increase. As a result, we took public housing out of the subsequent models.

In Regression 7, we controlled for education. Our health insurance coefficient as expected became positive again, with a value of 0.007. We expected that those with only a high school degree would be less likely to give birth than those with a college or doctorate degree because it seems likely that those with higher levels of education would start families later. All the education variables were statistically significant, and as expected, the coefficient on high school was negative and the coefficients on those with college degrees and doctorate degrees were positive. Controlling for education increased R^2 from 0.1629 to 0.1631, meaning that adding it accounted for .02% more of the variation in Y. Since it didn't increase by that much, education probably isn't a super important factor in the probability of giving birth.

We then controlled for marital status in Regression 8 and R^2 increased from 0.1631 to 0.2097, meaning that controlling for marriage caused our model to control for four more percent of the variation in Y, and so marital status is clearly important in determining the probability of giving birth. Our health insurance coverage coefficient decreased from 0.007 to 0.002, or by about 71%. Omitting marital status from the model likely caused the health insurance coefficient to be positively biased, which makes sense considering that we would expect a positive relationship between being married and having health insurance as well as between being married and giving birth. Furthermore, according to this model, those who are married are 3.8% more likely to give birth than those who are separated or divorced. Also as expected, those who are single are about 1.8 percent less likely to give birth than those who are separated or divorced.

In regressions 9 and 10, we controlled for year and then both year and state fixed effects. Since fixed effects cannot be included when finding the marginal effects (coefficients) for a probit model, the coefficients have to be interpreted not as a marginal rate of change or difference in expected effect, but as a difference in expected z score for every 1 unit increase of a continuous variable or for a dummy variable having a value of 1 relative to a value of 0. As expected, including both year and state fixed effects increased R^2 since state effects control for unobserved time invariant heterogeneity and year fixed effects control for national changes (such as the 2008 recession or national health insurance policies).

After running our main regression models, we also did a difference in difference analysis to find the effect of the 2006 Mandatory Health Insurance Law in Massachusetts. This policy provided us with a natural experiment in which we could tease out the effect of health insurance coverage. An ideal difference in difference model cancels out the effect of being in Massachusetts relative to the comparison state as well as the effect of being in the post 2006 period relative to the pre 2006 period. Our difference in difference results can be seen in Tables 3-6. In order to do this analysis, we had to abide by the parallel trend assumption. This means that we assumed that in the absence of the health insurance policy in Massachusetts, Massachusetts and the other state would have been on the same trend in the probability of giving birth. In order to use this assumption, we chose states with close average rates of change in the probability of giving birth and health insurance coverage in the years 2000-2006 so that we could be more confident that we were finding the effect of health insurance on probability of giving birth. We compared

Massachusetts to three states: Minnesota, Vermont, and Washington DC. Then we compared Massachusetts to every other state as a whole unit. In order to do the analysis, we used the following regression:

$$\text{Birth}_{it} = B_1 \text{MA}_i + B_2 \text{Post2006}_t + B_3 (\text{MA}_i * \text{Post2006}_t) + E_{it}$$

The omitted state would be whichever state or entity Massachusetts was being compared to. The interaction term between Massachusetts and post 2006 was significant, with a p value of 0.001 when comparing it with Washington DC and with a p value of 0.1 when comparing it with Minnesota. The difference in difference model with Minnesota suggests that the effect of health insurance coverage on the probability of giving birth is -.010 and the difference in difference model with Washington DC suggests that the effect is -.019. The results of the difference in difference with Vermont and with the rest of the country as a whole were not statistically significant .

Our model has some clear sources of error. Our difference in difference results suggest the exact opposite effect of health insurance on birth probability that our first regression model had. This implies that our model could be positively biased, meaning that variables having a positive relationship with both health insurance coverage and the probability of giving birth (or a negative relationship with both) are omitted from the model. There is also the issue that we are using a probit model, which makes it very hard to interpret the coefficients if we want to add year and state fixed effects. Another type of nonlinear model for dummy

dependent variables, the logit model, would be a good alternative to try in further research, which is a similar probability model.

One way that the results of our national data might be biased is that we cannot distinguish between reverse causality. Perhaps individuals planning on having children seek health insurance if they formerly weren't covered. Other problems with the national data involve variables that we cannot control for, or control for poorly. Ideally we'd like to be able to track individual households and see if changing the healthcare conditions for them changes their propensity to have children. However, our data does not follow the same families over time, so we cannot control for time invariant heterogeneity at the individual household level. It would also be helpful to control for omitted variables that address time variant unobserved heterogeneity between households, such as personal feelings about children or maternal and paternal health. The issues of reverse causality and possible omitted variable bias could be causing a false positive.

As for our Massachusetts natural experiment, it is possible that our parallel trend does not hold. We assumed for our model that in the absence of the Massachusetts health care reform of 2006, that Massachusetts and the comparison states would be on the same trend for the probability of having a child. Although we chose states with the closest trends prior to 2006 in birth to Massachusetts, no state was exactly the same as Massachusetts, which leaves room for error.

CONCLUSION

Our overall findings about the effect of health insurance coverage on the probability of giving birth to a child were small. The difference in difference model

suggests that health insurance coverage decreases the probability of giving birth by about 1 percentage point, and this finding is not completely significant since our regression model found a positive result (but of about the same magnitude). To address this issue, there is much more research that needs to be done. Because our study currently only studies whether having health insurance increases the probability of having a child in the year of study, it would be effective to also study the long-term effect of having health insurance on the number of children per household, especially to see if and how that differs from the short term effects.

It would also be effective to determine how health care expansion affects children per household, whether it increases it by increasing the pregnancy rate, or by decreasing the child mortality rate (or both), and how this effect would counteract with the effect of more access to contraceptives.

Table 1. Summary Statistics

Variable	Observations	Mean	Std. Dev	Min	Max
Birth (DV)	1,043,366	0.073	0.260	0	1
Health Insurance Coverage (DV)	1,043,366	0.864	0.343	0	1
Age	1,043,366	48.268	16.303	15	90
Married	1,043,366	0.534	0.499	0	1
Single	1,043,366	0.184	0.388	0	1
Widow	1,043,366	0.090	0.286	0	1
Public Housing (DV)	333,204	0.102	0.303	0	1
Food stamps (DV)	1,043,366	0.081	0.272	0	1
Hours worked last week	1,043,366	25.342	22.104	0	99
Total family income	1,043,366	62,439.98	68,455.16	0	2,742,997
High school (DV)	1,043,366	0.297	0.457	0	1
College (DV)	1,043,366	0.183	0.387	0	1
Doctorate (DV)	1,043,366	0.014	0.118	0	1

Table 2: Regression Analysis Results

[illegible]

State fixed effects										X
Pseudo R²	.0018	.1531	.1582	.1588	.1629	.1298	.1631	.2097	.2098	.2110
Number of Observations	1,043,366	1,043,366	1,043,366	1,043,366	1,043,366	333,204	1,043,366	1,043,366	1,043,366	1,043,366

Note: *indicates statistical significance at the 10% level, ** indicates statistical significance at the 5% level, and *** indicates statistical significance at the 1% level
Standard errors are in parenthesis.

Table 3. Difference in Difference Table (Massachusetts v. Minnesota)

Massachusetts	Pre 2006 -.009 (.004)**	Post 2006 -.010 (.005)*
Minnesota	.0668 (.050)***	.002 (.004)

Table 4. Difference in Difference Table (Massachusetts v. Vermont)

Massachusetts	Pre 2006 .011 (.004)**	Post 2006 -.005 (.005)
Vermont	.0681 (.037)***	-.002 (.004)

Table 5. Difference in Difference Table (Massachusetts v. Washington DC)

Massachusetts	Pre 2006 .030 (.004)***	Post 2006 -.019 (.005)***
Washington DC	.0359 (.056)***	.014 (.004)***

Table 6. Difference in Difference Table (Massachusetts v. All States)

Massachusetts	Pre 2006 -.003 (.003)	Post 2006 -.003 (.004)
Every other state	.0618 (.019)***	-.005 (.001)***

Note: *indicates statistical significance at the 10% level, ** indicates statistical significance at the 5% level, and *** indicates statistical significance at the 1% level
Standard errors are in parenthesis. Bolded coefficient represents the interaction term between Massachusetts and Post 2006 time period.

Works Cited

Chou, Shin Yi, Michael Grossman, Jin-Tan Liu. "The Impact of National Health Insurance on Birth Outcomes: A Natural Experiment in Taiwan." *National Bureau of Economic Research*. (2011). Web.

Dawson, Milly. "Cost of Health Care a Burden for Most U.S. Households". *Health Behavior News Service*. 13 March, 2014. Web.

Foster, Diana G, M. Antonia Biggs, Daria Rostovtseva, Heike Thiel de Bocanegra, Phillip D. Darney, Claire D. Brindis. "Estimating the Fertility Effect of Expansions of Publicly Funded Family Planning Services in California." *Women's Health Issues Journal*. 21-6(2011): pp 418-424. Web.

Leibowitz, Arleen. "The Response of Births to Changes in Health Care Costs." *The Journal of Human Resources*. 1990: 25(4): pp 197-711. Web.

Miriam King, Steven Ruggles, J. Trent Alexander, Sarah Flood, Katie Genadek, Matthew B. Schroeder, Brandon Trampe, and Rebecca Vick. *Integrated Public Use Microdata Series, Current Population Survey: Version 3.0*. [Machine-readable database]. Minneapolis: University of Minnesota, 2010.

*Regards to Hannah Woodruff, Taylor Jones, and Toby Chaiken for their paper, which served as a guidance for the layout of this paper.