Effects of Social Factors and Insurance Plans on Healthcare Utilization

By: Will Macy

Overview

One of the most contested U.S policy debates of the last 20 years is the issue of healthcare administration, use, and cost. In this paper, the connection between health insurance and the demand for healthcare across various subgroups is explored. The subgroups included are age, gender, health status, and healthcare policy type. This paper serves to highlight the important implications for understanding how the properties of health insurance affects access for different samples and segments of America's population.

Health expenditures in the United States have soared over the last decade. Healthcare costs in the United States make up a significant portion of the federal budget every year. In an effort to control rising costs, President Barack Obama implemented the Affordable Care Act in 2010. This act sought to contain rising healthcare costs by expanding government oversight to regulate premiums and practices set forth by private health insurance companies, reform healthcare payment structures that lead to waste, and increase competition in the healthcare insurer and provider market. However, there is an ongoing debate as to what exactly are the most costly factors in the American healthcare system. An appropriate healthcare policy agenda would enable America to control spiraling costs and further economic damage. Government fine-tuning of payment structure could provide for an effective means for tackling the issue of rising healthcare costs.

A major report by Manning et al. (1987) shed insight into how cost sharing, or the structure of deductibles as part of a health insurance policy, affects a person's utilization of healthcare. They used data obtained through the RAND Health Insurance Experiment to analyze demand for various healthcare services. This report was instrumental in analyzing the effects of payment structures on healthcare utilization, and it also explored the effect of belonging to

several subgroups such as age, gender, and health status. In my report, the results of the RAND Health Insurance Experiment are examined in an effort to seek a greater understanding of how different health insurance plans affect healthcare expenditures.

The RAND Health Insurance Experiment was a national scale social experiment conducted by the RAND Corporation during the 1970's and 1980's. Since then, it has spawned an enormous amount of literature on a variety of topics in medicine, economics, and other social sciences. The result of the analysis work by Manning et al. are compared to results obtained through further regression modeling. My paper seeks to answer a primary driving question, which asks, "does an individual utilize health care less if his or her healthcare plan requires them to pay more costs out of their own pocket?"

In my report, both ordinary least squares (OLS) and probit estimations are used to examine the interaction between the type of health insurance an individual has and how much they utilize healthcare. Individuals in the sample population paid 25 percent, 50 percent, and 95 percent of their healthcare costs, or they had an individual deductible plan.

Individuals that are subject to five different healthcare plans are compared in how much they utilize healthcare in terms of quantity. Specific traits of the individuals in the sample are held constant to get a sense of how the healthcare plan specifically affects the level of healthcare utilization. Using probit estimations, the effects of healthcare plans on the likelihood of a person to utilize healthcare are measured, keeping constant factors such as healthcare status or if the individual worries about his or her health.

Overall, it was observed that if an individual must pay more costs out of pocket, he or she would utilize health insurance less. When similar traits are controlled for however, the differences in healthcare utilization becomes much less significant. This shows that there is an

effect of belonging to a particular subgroup when comparing people's utilization of healthcare insurance for medical services.

Theory and Regression Model Used

The model of this project was represented by the following regression equation (Equation 1), where *y* is the demand for healthcare, *deduct* is the healthcare policy type, *age* is the age of individual receiving healthcare, *gender* is the sex of the person receiving healthcare, and *hlthstat* is the health status of the individual receiving healthcare. The error term is denoted by *u*.

Equation 1: y= B0+B1*deduct +B2*age+B3*gender+B4*hlthstat+u

As a proxy for demand for healthcare, dollar amounts spent on healthcare and numbers of hospital visits were used. The main issue this tackles is that the rise or fall in demand for healthcare could be caused by various drivers that were associated with the nature of a person's healthcare plan in the way of deductible payment structure, and not the other way around.

Any omitted variable bias would skew the weight of the effect of any of the variables on the right side of the above equation. It seems that each of the independent variables has to some extent or another an effect on healthcare utilization and demand.

My paper also attempts to observe new trends in the RAND Health Insurance Experiment by examining different subgroups in the sample, such as different age breakdowns that may not have been thought of as significant or noteworthy before. Ordinary least squares regression was used to analyze continuous outcomes, and probit estimation was used to analyze binary outcomes.

Data and Results

For this project, ordinary least squares was used to examine data from the RAND Health Insurance Experiment. The RAND Health Insurance Experiment was conducted from the 1970's to the 1980's over a course of 15 years. In this experiment, individuals were assigned to one of 14 different fee-for-service insurance coverage groups, or deductible payment structures. Coverage varied by the amount of the deductible that the individual had to pay, and by the maximum dollar expenditure, after which the deductible was waived. Since the original study was completed, most of the analyses of the RAND Health Insurance Experiment data have grouped the plans into five categories: free care (no deductible), 25% co-pay, 50% co-pay, 95% co-pay, and the "individual deductible" plan, in which inpatient services were fully covered, and outpatient services were subject to a 95% co-pay. The dataset also gave information on other personal and biographical attributes, such as age, gender, socioeconomic status, and general health status. In addition, the RAND Experiment provided information on traits such as individual employment status, welfare status, health status, and an indication of how much the subject worries about his or her own health. Additionally, the RAND Health Insurance Experiment quantified the subject's utilization of healthcare through such variables as hospital or office visits and expenditures. Data was taken from families that lived in one of six geographic areas: Dayton, OH, Seattle, WA, Fitchburg, MA, Franklin County, MA, Charlestown, SC, and Georgetown County, SC. These sites were chosen to proxy for differences in city dynamics that define a broad healthcare delivery landscape, such as delivery wait times and physician per capita ratios.

Table 1 below was recreated from Manning et al. using regression analysis. There was a pattern observed that showed the relationship between healthcare plan and healthcare utilization.

The table below shows that those who pay fewer costs out of pocket utilize health insurance more.

Plan	Total Visits	Outpatient Expenditures (2011 $)	Total Admissions to Hospital	Inpatient Expenditures (2011 $)	Probability Medical Expenditures (%)	Probability Inpatient Expenditures (%)	Total Dollars Spent (2011 $)
Free	7.4070	452.4957	0.1282	778.0141	0.8620	0.1028	2081.83
25 Percent	4.8471	355.1723	0.1048	706.1574	0.7786	0.0839	1608.98
50 Percent	1.9446	301.4199	0.0896	838.3319	0.7612	0.0732	1568.51
95 Percent	2.9061	269.5151	0.0979	587.3661	0.6594	0.0782	1274.72
Individual Deductible	4.3692	315.5266	0.1155	717.4023	0.7030	0.0953	1603.07

Furthermore, the way that different types of individuals utilized healthcare within their assigned health insurance plans was also explored. By holding constant different segments of the sample, changes were observed that deviated from those seen in Table 1. Since the data was taken from a randomized experiment and not from natural observations, the direct impact of an individual's health insurance plan and individual traits on how often an individual utilizes healthcare could be better understood.

For this problem, the RAND Health Insurance Experiment data was examined in order to determine the effects that different degrees of health insurance coverage may have on healthcare utilization and expenditures. Table 1 above indicated that as the percent of co-pay increased, individuals were less likely to use healthcare services, and therefore had lower medical expenditures. However, it would be premature to make conclusions from these statistics. It was

important to control for the coefficients of healthcare utilization to ensure that the difference due to the insurance plans was adequately being analyzed. Table 2 contains the difference in healthcare utilization according to insurance plans when the age, education level, gender, AFDC (Aid to Families with Dependent Children) status, income level, and the individual's health status were being controlled for.

Plan	**Total Visits**	**Outpatient Expenditures (2011 $)**	**Total Admissions to Hospital**	**Inpatient Expenditures (2011 $)**	**Probability Medical Expenditures (%)**	**Probability Inpatient Expenditures (%)**	**Total Dollars Spent (2011 $)**
25 Percent	2.37	-99.32*	-0.0165	265.24	-7.82*	-2.61*	-254.88
50 Percent	-5.64	-334.90*	-0.0757*	115.32	-12.90*	-6.56*	-978.71*
95 Percent	-3.50	-282.14*	-0.0370*	-29.27	-18.18*	-3.37*	-984.39*
Individual Deductible	-2.47	-208.27*	-0.0422*	-2.74	-12.89*	-3.37*	-601.15*
Female	5.73*	233.96*	.07644*	270.69	14.85*	6.20*	681.60*

High School	-2.32	-0.13	-0.0350*	-165.68	2.45	-2.73*	52.37
Age 13-19	-2.35	-48.22	-0.0214	12.98	-2.25	-2.89	-288.23
Age 41-62	5.53*	256.28*	0.0364*	542.01*	8.08*	1.42	1030.66*
AFDC	10.30	60.46	0.0547	-229.55	-12.72*	5.30*	342.87
Health Status	-22.26*	-390.79*	-0.4311*	-7284.44*	-7.00	-20.43*	-7885.95*
Constant (Baseline)	26.59*	836.98*	0.5637*	7968.76	85.08*	32.12*	9916.95*

When additional covariates were controlled for, the difference in healthcare utilization according to the insurance plans was not as clear and evident as in Table 1. In fact, the difference in total number of visits and inpatient expenditures was not statistically significant at the 5% level. Nevertheless, the difference in outpatient expenditures, total admissions, probability of medical expenditures, probability of inpatient expenditures, and total dollars spent on healthcare confirmed that individuals use significantly less healthcare services when their co-pay increases.

Table 2 also shows some interesting new facts and trends. The average female used more healthcare services than the average male. However, when the number of admissions due to maternity and pregnancy were removed, the difference in admissions between male and female was not significant. It was presumed that if maternity and pregnancy-related medical expenditures were removed, then the difference would also become less significant for these healthcare utilization proxies as well.

Some of the results obtained, specifically in terms of age, were not surprising. It was observed that individuals between the ages of 40 and 62 seemed to use much more healthcare than other individuals, while the difference in healthcare utilization between teenagers and young adults was not significant.

There was no significant difference found in healthcare utilization according to education level. This was very surprising, as it was expected that individuals with low education levels would be more prone to diseases and unhealthy lifestyles, and thus use more healthcare.

Among the variables that were used for measuring healthcare utilization, the probability of medical expenditures and inpatient expenditures were also given. These variables were very important to the study because they measured the likelihood to use healthcare services. However, in Table 2, the effects of plans on these variables were estimated using ordinary least squares regression, and that caused discrepancies such as negative probability. In order to correct for these discrepancies, new estimates were computed using a probit model. The marginal effects are shown below in Table 3:

Plan	**Probability Medical Expenditures (%)**	**Probability Inpatient Expenditures (%)**
25 Percent	-10.75	-2.22*
50 Percent	-17.78	-5.64*
95 Percent	-22.78	-2.83*
Individual Deductible	-16.72	-2.88*

The results confirm the overall trend that was noticed when using the ordinary least squares regression. Interestingly, it was observed that the marginal effects were even greater when using the probit model.

Another interesting way to analyze the dataset was to estimate a regression of healthcare utilization coefficients on the log of medical expenditures. By doing this, individuals who did not have any medical expenditure were eliminated from the dataset, and thus the effects in terms of percent of dollars spent were better measured. The results are shown in the table below. The results were the same as the results in Table 2, in that they were not significant for inpatient expenditures, but significant for outpatient expenditures.

Plan	**Log (Outpatient Expenditures) (%)**	**Log (Inpatient Expenditures) (%)**
25 Percent	-16.94*	-32.97*
50 Percent	-60.90*	28.71
95 Percent	-43.38*	14.08
Individual Deductible	-25.54*	-3.22

The results in Table 4 were interesting, as they confirmed that the difference in expenditure was mostly for outpatient medical expenditure. This seemed to indicate that individuals in the experiment were less likely to use healthcare for minor heath issues when they had to share the costs.

Discussion

In order to analyze and explain the data, the RAND Health Insurance Experiment dataset needed to be broken down into smaller, more manageable pieces. However, this created multiple sources of bias.

First, variables of groups were created. These groups consisted of data that was given to in the form of scale-like surveys or dummy variables. However, in the process of converting this data into these binary dummy variables, detail that was associated with the original survey was lost or compromised.

Second, there was an area of contention when examining employment status of individuals in the survey. When looking at workers versus non-workers, it was difficult to distinguish between workers who were not able to work and workers who were simply unable to find a job at the time of the study. This posed a problem when analyzing how employment status affected a person's healthcare usage.

A third bias inherent within the RAND Health Insurance Experiment dataset involved the observer effect, in that the mere assignment of individuals into a specific category and healthcare plan may have affected individuals motives, preferences, and behavior.

A fourth bias exists in regard to difference in incomes between geographic areas in which the study was conducted. It appeared that in some of the sites there were many more individuals that were at low income levels than in some other areas. That may have been an indiciation of oversampling, and thus some of the estimations probably should have been weighted. Additionally, some of the particular geographic areas might have particular "healthcare cultures" in which certain services are more preferable than others.

The fifth and final bias was observed in terms of difference in doctor quality. The difference may have existed between geographic regions, or even between different hospitals in the same region. It is possible that a less experienced or qualified doctor ordered tests and examinations that may have not been required, or may have been inappropriate in a specific situation. As a result, healthcare costs would increase, and this may have been skewed due to the quality of healthcare actually being administered.

To sum up, it was observed that the less an individual pays, the more likely that individual would visit the doctor. However, the difference between likelihood of visits decreased between the groups when similar variables were controlled for. If people utilized healthcare more in the experiment, they may have been taking advantage of preventative care visits, and therefore may have had less costs in the long run. When considering the impact of number of visits and totla healthcare costs it was important to keep in mind that those who go to the doctor less may actually be more costly to the overall healthcare system because they might only be going for emergency situations when a medical issue gets severe. In this sense, one interesting study to be conducted in the future would involve comparing medical expenditures of individuals without insurance and individuals with healthcare insurance.

Do-File

```
clear
use RAND2011

gen allplan = 0
replace allplan = 1 if free==1
replace allplan = 2 if deduct25==1
replace allplan = 3 if deduct50==1
replace allplan = 4 if deduct95==1
replace allplan = 5 if deductind==1
label variable empstatus  "1=working, 2=strike, 3=laid off, 4=retired,
5=homemkr, 6=other"

log using table1, text replace
*replicating manning et al.
regress totvis deduct25 deduct50 deduct95 deductind
regress outpdol2011 deduct25 deduct50 deduct95 deductind
regress totadm deduct25 deduct50 deduct95 deductind
regress inpdol2011 deduct25 deduct50 deduct95 deductind
regress posmed deduct25 deduct50 deduct95 deductind
regress posinp deduct25 deduct50 deduct95 deductind
regress totdol2011 deduct25 deduct50 deduct95 deductind

log close

*replicating manning et al., but adding more covariates
*first generate dummies
gen female = 1 - male
gen highschool = 0
replace highschool = 1 if educyrs>=12
replace highschool =. if educyrs==.

gen collgrad = 0
replace collgrad = 1 if educyrs>=16
replace collgrad =. if educyrs==.

gen teen = 0
replace teen =1 if age>13 & age <=19
gen yadult = 0
replace yadult =1 if age>=20 & age<=40
gen preaarp = 0
replace preaarp=1 if age>40
gen totadm_pw = totadm - matadm - pregadm

gen subjhealth = 0
replace subjhealth = 1 if egfp < 4
replace subjhealth = . if egfp == .

log using table2, text replace
*new estimates
regress totvis deduct25 deduct50 deduct95 deductind female highschool teen
preaarp afdc subjhealth
```

```
regress outpdol2011 deduct25 deduct50 deduct95 deductind female highschool
teen preaarp afdc subjhealth
regress totadm deduct25 deduct50 deduct95 deductind female highschool teen
preaarp afdc subjhealth
regress inpdol2011 deduct25 deduct50 deduct95 deductind female highschool
teen preaarp afdc subjhealth
regress posmed deduct25 deduct50 deduct95 deductind female highschool teen
preaarp afdc subjhealth
regress posinp deduct25 deduct50 deduct95 deductind female highschool teen
preaarp afdc subjhealth
regress totdol2011 deduct25 deduct50 deduct95 deductind female highschool
teen preaarp afdc subjhealth
regress totadm_pw deduct25 deduct50 deduct95 deductind female highschool teen
preaarp afdc

log close

log using table3, text replace
*probit estimation of the likelihood that there is med expenditures
probit posinp deduct*
probit posinp deduct* female highschool teen preaarp afdc subjhealth
dprobit posinp deduct*
dprobit posinp deduct* female highschool teen preaarp afdc subjhealth
probit posmed deduct*
probit posmed deduct* female highschool teen preaarp afdc subjhealth
dprobit posmed deduct*
dprobit posmed deduct* female highschool teen preaarp afdc subjhealth

log close

log using table4, text replace
*re-estimating models using log
gen loginpdol2011 = log(inpdol2011)
gen logoutpdol2011 = log(outpdol2011)
regress logoutpdol2011 deduct25 deduct50 deduct95 deductind female highschool
teen preaarp afdc subjhealth
regress loginpdol2011 deduct25 deduct50 deduct95 deductind female highschool
teen preaarp afdc subjhealth

log close
```

References

Manning, W.G., Newhouse, J.P., Duan, N., Keeler, E.B., & Leibowitz, A. (1987). Health Insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review, 77(3), 251-277.*

RAND Health Insurance Experiment. (n.d.). Retrieved from http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/6439/system.