Consumer Responsiveness of Health Care Utilization to Insurance Coverage: An Empirical Analysis Using Multilevel Regression

Shengting Yan

12/12/2018

Abstract

This paper attempts to quantify consumers' responsiveness to changes in insurance coverage through their utilization of medical services, including physician visits, prescription drugs, preventive care, surgeries, and inpatient and outpatient services. The famous RAND Health Insurance Experiment in 1970s designed a randomized trial to study how different insurance plan coverages can affect people's behavior. Researchers observed reduced utilization of care attributable to higher cost-sharing rate, which meant higher cost burden on the patients. While that study proved such a relationship, this study seeks to quantify and estimate the magnitude of this effect. By employing multilevel regression model on a subset of the 2016 MEPS data, the study finds that 1% increase in insurance coverage is associated with 6.1 times rise in health expenditures. While being old, female, white, highly-educated, and in poor health condition would lead people to take advantage of the medical services more frequently, minorities, employed, and those with private insurance use less care. Given the continually rising share of U.S. national health spending of the total GDP, understanding the factors that affect service use can direct policy-makers to design a more cost-efficient medical system to ultimately control cost without undermining the supply and the quality of care.

Introduction

Health spending has always been a debated topic in the United States. In recent decades, U.S. national health spending has more than doubled from $1.4 trillion in 2000 to $3.3 trillion in 2016, which constituted 18% of GDP (Kamal & Cox, 2017). In 2017, the average amount spent on healthcare per person in comparable countries ($5,198) was half that of the U.S. ($10,348). Researchers and the government have been seeking effective methods to control the cost without jeopardizing the supply and the quality of care.

Total health expenditure represents the amount spent on healthcare and health-related activities, such as administration of insurance and health research. It is a function of both price paid to providers or for drugs and the volume of services used (Kamal & Cox). This paper in particular focuses on the utilization of services in the private sector and seeks to identify factors that can affect the use. Generally, insurance is found to increase the intensity of utilization and reduce out-of-pocket spending (Ekman, 2007). Brook et al. (2006) find that increasing cost sharing rate reduces the use of health services, manifested by less doctor visits, including dental visits, prescriptions, and mental health treatments, and fewer hospitalization. Their results are drawn from the RAND Health Insurance Experiment conducted between 1971 and 1982, where participants were randomly assigned to four kinds of cost-sharing insurance plans with 0% (free), 25%, 50%, and 95% coinsurance rates. While that study employed randomized trial experiment design and analyzed the change in consumer behavior attributable to changes in coinsurance rate, this paper attempts to quantify this relationship and estimates the magnitude of the effect of insurance coverage on health care utilization.

Methods

Data and Variables

The Medical Expenditure Panel Survey (MEPS), which began in 1996, is a set of large-scale surveys of families, their medical providers, and employers across the United States sponsored by the Agency for Healthcare Research (Healthcare Research, 2009). It has two major components: the Household Component and the Insurance Component. This study uses the former from 2016, which collects data from sampled families drawn from a nationally representative subsample of households that participated in the prior year's National Health Interview Survey and whose data files are publicly available. MEPS collects detailed information on demographics, health history, insurance, medical care use and expenses from each household member during household interviews. All data for a sampled household are reported by a single household respondent. In 2016, 34655 persons from 13587 families were interviewed.

The study sample is restricted to persons who are between 18 and 64 years of age in 2016. Observations with missing responses and those uninsured are removed. The dependent variable is log transformed total health-related expenditures for each individual. I use it as a measure for utilization of medical services. The variable of interest is coverage, calculated by dividing the total amount self-paid by total expenditures. This variable aims to approximate the cost-sharing percentage for each insured individual. Other selected covariates are: gender, age, race, education, employment status, income, insurance type (public and private), self-rated health and mental health status on a scale of 1 (excellent) to 5 (poor), and the number of prior diagnosis of illnesses including high blood pressure, coronary heart disease, angina, myocardial infarction, other unspecified heart disease, stroke, emphysema, high cholesterol, cancer, diabetes, joint pain, arthritis, and asthma. More detailed information is available in the codebook.

Statistical Analysis

Given the nested nature of the data, which means that individuals are clustered within households, I use a multilevel regression model. Multilevel model attributes variations in the outcome variable to variations between individuals (first-level) and those between households (second-level) (Gelman & Hill, 2017). I assume a random intercept model since the study focuses on how people respond to insurance coverage in general rather than how different context characteristics (i.e. household) affect such responsiveness.

I performed statistical analyses using RStudio version 1.1.456. I constructed a full multilevel model using all covariates and then conducted model selection employing both forward stepwise and backward elimination methods based on improvement in the Akaike information criterion. To test whether using a multilevel model is necessary, the final model was compared to a multiple linear regression model with the same sets of covariates. For statistical inference, I conducted 2-tailed tests with $p \leq 0.05$ considered to be significant.

Results

Characteristics of the analytic sample by gender are shown in Table 1. In total, there are 10660 individuals from 7821 households. The number of individuals within a same household is between 1 (50.02%) and 5 (0.04%). There are 4652 (43.64%) males and 6008 (56.36%) females, with mean age as 44 ($\pm$12.8) and 43 ($\pm$12.8). 50.23% are white. 33.71% have a higher education degree. 23.54% self-reported an excellent health status. 79.45% are employed and have an average income level of \$52009.47 ($\pm$\$42714.98). The mean expenditures for male is \$5339.34 ($\pm$\$14016.42) and for female is \$6915.73 ($\pm$\$17420.72). The mean coverage for the sample is 0.72 ($\pm$0.28).

Table 2 presents the results from a multilevel regression and a multiple linear regression. These are the final models given that no variable has been removed from the model selection process. Note that the coefficients have been exponentiated for easier interpretation. The multilevel model shows that consumers are very responsive to insurance coverage in terms of their utilization pattern. 1% increase in insurance coverage would lead to 6.06 (95%CI, 5.42-6.77) times increase in total health expenditures after adjusted for other covariates in the model. Characteristics that are associated with higher utilization of medical services include being old (1; 95%CI, 1-1.01), being a female (1.45; 95%CI, 1.38-1.53), being white (1.31; 95%CI, 1.22-1.41), being highly-educated (1.14; 95%CI, 1.05-1.23), having poor health status (2.32; 95%CI, 1.95-2.76), and having more previously diagnosed diseases (1.24; 95%CI, 1.22-1.27). Being a minority is predicted to have less use of medical care. Employed population (0.74; 95%CI, 0.69-0.8) reduce their use of care as well as those with private insurance plans (0.89; 95%CI, 0.82-0.96).

A multiple linear regression model was used to test the hypothesis that whether a multilevel model is necessary for this sample. As one can notice, the estimated coefficients are very similar. An anova test shows a p-value of 1.6×10^{-5}, implying that despite the little difference in estimated values, a multilevel model is indeed more suitable given the clustered nature of the data.

Discussion

The direction of correlation for certain covariates are consistent with the findings documented by other researchers. For example, Yang, Norton, and Stearns contend that health expenses for elderly people increase substantially with age due to increasing use of inpatient service and long-term care (2003). Bertakis, Azari, Helms, Callahan, and Robbins conclude that women have higher medical care service utilization, such as primary care visits and diagnostic services, and higher associated charges than men (2000). Minorities are predicted to use fewer medical service possibly due to limited access to care or fewer employee benefits given a higher level of unemployment. The possibility of bearing potentially exorbitant medical bills reduces their tendency to use care. Interestingly, highly-educated people utilize more health service. While Cutler & Lleras-Muney find a positive relationship between education and health: "an additional four years of education lowers five-year mortality by 1.8 percentage points", their opportunity cost of being sick is likely to be higher and thus they spend more time and effort to stay healthy by visiting doctors more (2006).

One limitation of the study lies in the outcome variable. Provided that I use medical expenditures to approximate utilization and compare them for individuals from a national sample, I neglect the influence of prices across different regions. Future studies should use price-adjusted measurements of utilization. Moreover, considering that a single respondent completed the survey for all members in a household, input error is unavoidable. This could explain why coefficients for mental status variables are not significant. Since people value their privacy and mental health can be a taboo even between family members, having a person report mental status for other people is certainly susceptible to falsifications of reality.

On the implication side, while the model possesses predictive ability for medical care use, the outcome variable certainly does not suggest anything about individual's potential health outcome. The audience should take care when interpretating the coefficients and their implications. Particularly, a higher utilization does not necessarily imply an overuse of resources or an existence of moral hazard. For example, an increase in preventive care should not be discouraged as it hedges people from future catastrophic spendings. Learning about what kind of services each person uses would be extremely useful in deciding the channels through which insurance coverage affects people's consumption pattern of medical care.

Conclusion

The basic economic theory asserts that people have unlimited wants but are subject to limited resources. A budget constraint leads people to be very responsive to their insurance coverage, which specifies the percentage they are responsible for a medical service. Such a responsiveness is manifested through their consumption pattern of health-related services. By analyzing 2016 MEPS Household Component data with multilevel regression model, this study attains a quantitative measure of the maginitude of which people's medical care service utilization is influenced by their insurance plans' cost-sharing rates. Other characteristics of the insured, including gender, race, education, and age, are also found to be significant in affecting the use. In the context of a growing U.S. national health spending and a goal to restrain total cost and maintain a quality provision, unraveling the factors that contribute to the growth, whether due to rises in price or in demand for services, merits continued attention.

References

Bertakis, K. D., Azari, R., Helms, L. J., Callahan, E. J., & Robbins, J. A. (2000, February). Gender Differences in the Utilization of Health Care Services. Journal of Family Practice, 49(2), 147. Retrieved from https://link.galegroup.com/apps/doc/A60039859/AONE?u=bron88970&sid=AONE&xid=c32d0ace

Brook, R. H., Keeler, E. B., Lohr, K. N., Newhouse, J. P., Ware, J. E., Rogers W. H.,... Reboussin, B. (2006). The Health Insurance Experiment: A Classic RAND Study Speaks to the Current Health Care Reform Debate. Santa Monica, CA: RAND Corporation.

Cutler, D.M., & Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence. National Bureau of Economic Research, working paper 12532. doi: 10.3386/w12352

Ekman, B. (2007). The impact of health insurance on outpatient utilization and expenditure: Evidence from one middle-income country using national household survey data. Health Research Policy and Systems,5(1), 5-6. doi:10.1186/1478-4505-5-6

Gelman, A., & Hill, J. (2017). Data analysis using regression and multilevel/hierarchical models. Cambridge: Cambridge University Press.

Healthcare Research. (2009, August 21). Survey Background. Retrieved December 9, 2018, from https://meps.ahrq.gov/mepsweb/about_meps/survey_back.jsp

Kamal, R., & Cox, C. (2017, December 20). How has U.S. spending on healthcare changed over time? Retrieved December 9, 2018, from https://www.healthsystemtracker.org/chart-collection/u-s-spending-healthcare-changed-time/#item-per-capita-basis-health-spending-grown-substantially_2017

Yang, Z., Norton, C. E., & Stearns, C. S. (2003, January 1). Longevity and Health Care Expenditures: The Real Reasons Older People Spend More. The Journals of Gerontology: Series B, 58(1), S2–S10. doi: 10.1093/geronb/58.1.s2

Table 1: Characteristics of the Sample by Gender	Male (N=4652)	Female (N=6008)
Age		
mean (sd)	43.97 ± 12.84	43.30 ± 12.86
Race		
Hispanic	995 (21)	1,247 (21)
While Only	2,431 (52)	2,924 (49)
Black Only	678 (15)	1,157 (19)
Asian Only	413 (9)	462 (8)
Education		
High School Degree	2,201 (47)	2,681 (45)
Other degree	418 (9)	648 (11)
Higher-education Degree	1,522 (33)	2,071 (34)
Employment		
Unemployed	796 (17)	1,395 (23)
Employed	3,856 (83)	4,613 (77)
Income		
mean (sd)	53,001.60 ± 45,725.44	38,454.68 ± 37,127.16
Insurance Type		
Public Only	3,979 (86)	4,754 (79)
Private Only	673 (14)	1,254 (21)
Coverage		
mean (sd)	0.69 ± 0.30	0.74 ± 0.27
Health Status		
Excellent	1,132 (24)	1,377 (23)
Very Good	1,568 (34)	1,947 (32)
Good	1,290 (28)	1,714 (29)
Fair	507 (11)	735 (12)
Poor	155 (3)	235 (4)
Prior Diagnosis		
mean (sd)	1.77 ± 1.79	1.72 ± 1.84
Expenditure		
mean (sd)	5,339.34 ± 14,016.42	6,915.73 ± 17,420.72

Table 2: Regression Results from a Multilevel Model and a Multiple Linear Model

	Dependent variable:	
	Log(Total Expenditures)	
	linear mixed-effects	*normal*
Intercept	127.79*** (0.09)	127.03*** (0.09)
Coverage	7.06*** (0.05)	7.05*** (0.05)
Age	1.00*** (0.001)	1.00*** (0.001)
Female	1.45*** (0.03)	1.45*** (0.03)
White Only	1.31*** (0.04)	1.31*** (0.04)
Black Only	0.88*** (0.04)	0.88*** (0.04)
Asian Only	0.90*** (0.06)	0.90*** (0.06)
Multiple Race	1.17*** (0.08)	1.17*** (0.08)
High School Degree	1.17*** (0.05)	1.17*** (0.05)
Other degree	1.32*** (0.06)	1.33*** (0.06)
Higher-education Degree	1.57*** (0.05)	1.58*** (0.05)
Health: Very Good	1.14*** (0.04)	1.14*** (0.04)
Health: Good	1.30*** (0.04)	1.29*** (0.04)
Health: Fair	1.81*** (0.06)	1.79*** (0.06)
Health: Poor	2.32*** (0.09)	2.29*** (0.09)
Mental: Very Good	1.01*** (0.04)	1.01*** (0.04)
Mental: Good	0.96*** (0.04)	0.97*** (0.04)
Mental: Fair	1.07*** (0.06)	1.07*** (0.06)
Mental: Poor	1.38*** (0.13)	1.38*** (0.13)
Employed	0.74*** (0.04)	0.74*** (0.04)
Income	1.00*** (0.0000)	1.00*** (0.0000)
Private Insurance	0.89*** (0.04)	0.89*** (0.04)
Prior Diagonosis	1.24*** (0.01)	1.24*** (0.01)
Observations	10,660	10,660
Akaike Inf. Crit.	36,958.85	36,837.76
Bayesian Inf. Crit.	37,140.70	
Note:		*p<0.1; **p<0.05; ***p<0.01