Hitotsubashi Journal of Economics 51 (2010), pp.129-147. © Hitotsubashi University

DE-MYSTIFYING THE INCONVENIENT TRUTH: DOES EX POST MORAL HAZARD INDEED EXIST IN KOREAN PRIVATE HEALTH INSURANCE MARKET?

Eun-Hae Shin

Department of Food and Resource Economics, Korea University
Seoul 136-701, Korea
seh0710@korea.ac.kr

AND

Jae-Young Lim*

Department of Food and Resource Economics, Korea University
Seoul 136-701, Korea
jylimecon@korea.ac.kr

Received February 2010; Accepted August 2010

Abstract

There have been heated debates on whether private health insurance creates moral hazard effects. Despite its importance, however, the moral hazard problem of private health insurance is still controversial and understudied. To empirically examine whether or not moral hazard exists in the Korean private health insurance market, we employed two-stage regression for endogeneity control and the Heckman two-step procedure for sample selection bias control, which are expected to produce consistent estimates. All estimation results do not allow us to detect the presence of the moral hazard effects and imply that people hold private health insurance simply as "safety net".

Keywords: private health insurance; moral hazard; endogeneity; sample selection bias
JEL Classification: I10, I11

I. *Introduction*

The problem of moral hazard in health care and health insurance has long been discussed in economic literature. There are two ways in defining moral hazard in insurance markets: 1) *ex ante* and 2) *ex post* moral hazard. The first type indicates the risky behavior itself. For example, those who have automobile insurance may tend to be less careful of driving, thereby increasing

* Corresponding Author

the risk of an accident. In the same manner, after purchasing fire insurance, people may tend to less care about preventing fires. On the other hand, moral hazard falling into the second category does not result in more risky behaviors, but rather in more costly consequences. For instance, after purchasing medical insurance, health care services available become cheaper for individuals, and therefore people may consume more services than they would otherwise, all else being equal (Pauly, 1968; Zweifel and Manning, 2000; Anne, 2009). This paper adopts the second definition of moral hazard to see whether people face incentives to consume more medical services than the optimal level since they do not have to pay the full marginal costs by purchasing health insurance.

Does this *ex post* moral hazard problem exist in Korean insurance market? To answer this question, one may have to understand the unique health care system of Korea. Korea introduced mandatory social health insurance for industrial workers in large corporations in 1977, and extended it incrementally to the self-employed until it covered the entire population in 1989. In addition to the universal health care system, private health insurance in Korea plays a role as supplemental insurance; people voluntarily purchase private health insurance if they feel that public-mandated insurance is not sufficient or if they want to reduce potential burdens further which can be caused by unexpected diseases or accidents in the future.

This paper specifically aims to detect the problem of moral hazard in Korean *private* health insurance market. We focus on this issue since there have been few studies dealing with the supplemental insurance, and thus it can allow us to re-evaluate the role of private health insurance and to design policies that can take advantage of it in accordance with public insurance. In Section Ⅱ, we first discuss previous studies conducted with the same issues and find our own strategies for the following analysis. We then perform empirical analysis with several econometric tools in order to investigate the effects of having private health insurance on medical care use by effectively controlling other confounding factors in Section III, and the empirical results are introduced in Section IV. Section V concludes.

II. *Literature Review*

The presence of health insurance and the demand for health care are intimately related, and it has been widely recognized and examined since Arrow (1963) pointed out the moral hazard problem. According to his claim, health insurance may allow the insured to over-utilize health care services since it distorts the effective price of services in the market. McCall et al. (1991), Christenson et al. (1987), Lillard and Logowski (1995), Khandker and McCormick (1999), and Atherly (2001) attempted to empirically estimate the impact of supplemental health insurance on medical service consumption and all claimed that the insured demands more services. There is no agreement, however, on how much of this demand gap between the insured and the uninsured can be explained by the moral hazard problem (Wolfe and Goddeeris, 1991; Ettner, 1997; Hurd and McGarry, 1997; Chiappori et al., 1998; Buchmueller et al., 2004; Sapelli and Vial, 2003). Moreover, estimating the moral hazard effects entails significant problems in that not only is the demand for health care determined by the insurance status but also the insurance decision itself depends on expected future consumption of health care services; there may be mutual interdependence between health insurance and health care, leading to the simple regression estimator biased (Cameron et al. 1988; Edward et al. 1995).

The previous studies listed above did not consider this problem.

In Korean context, Yu (1983) conducted a pioneering study when the universal health care system was not yet established in Korea but only about thirty percent of the total population was covered by various health insurance plans. He chose four categories of treatments (appendectomy cholecystectomy, tonsillectomy, cesarean section and respiratory tract infection), and investigated whether there were differences in quantities, length, charges, and qualities of medical care services between the insured and the non-insured. The major findings confirmed the moral hazard hypothesis; the volume of services was greater and the length of stay was longer among the insured whereas quality of care did not show statistically significant differences. On the other hand, Seo (1981) suggested a confounding result in his study during the similar period. The author classified hospitals in three categories - university hospital, general hospital, hospital — and selected two diseases — acute appendicitis and normal delivery — to compare the patterns of medical consumption before and after introduction of public health insurance in Korea. The results show that there were no significant differences in the average length of stay and the average total hospital costs of both diseases in all types of hospitals. These two studies might have yielded inconsistent results due to the different methods adopted in selecting and categorizing the samples, and more importantly, the institutional instability during that time.

Since the National Health Insurance (NHI) began to play a role as primary insurance while private health insurance as supplemental in 1989, there have been a heated debate on whether private health insurance creates social inequality and worsens the financial status of the NHI; however, most of them are lack of specific ground and empirical evidence. Thus, Yun and Kwon (2008) recently attempted empirical analysis based on the survey conducted from 2004 to 2006. The total sample used in this research was 406, 751, acquired by a random sampling method, and was divided into two groups, private insurance holders and the other group, to see whether they are different in medical care utilization level. The results imply that the private insurance holders do not consume more medical services than the others, which can alleviate concerns over the moral hazard problem in Korean private insurance market. The analysis, however, is solely based on mean and percentage calculation; by this method, the effects of other factors that can influence dependent variables and individuals' own characteristics are not controlled.

As one can notice, the moral hazard problem of private health insurance in Korea is still controversial and understudied. Its importance as supplemental insurance to the NHI, on the other hand, continues to increase as the NHI coverage rate remains only about 65% out of total medical expenditure, much lower than 80% of the average rate for OECD member countries, and as the proportion having private health insurance is as high as 64% of the total population.[1] Furthermore, because the privatization policy is one of the key elements of the current President Lee's economic plan, various ways to promote private insurance are also largely being discussed. Therefore, it is important to investigate whether the moral hazard problem exists in private health insurance and whether it places more financial burdens to the NHI which has suffered from budget deficit for recent years. It may also be useful to extend the results to other countries where the universal health care system exists and private insurance supplements it. Having considered its importance, therefore, this study seeks to find answers for the question

[1] These were all measured in 2008.

posed at the beginning by adopting econometric techniques and deriving estimates as robust as possible.

III. *Empirical Analysis*

1. Models

As previously mentioned, this study aims to uncover moral hazard effects in private health insurance. Yun and Kwon's latest study (2008) performed a research to empirically prove such effects, relying on mean and percentage values. By this method, however, causal relationships are not revealed; if the group of the insured consists of people who have certain characteristics in common, high or low mean of medical service demand may not result from insurance status but from those idiosyncratic characteristics. Thus, to control this effect, we first estimate the probability of purchasing private health insurance. Besides, since Korea has the universal health care system, it is worth studying *who* has additional supplemental private insurance and *what* determines the decision. The following equation is estimated by using probit model.

$$I_i = \pi_0 + \pi_1 X_i + \pi_2 H_i + \pi_3 B_i + \pi_4 PR_i + \omega_i \quad (1)$$

I_i: a dummy variable indicating private insurance status
(1 if one is enrolled in private health insurance and 0 otherwise)
X_i: a vector of socio-demographic variables
H_i: a vector of health status variables
B_i: a control variable measuring individuals' efforts or preferences to improve health
PR_i: National Health Insurance (NHI) premium each individual pays.
ω_i: an error term assumed to follow a white-noise process

The socio-demographic variables include age, gender, marital status, education level, occupational status, number of family members, and income level. Other key explanatory variables are related to individuals' health status. This empirical analysis defines health status in two different ways: one is a subjective five-point scale answer as for the question "how do you think of your overall health conditions?"[2]; the other is an objective measure of health by summing up the number of all chronic diseases that the respondent has, if any, such as diabetes, high blood pressure, anemia, asthma, arthritis and so on. We also tried to measure each individual's "intention" to improve health; if one puts higher efforts to enhance his or her health, it would mean that they care more about physical well-being and thus their willingness to consume health care services and willingness to pay for them could be higher than otherwise. There is no such a straightforward question in the questionnaire. Therefore, we defined four areas standing for health promotion behaviors as 1) alcohol consumption, 2) cigarette consumption, 3) regular exercises, and 4) weight management, and indexed them into three-scale points; if one endeavored much, satisfying certain criteria, he will get 2 points; if one performed all right, he will get 1 point; if one was reckless in caring health, he will get nothing.[3] These are summed up and reduced to one variable referring to "degrees of caring

[2] The answers can be 1) very bad, 2) bad, 3) normal, 4) good, or 5) very good.

[3] 1) For drinking, we assigned 2 if one has never taken any alcoholic beverage, 1 if one drinks less than once a

health". Lastly, insurance premium is included.

Then we verify whether private health insurance affects individuals' decisions to visit doctors. The equation (2) is established to estimate the effect.

$$PV_{1i} = \alpha_0 + \alpha_1 X_i + \alpha_2 I_i + \alpha_3 H_i + \alpha_4 B_i + \varepsilon_i \quad (2)$$

PV_{1i}: the annual number of physician visits including in- and out- patients[4]

ε_i: an error term assumed to follow a white-noise process

In estimating the equation (2), Poisson and negative binomial models can be considered given that both are good models for count data (Cameron and Trivedi, 1998). In the Poisson model, the variance in the number of accidents is equal to the mean whereas in the binomial model, the variance is larger than the mean; the latter is more appropriate for the data overdispersed. In order to see which model works better, we performed a goodness-of-fit test by using the Pearson statistic after estimating the Poisson regression. If the test is significant, the Poisson model can be viewed as inappropriate and the negative binomial model should be considered. The test result, however, does not allow us to reject the null hypothesis that the Poisson distribution fits the data, so it is adopted as a suitable model[5].

However, as many health economists insist (Cameron, *et al.*, 1988), an individual's expected medical care utilization could affect his or her decision on choice of health insurance, and the existence of health insurance could affect his utilization of medical care services, which suggests there is mutual interdependency between the decision of medical care use and health insurance purchase. Hence, a simple Poisson regression model may not produce consistent estimates of parameters. In other words, if two error terms of equation (1) and (2) are correlated each other which can be a source of endogeneity bias, it is highly likely that the unobservable characteristics included in ε_i of equation (2) are correlated with some of the observed explanatory variables in that equation. Therefore, to control such endogeneity problem, we adopted the two-step method by using both equation (1) and (2). That is, the predicted value of $\hat{I}_i$, which is obtained from estimating equation (1), substitutes the actual value of I_i, and the same Poisson regression method is used for the second step.

In this process of applying the two-step method, as shown in equation (2), PR_i, the NHI insurance premium variable is utilized as an instrumental variable. According to Wooldridge (2002), when the estimation model includes one endogenous variable, at least one instrument is necessary to identify it. Intuitively, using PR_i as an instrument can be justified by the way the

---

month, and 0 otherwise; 2) For smoking, we assigned 2 if one has never smoked, 1 if one had smoked in the past but stopped recently, and 0 otherwise; 3) For exercises, we assigned 2 if one has performed mid-level physical activities, such as tennis, badminton, yoga, or swimming, more than 5 days a week and over 10 minutes per day, 1 if one has performed those at least one day per week, and 0 if one has performed nothing; 4) For weight management, we assigned 2 if one's BMI is more than 18.5 and less than 25, 1 if BMI is less than 18.5 or more than 25 but still less than 30, and 0 if BMI is over 30.

[4] Since ex-post moral hazard assumes that patients can choose arbitrary amount of medical services, it seems important to make the concept of "the number of visits" be clear. Although the total amount of medical care use could be affected by the supply side due to medical knowledge imbalance between doctors and patients, patients can choose whether or not to visit doctors when they feel sick. So we define "the number of visits" to be "the number of episodes" in this paper. This concept suggests a patient's decision on visiting a doctor is mainly determined by the patient him/herself even if the supply side could affect patient's follow-up visit with the same episode. The information on who initiated the respondent's first doctor-visit with an episode is clearly shown at the survey questionnaire this study used.

[5] Test results are available upon request.

Korean government exogenously imposes the contributions of NHI. First, for the employee insured, the contribution amount is calculated by multiplying monthly salary to the contribution rate, and is deducted from the monthly salary. Second, for the self-employed insured, the contribution amount for those who have an income is calculated in the same way above, but for those who have no income, their contribution amount shall be equal to the average monthly contribution of the Korean self-employed in the previous year. Hence "NHI premium" might work as an appropriate instrument variable in that it is unlikely to be related to individual medical service use because the mechanism of determining it by Korean central government does not truthfully represent the individual real income level, and because the individual use of medical care might be mainly affected by people's health conditions or private health insurance status as the *ex post* moral hazard assumption predicts. Thus we can secure the variable's exogeneity and this paper will discuss more about the validity of the instrumental variable later by conducting a series of tests and by introducing the mechanism of determining the premium in detail.

Another econometric issue that may bias the estimation result of equation (1) and (2) above is the existence of sample selection bias. This sample selection problem has long been discussed in health economics literature; in some data which involve a number of zero responses, a researcher may face such questions as "how would *many zeros* affect estimation results?" and "do those zeros represent *actual choices* of individuals?" In our data representing medical service utilization, there are 2,456 non doctor-visits out of total 4,513, thus we have to verify how results will change after properly excluding them. According to Duan et. al, (1984), censored data approach requires restrictive distributional assumptions and as the censored data are unobservable, these assumptions are not testable. Maddala (1985), however, stresses the need to understand the nature of the underlying decision process in selecting an empirical model and argues that joint decisions of visiting a doctor and the amount of consumption of medical care conditional on any visit may be more appropriate. In other words, if there are some common omitted variables, the two decisions will be correlated which might cause the estimation results to be biased. Thus, it is worth trying to jointly estimate, and check whether the sample selection problem which is caused by correlation of two error terms, is critical in the model.

Heckman (1976, 1979) has suggested a practical model to deal with such situations, which treats the selection problem as an omitted variable problem. Although this solution known as the two-step or the limited information maximum likelihood (LIML) method has been recently criticized due to its small-sample properties, according to Puhani's Monte Carlo studies (2000), it still yields reasonable results except when multicollinearity problems are critical; in this case, subsample OLS (or the Two-Part Model) is the most robust among the easy-to-implement estimators (Puhani, 2000). Hence, in this paper, in order to test multicollinearity, the Variation Inflation Factor (VIF) was used, and the result was 1.48, much lower than that of a concern.[6]

Heckman's model, however, is a standard approach to account for sample selectivity in the linear regression model, so it is not appropriate for our model which is based on count data. Greene (1995) has extended this Heckman model to the Poisson regression model for count data; the procedure begins by fitting the probit model as earlier, computing the inverse Mill's ratio — the ratio of the probability density function over the cumulative distribution function

[6] As a rule of thumb, if the VIF is larger than 10.0, then multicollinearity is regarded as a problem.

$\left(\frac{\Phi\left(\frac{x'\beta}{\sigma}\right)}{\Phi\left(\frac{x'\beta}{\sigma}\right)}\right)$ — for the selected observations, then fitting the Poisson model with this augmented conditional mean function by maximum likelihood. We adopt this method as another alternative.

The equation (3) is the selection equation to estimate individual *i*'s probability of visiting a doctor.

$$\begin{aligned} &PV_{2i}=\beta_0+\beta_1 X_i+\beta_2 \hat{I}_I+\beta_3 H_i+\beta_4 B_i+\beta_5 TIME_i+\mu_i \qquad (3)\\ &\begin{cases} d_i=1 & \text{if } PV_{2i}>0 \\ d_i=0 & \text{otherwise } (PV_{2i}=0) \end{cases} \end{aligned}$$

PV_{2i} is a dichotomous variable expressing whether or not one has ever seen a doctor. The variable TIME represents travel time that one must spend to see a doctor and it is added for the identification issue. Our intuition behind choosing TIME as another exogenous variable is that while an individual's decision on whether to visit doctors could be influenced by travel time, once she or he visits a doctor, the total amount of medical care use is largely affected by suppliers, considering the existence of asymmetric information between patients and doctors. If it really is so, the role of this variable as an instrument can be justified. The validity of this identification will be discussed in the next section.

Conditional on positive number of doctor visits, an individual *i*'s total medical care use can be estimated by the next equation.

$$PV_{1i}(d_i=1)=\alpha_0'+\alpha_1' X_i+\alpha_2' \hat{I}_I+\alpha_3' H_i+\alpha_4' B_i+\rho\kappa_i+\varepsilon_i' \qquad (4)$$

The κ_i is a new variable standing for the inverse Mill's ratio. According to the Heckman's sample selection model, the existence of error-correlation caused by sample selection can be checked by testing the hypothesis of $\rho=0$. If we can reject this hypothesis at the legitimate significance level, it would be desirable to interpret the results by the Heckman's model while simple estimation without controlling sample selection bias yields better outcomes if we cannot (Hay and Olsen, 1984). We will use this criterion to assess which model is more reliable later. The standard errors are adjusted using the robust estimator of variance.

In addition to the number of doctor visits, this study also uses total annual medical expenditures as another dependent variable conditional on any doctor visit. The reason for using another dependent variable arises from the concept of *ex post* moral hazard. As explained, it indicates a phenomenon that people consume more services than they would otherwise if they have medical insurance. Thus, the moral hazard of this kind could occur where people expect private health insurance they hold to reimburse the amount of money they spend for medical services so that they want to use them more frequently, being less careful of costs, and utilize more expensive ones. The frequency of physician visits only captures the aspect of patient's being less careful of costs but not the specific contents of medical care a patient uses, so the need for including a variable representing how much one spends for medical services comes up. Even if it is highly plausible to think the specific contents of medical care a patient uses might be largely affected by doctors, it is hard for them to know the patient's idiosyncratic characteristics including patient's private insurance status unless a patient clearly provides a doctor with the information in Korea[7]. So, a patient's willingness to utilize more expensive medical care services originated from his/her having private health insurance, i.e. ex-post moral

TABLE 1. EMPIRICAL MODELS AND STRATEGIES

	Physician Visits			Medical Costs		
	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Endogeneity Control	N	Y	Y	N	Y	Y
Sample Selection Control	N	N	Y	N	N	Y

Note: N - No; Y - Yes

hazard might be analyzed using total annual medical expenditures. Besides, because the frequencies of health care utilization — i.e. the number of doctor visit - do not contain information on patient's severity of illness, how much one pays per visit can provide more relevant information. Hence, this part can be viewed as a sensitivity check to see whether there exists moral hazard even when analyzing the specific contents of medical services people used and controlling the severity of diseases people hold.

In doing so, we transformed medical costs into a natural logarithmic form in order to correct the left-skewed distribution and to scale down the size of coefficients for the sake of simplicity. The variable, denoted as LMCOST, displaces PV_{1i} in equation (2) and (4), and since LMCOST is a continuous variable, it is fair to adopt the OLS method instead of the Poisson regression. So we performed the same procedures by adopting three estimation methods: 1) Ordinary Least Squares (OLS) as basic model, 2) Two-Step Least Squares (TSLS) to correct endogeneity bias which could be caused by individual insurance choice being endogenously determined, and 3) Heckman's sample selection model to consider potential sample selection bias; because no visit means no expenditure we may want to censor those who have never seen a doctor. We will compare the results produced by each model as shown in Table 1.

2. Data

This paper uses the National Health and Nutrition Examination Survey (NHANS) in Korea which was conducted in 2005. Although there exist more up-to-date survey results available, obtained in 2007, it is more legitimate to choose data of 2005 in that they have a larger sample size and more affluent information on private health insurance such as the number of private insurances people have, reasons to purchase them, and most importantly how much people pay for those; on the other hand, data of 2007 do not contain such useful information while only reporting whether or not the respondent is insured by private companies. Furthermore, the question itself is different: the question in 2005 asks "do you have some private insurance like cancer insurance sold by private companies?" The same question in 2007, however, asks somewhat differently - "do you have some private insurance such as cancer insurance, cardiovascular disease insurance or *auto insurance*?" To be more precise, it seems fair to include only health-related insurance for the following analysis, excluding probabilities of any bias that can be caused by embracing other insurance types which are seemingly unrelated with health-care issues. More importantly, the Korean government allowed private health insurance in 2003, and the enrollment rates soared up, and its total market size was approximately $12

[7] A patient's medical record or consultation sheet which a doctor can refer to only includes a patient's disease code and main symptoms of disease, so whether a patient has private health insurance cannot be observed by a doctor in Korea.

billion in 2005. To reflect all factors above, we adopted the NHANS of 2005.

Of total observations of 34,152, we cut off data of people either younger than 19 or older than 65 given the assumption that the decisions of people belonging to these groups regarding whether or not to consume medical services are likely to be influenced by their family members or other acquaintances; as a result, our target dataset only include 11,431 observations. In addition, we further excluded observations which had more than one missing values in critical socio-economic variables such as income or education level. In the end, the estimation was performed with 4,513 observations. Since reducing more than half of the sample size due to missing values could yield some problems, we have to leave them as a limitation of this study given the data.

IV. *Empirical Results*

1. Descriptive Statistics

Table 2 shows basic descriptive statistics of variables introduced above. There are several points that should be made. First, the number of physician visits and medical costs appear to be greater for those who are not covered by private health insurance, contradicting standard moral hazard assumption. In case of the medical cost variable, this may result from several outliers who did not purchase private health insurance but underwent severe diseases. They might have paid a huge amount of out-of-pocket money for medical services. As for the doctor-visit variable, we conducted t-test to compare the sample mean of each group, and we were not able to reject the null hypothesis that two sample means were equal at 5% significance level.

Regarding control variables, degrees of performing health promotion behaviors are not different between the insured and the uninsured groups according to Table 2. Other variables indicate that people with more education, higher position in workplace and higher monthly income more tend to purchase private health insurance. On the other hand, those who report better health conditions and less chronic diseases are more covered by private health insurance. If so, the pool of the insured may consist of healthier and wealthier individuals, and accordingly one may conjecture that moral hazard is not observed in the private health insurance market if taking into account such characteristics. The following section explores whether this implication can also be supported when we adopt econometric techniques.

2. Estimation results

(1) Individual choice of private health insurance

Table 3 displays results of estimating equation (1); people are more likely to purchase private health insurance if they are younger, more educated, have higher monthly income, and have less chronic diseases. The NHI insurance premium variable, however, indicates the higher premium, the more demands. To understand why, it is necessary to understand that the premium variable is not merely an indicator for the price of national health insurance or the level of income, but may contain far more information. As mentioned before, it is determined by the Korean government based on the income level of the insured, but because of the way

TABLE 2. DESCRIPTIVE STATISTICS OF SUB-GROUPS

Variables	Descriptions	Covered by Private Health Insurance				Not Covered by Private Health Insurance			
		Mean	St. Dev.	Min.	Max.	Mean	St. Dev.	Min.	Max.
Choice Variables									
PV	Number of Physician Visits	1.24	2.210	0	25	1.41	2.043	0	14
MCOST	Total Medical Costs (Korean Won)	85690.50	9	0	10037000	143441.95	6.000	0	16998000
Socio-Demographic Variables									
AGE	Age	41.49	11.280	19	65	47.40	13.903	19	65
SEX	Sex(0: male ; 1: female)	0.42	0.493	0	1	0.45	0.498	0	1
MARRI	Marital Status (0: not married ; 1: married)	0.82	0.530	0	1	0.76	0.789	0	1
EDUC	Educational Status (1: no education ; 2: $\leq$ elementary school graduate ; 3: $\leq$ middle school graduate; 4: $\leq$ high school graduate ; 5: $\geq$ college graduate)	4.04	1.047	1	6	3.42	1.319	1	6
NFAM	Number of Family Members	3.50	1.176	1	7	2.82	1.216	1	7
JOB_STT	Job Status (0: employees or part time jobs; 1: employers or full time jobs)	0.49	0.500	0	1	0.35	0.476	0	1
INC	Monthly Income (Ten Thousand Korean Won)	272.23	160.642	0	998	159.06	124.112	0	998
Health Status Variables									
SHS	Subjective Health Status	2.30	0.818	1	5	1.98	0.964	1	5
NCD	Number of Chronic Diseases	2.69	2.142	0	15	3.31	2.500	0	15
Control Variables									
BEHAV	Health Promotion Behaviors (degrees of caring health)	4.67	1.591	0	8	4.67	1.604	1	8
Instrumental Variables									
PREMIUM	NHI Premium (Ten Thousand Korean Won)	25.38	225.126	0	99.8	10.21	131.667	0	99.8
TIME	Travel Time to Visit Doctors (minute)	7.64	29.706	0	888	10.94	27.741	0	300
N		3600				913			

TABLE 3. PROBIT ESTIMATION RESULTS ON PRIVATE INSURANCE STATUS

Variables	Descriptions	Private Insurance Enrolment	
		coefficients	s.e.
CONSTANT		−3.1000	0.4907
Socio-Demographic Variables			
AGE	Age	−0.0741***	0.0194
SEX	Sex	−0.1092	0.0835
MARRI	Marital Status	−0.0190	0.0485
EDUC	Educational Status	0.4396***	0.1547
NFAM	Number of Family Members	−0.0090	0.0302
INC	Monthly Income (Ten Thousand Korean Won)	0.0005*	0.0003
JOB_STT	Job Status	0.0940	0.0742
Health Status Variables			
SHS	Subjective Health Status	0.0373	0.0426
NCD	Number of Chronic Diseases	−0.0455***	0.0170
Control Variables			
BEHAV	Health Promotion Behaviors (degrees of caring health)	0.0115	0.0238
Instrumental Variables			
PREMIUM	NHI Premium (Ten Thousand Korean Won)	0.0207***	0.0006
Pseudo R^2		0.6304	
N		4513	

Note: * = statistically significant at the 0.1 level; ** = statistically significant at the 0.05 level; *** = statistically significant at the 0.01 level

the government imposes it, the variable also contains information which cannot be captured by researchers, which allows us to explain the ambiguous result on the regression coefficient. The following reasone also confirm the appropriateness of the variable as an instrument.

First, there exists a time gap between when the government imposed the NHI premium and when the survey respondents reported their income. To be specific, the survey on which this study is based was conducted in *2005*, and the questionnaire asked the respondent's current income level of *2005*. The NHI premium of *2005* determined by the Korean government, on the other hand, was retrospectively determined based on the individual income level of *2004* because it is hard to recognize the yearly income level of 2005 at the early stage of that year. The over- or under-charged NHI premium due to the time gap is supposed to be adjusted in the April of the following year 2006. In brief, in Korean National Health Insurance system, the NHI premium is based on the lagged individual income level which does not truthfully represent the current individual real income level.

Second, the Korean government decides the NHI premium based on seven income brackets, but not on the actual level of income. Thus, there could be a great difference between the actual income people have and the income bracket they belong to; for example, people in the sixth bracket could be those who earn 50,000,000 Korean won but also could be those who gain 90,000,000 Korean won on a yearly basis, and this gap cannot be disregarded. Hence this mechanism of setting NHI premium suggests that the unobserved individual heterogeneity in terms of individual real income is not considered in the process of determining it, which also

indicates the NHI premium is somewhat deviated from the individual real income level.

Furthermore, what is more important fact on this mechanism is that, as already mentioned in this paper, the NHI premium for the self-employed whose income is not known to Korean government is set to be equal to the average monthly NHI premium of the overall Korean self-employed in the previous year. Therefore, this mechanism could also yield a significant disparity from their real income level and the unobserved individual heterogeneity in terms of his/her real income is not considered in the process of determining it either. Nowadays, a number of academics, legislators, and media in Korea criticize the mechanism to set the NHI premium in Korea, claiming that it does not properly reflect the real income level of the insured and accordingly it worsens financial deficits of public insurance.

Having considered all these aspects, we can conclude that the individual real income level and the NHI premium might be correlated in part, but each represents their own characteristics. In this context, concerning the positive coefficient of NHI premium on individual private insurance enrollment, it is highly likely that the characteristics of the mechanism of determining NHI premium explained above may have produced the positive effect. All in all, the variable of NHI premium is not merely an indicator of the individual level of income, probably containing far more information which cannot be captured by researchers, and this fact may allow the variable to function properly as an instrument.

(2) Effect of private health insurance status on number of doctor visit

Model 1, 2, and 3 in Table 4 represent the simple Poisson regression, the two-stage Poisson regression, and the Heckman's LIML respectively. The most interesting result is that purchase of private health insurance largely increases the number of doctor visit and it is also statistically significant in the simple Poisson regression model. This can be harmonized with the moral hazard assumption in the health insurance market. In the two-step Poisson regression, however, the coefficient decreases dramatically and it also loses statistical significance. Hence, this result does not support the presence of the moral hazard effects in Korean private health insurance market, and the endogeneity bias is generated by individual unobserved characteristics associated with the choice of private insurance.

Concerning the validity of instrumental variable, PREMIUM, used for identifying equation (1) and (2), we tested the relevance and exogeneity of it. With respect to the relevance, we used the Weak IV test suggested by Stock and Watson (2007). One way to check for weak instrument when there is a single endogenous variable is to compute the F-statistic testing the hypothesis that the coefficient on the instrument is equal to zero in the first stage regression of two stage least squares (TSLS); if it is less than 10, the instrument can be regarded as weak, and the TSLS estimator is biased (Stock and Watson, 2007). In our case, however, the first-stage F statistic is 13.125, supporting the relevance of the instrument used. One should note, however, that the argument above is based on a linear regression model, not Poisson model with the probit method. Hence, the tests could not be rigorous and we still need to call for our intuition to judge the validity explained above.

Testing instrument exogeneity is another issue of concern; if the instruments are not exogenous, then TSLS is inconsistent. Because we have only one instrument and one endogenous regressor, it is impossible to develop a statistical test of the hypothesis that the instruments are in fact exogenous, such as the overidentifying restriction test called J-statistic. In this case, the only way to assess whether the instrument is exogenous is to depend on

TABLE 4. ESTIMATION RESULTS ON DEMAND FOR HEALTH CARE SERVICES

Variables	Descriptions	Model 1		Model 2		Model 3			
		Physician Visit		Physician Visit		Probability of Physician Visit		Conditional Mean	
		coefficient	s.e.	coefficient	s.e.	coefficient	s.e	coefficient	s.e
CONSTANT		0.2738	0.2129	0.2976	0.2138	0.4580	0.3003	2.3135***	0.6932
Socio-Demographic Variable									
AGE	Age	0.0017	0.0087	0.0034	0.0087	−0.0294***	0.0123	0.0495***	0.0189
SEX	Sex	−0.0867**	0.0362	−0.0882**	0.0363	−0.0465	0.0507	0.0934	0.0709
MARRI	Marital Status	0.0295	0.0215	0.0292	0.0215	0.0830*	0.0362	−0.1281**	0.0564
EDUC	Educational Status	0.1052*	0.0633	0.1193*	0.0631	0.0656	0.1003	−0.0573	0.1204
NFAM	Number of Family Members	−0.0335***	0.0120	−0.0339***	0.0120	−0.0299*	0.0174	0.0276	0.0758
JOB_STT	Job Status	−0.0142	0.0313	−0.0116	0.0313	0.0795*	0.0448	−0.1730**	0.0758
INC	Monthly Income (Ten Thousand Korean Won)	−0.0002*	0.0001	−0.0003**	0.0001	−0.0003**	0.0001	0.0005	0.0003
Health Status Variables									
SHS	Subjective Health Status	−0.2653***	0.0176	−0.2627***	0.0175	−0.2009***	0.0260	−0.1751	0.1200
NCD	Number of Chronic Diseases	0.0920***	0.0061	0.0933***	0.0061	0.1072***	0.0111	0.1074*	0.0560
Control Variables									
BEHAV	Health Promotion Behaviors (degrees of caring health)	−0.0043	0.0098	−0.0034	0.0098	−0.0335***	0.0121	−0.0360	0.0244
Instrumental Variable									
TIME	Travel Time to Visit Doctors (minute)					0.0033***	0.0002		
Insurance Status Variables									
PINS	Private Insurance Status	0.1564***	0.0350						
PINS_HAT	Predicted Value of PINS			0.0126	0.0036	0.0026	0.0048	0.0062	0.0055
Pseudo R^2 [8]		0.0844		0.0840		0.0683		0.0888	
Heckman's Mills Ratio								−3.5311***[9]	0.9066

personal knowledge of empirical problem at hand (Stock and Watson, 2007). As mentioned initially, our common sense indicates that the specification of model 2 is appropriate because it is less plausible that people's use of medical care is affected by the level of NHI premium because it is exogenously determined by the government.

[8] Even though the R^2 are not large enough in all models, according to Kenneth et al. (1995), the R^2 in this range is quite common in cross-sectional data. Thus, it is still worth comparing three models' results.

[9] Statistical significance is evaluated based on likelihood ratio test of $\rho = 0$.

When it comes to model 3 which considers the sample selectivity problem, we can reject the null hypothesis that $\rho=0$ at 1% significance level, so it suggests that the estimation results from the Heckman's LIML are more reliable.

In this model, the reduction of the effect of private health insurance on the total number of doctor visit is more apparent. The coefficient on PINS_HAT which is a predicted value of PINS from the first step, decreases further and displays higher standard errors. Hence, the result of considering both endogeneity bias and the sample selectivity problem also does not support the presence of the moral hazard effects in Korean private health insurance market.

One more thing to be mentioned in estimating model 3 is the validity of TIME variable, the instrument supposed to identify equation (3) and (4). As in the previous case, the validity of it was tested to check its relevance and exogeneity. First, the relevance was verified by computing the first stage F-statistic. It was 18.045, confirming that the instrument is not weak. For exogeneity, because model 3 is also just-identified, we cannot conduct additional tests such as the overidentifying restriction test.[10] The only way to check the exogeneity of instrument is therefore dependent on personal intuition again (Stock and Watson, 2007). Intuitively, the travel time can affect an individual's decision on whether to visit a doctor since it is related to accessibility to the health care resources; but it is less likely to determine the total amount of medical care use, which is largely influenced by supply-side factors such as physicians' practice style or their incentive structures (Lim and Jo, 2009) due to medical knowledge imbalance between patients and doctors.

When it comes to other variables, the results support existing theories which have been widely accepted. For example, the older have a tendency to consume more medical services, and the job status is a significant factor of determining the total number of doctor visit. Also, the number of chronic diseases is a crucial factor of individual use of medical care services.

(3) Effect of private health insurance status on total medical care cost

The same procedure with the log of medical expenditures as a dependent variable is performed and the results are presented in Table 5. As in the previous case, PREMIUM and TIME, the instrumental variables, were proven to be valid instruments. The F-statistics also confirm that the instruments are not weak.

Among model 1, 2, and 3, model 3 which applies the Heckman's LIML can be considered as the most reliable one given that the null hypothesis of $\rho=0$ is rejected by likelihood ratio test, which indicates that the error correlation caused by sample selection bias exists.

Estimates from three different models exhibit similar patterns with the previous case (number of physician visits); the coefficients of individual status of private health insurance identically show positive signs, but the statistical significance is not guaranteed in all cases. Moreover, the highest value in simple OLS regression is reduced after endogeneity and sample selection bias were controlled. As mentioned previously, the reason we adopt the total medical cost as another dependent variable is that the frequency of doctor visit cannot capture the specific contents of medical care services a patient uses, and cannot control the effect of patient's severity of illness. Therefore, we can strengthen robustness of the estimation results

[10] We have added other possible instrumental variables such as travel costs and specific insurance types to perform exogeneity test of instrumental variables, but all of other variables were proved to be weak instrumental variables based on Weak IV test. So in this paper, only travel time was used as instrumental variable.

TABLE 5. ESTIMATION RESULTS ON MEDICAL COSTS

Variables	Descriptions	Model 1		Model 2		Model 3			
		Medical Costs		Medical Costs		Probability of Physician Visit		Conditional Mean	
		coefficient	s.e.	coefficient	s.e.	coefficient	s.e	coefficient	s.e
CONSTANT		11.1808***	0.6916	11.1646***	0.7080	0.4580	0.3003	10.0193***	1.1778
Socio-Demographic Variables									
AGE	Age	0.0680**	0.0271	0.0666**	0.0272	−0.0294**	0.0123	0.1067**	0.0413
SEX	Sex	−0.1081	0.1115	−0.1092	0.1152	−0.0465	0.0507	−0.3889*	0.2035
MARRI	Marital Status	−0.1905***	0.0733	−0.1911***	0.0733	0.0830*	0.0362	−0.3606***	0.1191
EDUC	Educational Status	−0.3369	0.2065	−0.3289	0.2062	0.0656	0.1003	−0.3234	0.2653
NFAM	Number of Family Members	−0.0883**	0.0390	−0.0878**	0.0391	−0.0299*	0.0174	−0.1911**	0.0756
JOB_STT	Job Status	−0.1972*	0.1010	−0.1963*	0.1010	0.0795*	0.0448	−0.1975*	0.1141
INC	Monthly Income (Ten Thousand Korean Won)	0.2128***	0.0763	0.2170***	0.0785	−0.0003**	0.0001	0.1208	0.0985
Health Status Variables									
SHS	Subjective Health Status	−0.3000***	0.0176	−0.2982***	0.0545	−0.2009***	0.0260	−0.7324***	0.2634
NCD	Number of Chronic Diseases	0.0156	0.0215	0.0166	0.0214	0.1072***	0.0111	0.1704*	0.1017
Control Variables									
BEHAV	Health Promotion Behaviors (degrees of caring health)	0.0413	0.0314	0.0414	0.0314	−0.0335***	0.0121	0.0743*	0.0407
Instrumental Variable									
TIME	Travel Time to Visit Doctors (minute)					0.0033***	0.0002		
Insurance Status Variables									
PINS	Private Insurance Status	0.0813	0.1108						
PINS_HAT	Predicted Value of PINS			0.0032	0.0110	0.0026	0.0048	0.0051	0.0120
R^2		0.0407		0.0405		0.0683		0.0439	
Heckman's Mills Ratio								3.7663*	2.1614

with this finding.

Table 6 summarizes the effects of private health insurance on each dependent variable. As one can see, all estimates are not statistically significant except that of the simple Poisson regression model for physician visits. After controlling the endogeneity and the sample selection problems, however, it also loses statistical significance and the size of the coefficient also decreases. Thus, these results consistently implicate that the moral hazard effect caused by private health insurance disappears in Korean case if one controls those confounding factors.

The following can be suggested as one of the reasons that we could not detect the moral

TABLE 6. SUMMARY OF ESTIMATION RESULTS

	Physician Visits		Medical Costs	
	coefficient	s.e.	coefficient	s.e.
Model 1	0.1564***	0.0350	0.0813	0.1108
Model 2	0.0126	0.0036	0.0032	0.0110
Model 3	0.0062	0.0055	0.0051	0.0120

hazard effect in Korean private insurance market. As shown in Table 3, the more educated, the wealthier and the healthier people more tend to purchase private insurance; these people tend to have better accessibility to health care resources, and presumably have higher opportunity costs of being sick. Therefore, if we fail to control the effect of these individual features related to the private insurance choice, the effect of individual private health insurance status on medical care use could be overestimated. In other words, if there are some common omitted variables such as individual's time preference, marginal valuation of health status, or degrees of risk aversion, the decisions on purchasing private health insurance and those on medical care use could be correlated each other, which causes the estimation result to be biased. Here comes the reason why we should interpret the phenomenon with the results obtained by further analysis with Two-Step method and Heckman's approach.

V. *Conclusions*

To verify the existence of moral hazard in supplemental health insurance market, this study has investigated the relationship between private health insurance status and the use of medical services in Korea. In order to produce consistent estimates, we employed two-stage regression and Heckman's LIML approach. All estimates acquired by such econometric methods do not allow us to detect the presence of the moral hazard effects. This contradicts the results obtained by simple regression analysis and demonstrates the importance of controlling confounding factors; if we fail to control the unobserved characteristics behind the private insurance choice, the effect of insurance status on medical care use can be overestimated.

Although this paper attempts to discover the moral hazard effects in a possibly robust way by using several econometric methods, one has to be cautious in interpreting the results. Because Korea has universal health care system, private health insurance status should not be equally interpreted as that in other countries. For instance, in the United States, health insurance is primarily provided by the private sector, with the exception of programs such as Medicare and Medicaid — i.e. medical costs of most population are reimbursed by private companies. In Korea, on the other hand, the medical costs are covered first by National Health Insurance (NHI) system and the private insurance plays a supplemental role in covering medical costs that are not covered by NHI. In this respect, the results would rather allow us to understand how individuals' behaviors change when having supplemental insurance and facing the increased coverage rate. According to the hypothesis often discussed in information economics, people might consume more recklessly along with the expanded coverage rate. However, the estimates in this study indicate that this is not the case in practice, and rather that people hold private health insurance simply as "safety net."

Despite its contributions, this study encounters some technical limitations: the model

cannot fully control the reverse causality problem. The property of cross-sectional data restricts the precise estimation since they do not contain information on the sequence of events. In other words, we cannot exclude the possible reverse causality problem in which people consume medical services first, and purchase private insurance later; in this case, the same estimation methods used in this paper cannot be applied. Fortunately, however, the questionnaire is constructed in an order that coincides with the framework of this study. It asks 1) whether you are enrolled in private health insurance with no time limit, and 2) how much you have used or spent for medical services within periods ranging from 2 weeks to 1 year. Out of total respondents, the proportion of people just insured within these periods may not be large enough to affect the estimation results, so this case is not taken into account in this paper.

In addition, one further and important problem arises: the outcomes are likely to change if we distinguish whether the insured uses inpatient or outpatient medical care services. That is, it is conceivable that the outpatient services can be more affected by moral hazard than the inpatient services because the consumption of inpatient medical services is mainly determined by patient's fatal health status rather than economic motivation.

Because of the time limit and insufficient resources, this paper was not able to cover all these limitations. Nonetheless, it can contribute to providing useful insights to interpret consumers' behaviors under the distinctive health insurance system in Korea where national insurance and private insurance coexist and the latter supplements the former. Therefore, it is recommended for follow-up studies to deal with the remaining limitations so that they can suggest a way that the Korean government and those of other countries with the similar health care system fully take advantage of the role of private health insurance.

References

Anne, S.C. (2009), *Quantifying Moral Hazard in Health Insurance across Medical Services for the Nearl-Elderly*, Ph.D. dissertation, Georgetown University, Washington, DC.

Arrow, K.J. (1963), "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review* 53, pp.941-973.

Atherly, J.A. (2001), "Supplemental Insurance: Medicare's Accidental Stepchild," *Medical Care Research and Review* 58(2).

Buchmueller, T.A., M. Couffinhal, Grignon, and M. Perronnin (2004), "Access to Physician Services: Does Supplemental Insurance Matter? Evidence from France," *Health Economics* 13.

Cameron, A.C. and P.K. Trivedi (1998), *Regression Analysis of Count Data*, Cambridge University Press.

Cameron, A.C., P.K. Trivedi, F. Milne, and J. Piggott (1988), "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia," *Review of Economic Studies* 55(1), pp.85-106.

Chiappori, P.,F. Durand, and P. Geoffard (1998), "Moral Hazard and the Demand for Physician Services: First Lessons from a French Natural Experiment," *European Economic Review* 42.

Christenson, S., S. Long, and J. Rodgers (1987), "Acute Health Care Costs for the Aged Medical Population," *The Milbank Quarterly* 65(3).

Duan, N. et al. (1984), "Choosing between the Sample Selection Model and the Multi-Part Model," *Journal of Business and Economic Statistics* 2(3), pp.283-289.

Edward, C., V.T. Joseph, A.N. Cheryl, and C.S. Bruce (1995), "Estimating the Moral Hazard Effect of Supplemental Medical Insurance in the Demand for Prescription Drugs by the Elderly," *American Economic Review* 85(2), pp.122-126.

Ettner, S. (1997), "Adverse Selection and the Purchase of Medigap Insurance by the Elderly," *Journal of Health Economics* 16.

Greene, W. H. (1995), "Sample Selection in the Poisson Regression Model," NYU Working Paper No. EC-95-06.

Hay, J. and R.J. Olsen (1984), "Let Them Eat Cake: A Note on Comparing Alternative Models of the Demand for Health Care," *Journal of Business and Economic Statistics* 2, pp.279-282.

Heckman, J. J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic Social Measurement* 5(4), pp.475-492.

Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 47(1), pp.53-161.

Hurd, M. and K. McGarry (1997), "Medical Insurance and the Use of Health Care Services by the Elderly," *Journal of Health Economics* 16.

Khandker, R. and L. McCormick (1999), "Medical Spending by Beneficiaries with Various Types of Supplemental Insurance," *Medical Care Research and Review* 56(2).

Kenneth, A.B., K.G. David, and A.M. Thomas (1995), "Binary Outcomes and Endogenous Explanatory Variables: Tests and Solutions with an Application to the Demand for Contraceptive Use in Tunisia," *Demography* 32(1), pp.1.

Lillard, L. and J. Logowski (1995), "Does Supplemental Private Insurance Increase Medicare Costs?" Working Paper Series 95-16, RAND.

Lim, J. and C. Jo (2009), "The Effect of Patient's Asymmetric Information Problem on Medical Care Utilization with Consideration of a Patient's Ex-ante Health Status," *Hitotsubashi Journal of Economics* 50(2).

Maddala, G.(1983), *Limited and Qualitative Dependent Variables in Econometrics*, New York: Cambridge University Press.

Maddala, G. (1985), "A Survey of the Literature on Selectivity Bias as it Pertains to Health Care Markets," *Advances in Health Economics and Health Services Research* 6, pp.3-18.

McCall, N. et al. (1991), "Private Health Insurance and Medical Care Utilization: Evidence from the Medicare Population," *Inquiry* 28.

Pauly, M. (1968), "The Economics of Moral Hazard: Comment," *The American Economic Review* 58(3), pp.531-537.

Puhani, P. A. (2000) "The Heckman Correction for Sample Selection and Its Critique," *Journal of Economic Surveys* 14(1), pp.53-68.

Sapelli, C. and B. Vial (2003), "Self-selection and Moral Hazard in Chilean Health Insurance," *Journal of Health Economics* 22.

Seo, I. (1981), "Changes in Hospital and Clinic Care Patterns under the Medical Insurance System," *Korean Journal of Preventive Medicine* 14(1), pp.3-12.

Stock, J. H. and M.W. Watson (2007), *Introduction to Econometrics*, Prentice Hall.

Wolfe, J. and J. Goddeeris (1991), "Adverse Selection, Moral Hazard, and Wealth Effects in the

Medigap Insurance Market," *Journal of Health Economics* 10.
Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press.
Yu, S.H. (1983), "Study on Hospital Care Services between Insured and Non-insured Patients for Selected Diagnoses in Korea," *Yonsei Medical Journal* 24(1), pp.6-32.
Yun, H.S. and H.J. Kwon (2008), "Private Health Insurance and Medical Care Utilization," *Korean Development Institute Policy Forum* 204.
Zweifel, P. and W. Manning (2000), "Moral Hazard and Consumer Incentives in Health Care," *Handbook of Health Economics* Vol. 1.