Measuring the Impact of Voluntary Health Insurance on cost of healthcare with simultaneous correction for Care-Seeking and Insurance Self-Selection – A Case Study of Vietnam

Shehzad Ali[1], Richard Cookson[2], Tim Ensor[3] and Mohammad IftekherHossain[4]

ABSTRACT

Aims: This paper aims to estimate the impact of Vietnamese Voluntary Health Insurance (VHI) programme on out-of-pocket (OOP) cost of healthcare to the individual, controlling for observable need variables and also simultaneously correcting for both care-seeking and insurance-seeking self-selection biases.

Methods: Five econometric models have been used to account for self-selection biases. First model is general OLS regression model estimating cost of healthcare without correction of selection biases. Models 2, 3 and 4 use two-part models to first estimate care-seeking decision with or without correction of insurance endogeneity, followed by a Heckman's treatment effects model that simultaneously corrects for care-seeking and insurance-seeking selection biases. Model 5 uses a two-step model to correct for insurance self-selection only.

Data: This study is based on cross-sectional household survey data collected from 3 provinces of Vietnam. A total of 1,650 adults and 1,101 children were randomly selected and interviewed during the year 1999. In the sample, 1,184 individuals felt sick at least once in 3 months prior to the interview and of them 982 sought health care. Among those who sought care, 168 were insured while 814 were not.

Result: Econometric models suggest that the estimates from OLS model were biased due to care-seeking and insurance-seeking selection biases. Correction of these biases more than doubled the magnitude of the negative coefficient on insurance membership variable used in cost of healthcare analysis.

Conclusion: The Vietnamese VHI scheme appears to have a protective effect on cost of healthcare faced by individuals, after allowing for systematic differences in individuals who self-selected into care-seeking and insurance purchasing. This effect was under-estimated when correction procedures for self-selection biases were not employed.

INTRODUCTION

Out-of-pocket (OOP) payment or user fee is the predominant mechanism of health care financing in most developing countries. It forms an important barrier to accessing healthcare, especially for poor households, who may delay seeking advice until disease severity has progressed so far that they need prolonged and more expensive treatment (James *et al*., 2006; Waddington, 1989; Fabricant *et al*, 1999). OOP payments also have the potential to lead domestic financial pool to an inescapable vicious circle of debt and repayments. Empirical evidence shows that healthcare expenditures in developing countries are often far beyond moderate relative costs and frequently result in financial catastrophe for poor households (Xu *et al.*, 2003; O'Donnell *et al*, 2005).

[1] PhD candidate: Department of Social Policy and Social Work, University of York
[2] Senior Lecturer: Department of Social Policy and Social Work, University of York
[3] Principal Economist: Oxford Policy Management
[4] Lecturer: Department of Economics, University of Dhaka

To counteract the impoverishing consequences of user fee, many developing countries have embarked on formal and informal risk pooling mechanisms that decouple financial contributions from service use. One such mechanism is voluntary health insurance (VHI) programme, sometimes referred to as community-based health insurance or micro-health insurance scheme, which provides formal means of risk pooling for countries with predominantly informal economies (Witter *et al*, 2000). VHI programmes may be organised as fragmented small scale insurance schemes mostly set around not-for-profit health care providers, or they may be structured as regional or national programmes managed by the large public or private sector organisations. Countries like Vietnam and Ghana provide examples of VHIs operated on larger scale by the public sector.

In contrast with the private health insurance model used in many developed economies, voluntary insurance schemes in poor countries are based on the principles of affordability and equitable access (Ahuja and Jutting, 2004). Therefore, financial contributions are decoupled from health risk, and individuals either pay equal contributions irrespective of their health status, or payments vary with services covered. The pooling function of voluntary insurance also has the potential to transfer resources from relatively healthy to the sick within a population (Smith and Witter, 2001).

The potential impact of voluntary health insurance programmes on the cost of seeking healthcare is an important issue underlying policy debate in low- and middle-income countries. In recent years, a growing number of studies have analysed the role of VHI programmes in improving access to health care and reducing financial barriers, and most of them have suggested a desirable role of VHIs (Palmer *et al*, 2004; Ekman, 2004). While only a handful of studies have taken into account potential endogeneity bias attributed to choice-based decision of insurance enrolment, to authors' knowledge none of the studies on developing countries have simultaneously accounted for both care-seeking and insurance-seeking self-selection biases.

Using a case study of Vietnam, the aim of this paper is to study the impact of voluntary health insurance on the cost of health care faced by individuals in developing countries. This study does not look at cost of healthcare from the provider's perspective but focuses only on the cost incurred by the user for seeking care. While measuring the impact of VHI, the study will assess and, when required, correct for both care-seeking and insurance-seeking self-selection biases. The study uses data from a cross-sectional household survey conducted in three provinces of Vietnam in 1999.

The paper is organised in the following sections. First, we will discuss the potential causes and consequences of care-seeking and insurance-seeking self-selection biases. This will be followed by a detailed discussion of the econometric models used to assess and, if required, correct the above mentioned biases. Third section will elaborate on the context of the study and the structure of voluntary health insurance system in Vietnam, followed by a section on description of data and methods. Fifth section will provide results of the analysis, and the final section will discuss the findings of the study and their policy implications.

THEORY AND REVIEW OF LITERATURE

Measuring the impact of health insurance on the cost of healthcare is not a simple question to study because of the potential bias associated with the decisions to seek health care and purchase health insurance. While some had to ignore the bias due to data limitations (Sepehri *et al.,* 2006; Schneider and Hanson, 2006), a few other studies have corrected for insurance self-selection only. We start with a brief review of the theory and literature of the possible causes and forms of self-selection biases and some of the methods employed to overcome them.

Care-seeking self-selection

Care-seeking self-selection bias, or sample selection bias, occurs when outcome variable is only observed if some criterion, defined with respect to certain exogenous variable, is met (Breen, 1996; Bhat and Jain, 2006). In the context of cost of healthcare analysis, actual price paid for healthcare is only observed for individuals who decided to seek medical advice. Therefore, while a larger proportion of the sample may be sick and consequently in need of medical care, only a sub-sample actually seeks care and therefore have non-zero cost of seeking healthcare. There would be no cause for concern if the decision to seek medical care is randomly distributed in the population. But if the unobserved factors affecting the decision to seek medical care are also correlated with the unobserved factors determining cost of seeking healthcare, then the estimates on cost of healthcare regression based only on observed cost of care will produce biased estimates. If, for instance, more risk-averse individuals are likely to consume more medical care and in turn incur higher healthcare cost, then not accounting for individual risk behaviours in regression model for cost of healthcare would lead to biased estimates. In another example, if poor and sick individuals have a tendency to avoid seeking care due to high expected cost of healthcare, then the observed cost of seeking healthcare will be biased away from expected value of the cost in population. Jones (2007) would argue that if there are systematic differences between those who self-selected (for care in our case) and those who did not, then we can not directly identify the expected value and population regression function of the dependent variable (cost of healthcare in this case). Therefore, ignoring the sample selection bias would limit the external validity of our findings to individuals who felt sick and also sought health care.

Care-seeking self-selection bias is, in many ways, similar to the classic problem of estimating a model for female labour supply, where hours worked are observed only for the women who self-selected into the labour market at the offered wage. Women whose wage demand threshold is not met decide not to participate and hence have zero observed work hours (Winship and Mare, 1992; Kyriazidou, 1997; Wooldridge, 2002). Gronau (1974), in his classic paper, points at the possible consequence of sample selection bias in this context. His work was further extended by Heckman (1974) who showed that, when modelling female labour supply or their market wages, the presence of uncorrected sample selection bias will produce biased and inconsistent coefficients on OLS regression. While several examples of sample selection bias and its correction methods can be found in labour economics, its use in the analysis of cost of healthcare in developing countries is relatively rare.

Insurance-seeking self-selection

Both care-seeking decision and cost of seeking healthcare can be affected by insurance status of an individual. Insurance-seeking self-selection bias, or insurance endogeneity bias, occurs when the unobserved characteristics of individuals that make them more likely to self-select into an insurance programme also influence their decision to seek health care and their cost of healthcare (Gruber, 2000). This bias is more likely to occur in case of VHI programmes that tend to attract sicker individuals or individuals who expect to have high cost of healthcare (adverse selection). Regression models can control for differences in healthcare need but many inherent risk-related characteristics of an individual are not observed and therefore can not be controlled for in the analysis.

Besides individual intrinsic factors, there may be environmental factors affecting insurance self-selection and cost of seeking healthcare. Consider an example where the probability of an individual purchasing health insurance is reduced due to the large distance between his/her village and the nearest health facility where insurance card is accepted. This distance may also reduce their probability of seeking care when sick and in turn their total cost of healthcare, with or without insurance. If the distance variable is an unobserved factor in the study, then part of the calculated association between health insurance status and

cost of healthcare may not actually be due to insurance but to the underlying unobserved factors. Similar consideration can be given to other unobserved characteristics related to individual's personal preferences, which may affect the tendency to purchase insurance as well as health spending behaviour. However, the direction of this bias cannot be predicted without an empirical analysis of the insurance programme (Meer and Rosen, 2003). Anticipation of relatively high level of utilisation in the future may incline an individual to seek health insurance, which will lead to an upward bias in the estimated association between insurance and health care utilisation. On the other hand, insurers may introduce procedures that would cream skim only the most profitable low-risk clients and decline insurance to individuals with high expected utilisation (Sekhri and Savedoff, 2005; Osei-Akoto, 2003). In such scenario, due to low levels of utilisation and expenditure among the insured compared to the wider population, the estimated impact of health insurance on cost of healthcare may be biased downwards.

Hadley (2003) reviews the literature on the role of health insurance in improving access and utilisation of health care and notes the scarcity of research that adjusts for insurance endogeneity bias. In recent years, however, there are more studies considering endogeneity bias, though some clearly face limitations due to data rather than technique (Schneider and Hanson, 2006). Ekman (2007) uses national household survey data from Jordan and applies a two-part model to analyse the impact of health insurance on health care utilisation and expenditure. He found that the effect of health insurance on improving access to health care and reducing healthcare expenditure became more pronounced after correcting for endogeneity bias. Dor et al (2006) used data from US national health and retirement study to find a six-fold increase in insurance effect on health scores after adjusting for insurance endogeneity. Similarly, Dalen (2006) uses Chinese health and nutrition to study the effect of health insurance membership on demand for health care. After correcting for endogeneity bias using instrumental variable method, they find that insurance does not seem to have a significant effect on health care utilisation and expenditure. This is partly attributed to high levels of out-of-pocket payments made by both insured and uninsured. Finally, Waters (1999) uses data from Ecuador Living Standards Measurement Survey 1995 to evaluate the impact of health insurance on health care utilisation. After correcting for insurance endogeneity bias, Waters found that the General Health Insurance programme in Ecuador has a strongly positive association with curative health care but has no significant effect on the use of preventive care.

In econometric terms, both care-seeking and insurance-seeking self-selection biases occur due to a non-zero correlation between a choice-based variable and the outcome variable. This would produce biased and inconsistent estimates of the population parameters. In other words, as sample size increases, regression coefficient $\hat{\beta}_1$ does not converge in probability to the true value of population parameter β_1. The bias is equal to $\rho_{X\varepsilon}\ \sigma_\varepsilon/\sigma_X$, where $\rho_{X\varepsilon}$ is correlation between the error terms of choice variable and the outcome variable, σ_ε is the variance of the error term and σ_X is the variance of endogenous variable. Models to adjust for self-selection biases will be discussed later.

THE VIETNAMESE CONTEXT

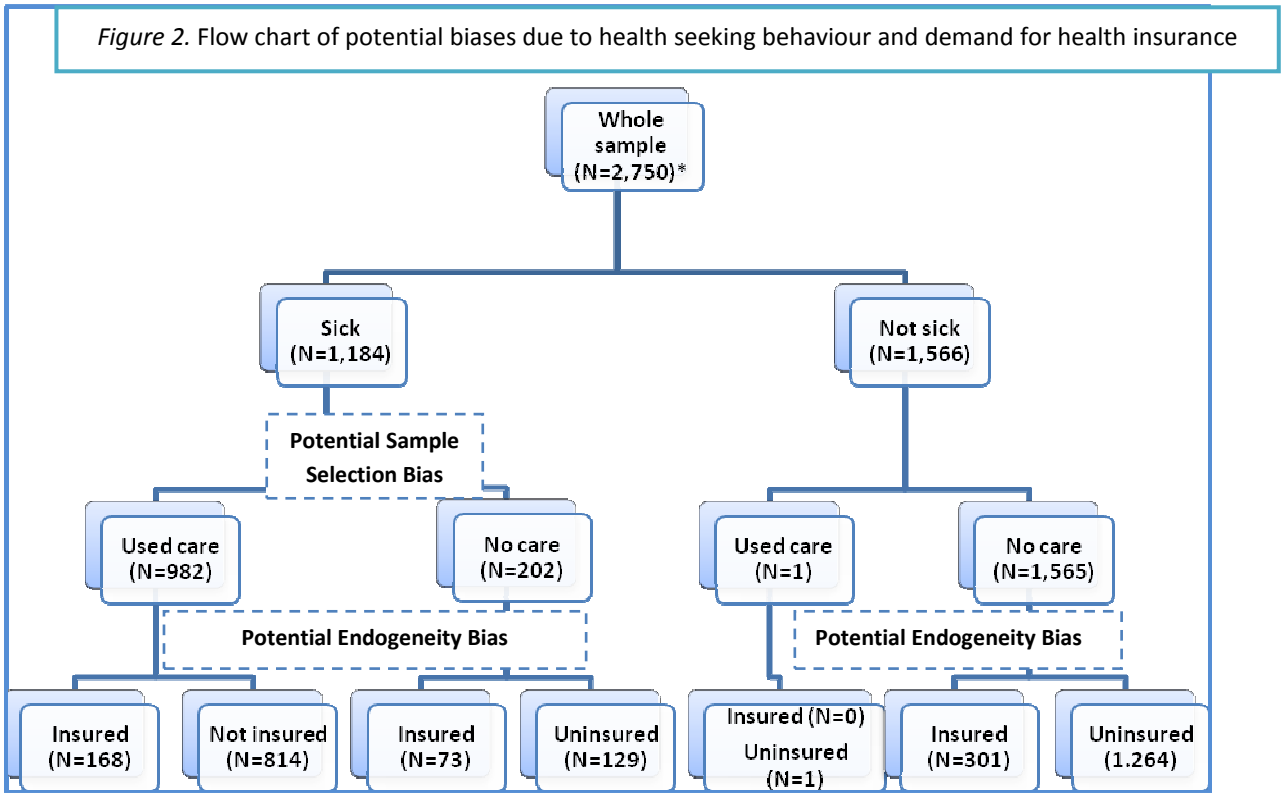
Vietnamese Health Insurance, was formally launched in August 1992, following the issuing of the National Health Insurance Decree (Ministerial Decision No. 299/HDBT), which called for compulsory health insurance for salaried workers in both public and private sectors. Compulsory insurance is exclusively for state employees and employees of large organisations. All those individuals not eligible for compulsory health insurance are encouraged to insure themselves on a voluntary basis (Ensor 1995; Jowett and Thompson 1999). In practice, membership of voluntary VHI falls into three distinct sub-groups: i) schoolchildren and students; ii) members of households eligible for humanitarian assistance to whom insurance is freely provided; and iii) others, including the self-employed (including farmers and service workers), employees of small enterprises, and dependants of those insured under the compulsory scheme. Figure 1 summarises the structure of Vietnamese health insurance in terms of target groups.

Health insurance essentially offers financial benefits (i.e. an 80% reduction in user charges), rather than coverage for additional health services, or service provision in a separate facility. These benefits are offered at the time of service and no reclaims or reimbursements are involved. Health facilities are reimbursed by provincial health insurance funds for the services they provide to insured patients. Reimbursements to the health provider are made on a fee-for-service basis, which generally includes a flat fee for accommodation, and a standard fee for medical tests and procedures. Premia are fixed within provinces, but vary between them, and are typically in the region of US$ 2-10 for an annual policy. Under the initial decree of 1992, benefits under insurance were defined as including medical examinations, diagnostic tests, drugs, and diagnostic and treatment operations. At the time of purchasing an insurance policy, individuals must designate one public health facility at which benefits can be obtained which, in most cases, is a District Hospital. Members were thus unable, at the time of the research, to use the insurance policy at private or non-designated public facilities.

DATA AND METHODS

Survey was carried out in year 1999 in 3 provinces of Vietnam, namely Hai Phong (North east), Ninh Binh (North east) and Dong Thap (South West). Selection of provinces took into account the difference between the north and south of the country along with the level of membership in the programme. Within each province, one urban and two rural districts were randomly selected, followed by random selection of three communes within each district. Finally, interviews were conducted in 27 communes in 9 districts of 3 provinces. Within communes, members of voluntary health insurance were randomly selected from lists supplied by Provincial Health Insurance offices. Uninsured individuals were selected from lists of residents in each commune provided by local People's Commune Committee. Data was collected through one-to-one questionnaire-based interviews conducted by trained interviewers. Since the focus of the study is on the impact of voluntary health insurance, members of compulsory or humanitarian-based health insurance programme were excluded from the study. A total of 1,650 adults and 1,101 children were interviewed of which 19% were residents of Ninh Binh, 40% of Hai Phong and 41% of Dong Thap.

Figure 2. Flow chart of potential biases due to health seeking behaviour and demand for health insurance



Survey collected data on personal characteristics (age, gender, residence, years of schooling, health insurance status etc.), socioeconomic variables (annual household expenditure and income, occupation etc.), and self-assessed health variables. Socioeconomic status of the respondent was measured using annual household expenditure in 12 months prior to the interview as a proxy. Using 'Permanent Income Theory' attributed to Friedmen (1957), it is argued that households smooth their consumption by accumulating or de-cumulating physical buffer stocks that may include farm animals, grains or jewellery. Therefore consumption expenditure is a better estimate of socioeconomic status than income itself (Conning and Udry, 2005). Expenditure data was adjusted for the heterogeneity in household size by using equivalence scale proposed by Aronson et al (1994):

$$e_i h = (A_i h + \emptyset K_i h))^{\theta}$$,

where e_h is the equivalence factor for household h, A_h is the number of adults in the household h and K_h is the number of children. Based on Wagstaff et al (1999), we have set the two unknown parameters $\emptyset$ and θ equal to 0.5.

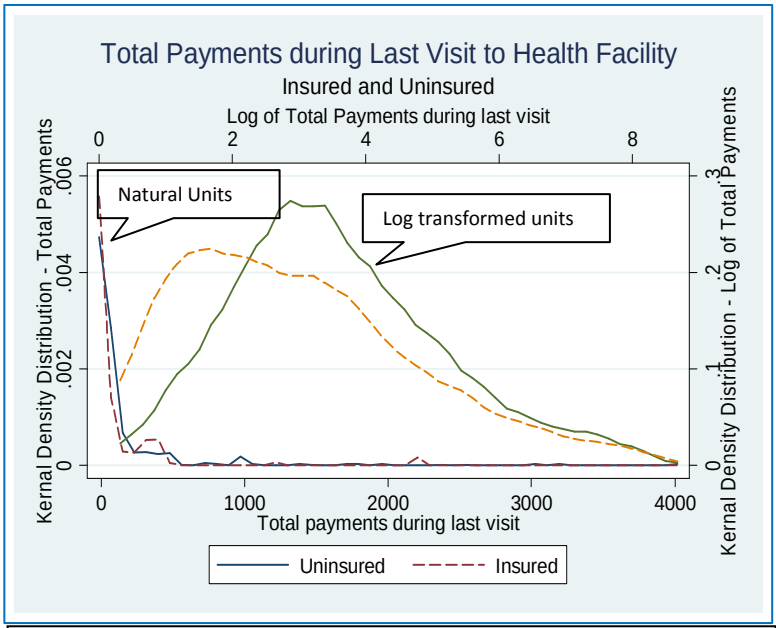


Figure 3. Distribution of Total health care expenditure during last two months (both natural units and logarithms)

Data on illness, health care utilisation and expenditure was collected on an individual level for the last visit/s made within three months prior to the interview. From a total sample of 2,751 interviewees, less than half had fallen sick in three months prior to the interview. While a subsample of individuals did not seek medical advice despite their illness, 982 individuals did seek care and provided details of the total expenditure incurred during their last visit/s to the health facility. Figure 2 identifies the potential biases attributed to selection decisions.

Respondents were asked to recall both direct payments (i.e. user fees for consultations, diagnostic tests and medicines) and indirect costs (like travel costs), as well any unofficial payments made. Figure 3 shows the distribution of payments among insured and uninsured respondents. Insured individuals were also asked about the payment they made in the form of insurance premium. However, there were substantial non-responses on this question, possibly for the reason that many individuals purchased their policy several months before the survey. Given the difficulty in accurately establishing the premium paid, the amount is not included in estimations of health expenditures for the insured. Whilst a downward bias in estimates will result, the underestimation is unlikely to be substantial, given the low level of premia relative to average health expenditures amongst insured patients.

Descriptive analysis

A summary of the descriptive analysis based on illness in last three months is presented in table 1.

Table 1. Results from the descriptive analysis of variables of interest based on illness in last three months, health seeking behaviour and insurance participation

	Whole sample	Respondents wl were sick in last months (N=1,18	Sick respondents w sought healthcar (N=982)		
Variable Name	**Mean**	**Mean**	**Mean**	**Min**	**Max**
Log of equivalent annual household expenditure	8.657249	8.686742	8.691947	6.08221	11.239
Age	29.58836	33.02025	33.10478	6	91
Age-squared	1229.6	1519.169	1538.163	36	8281
Female	0.488	0.5265823	0.5452696	0	1
Rural resident	0.6665455	0.7172996	0.7222787	0	1
Resident of Hai Phong	0.1963636	0.1206751	0.0854527	0	1
Resident of Ninh Binh	0.3923636	0.3991561	0.4191251	0	1
Occupation - service	0.1072727	0.0987342	0.1007121	0	1
Occupation - farmer	0.2661818	0.2877637	0.2899288	0	1
Occupation - government employee	0.0970909	0.1097046	0.0966429	0	1
Occupation - student	0.372	0.3316456	0.3306205	0	1
Occupation - retired	0.0469091	0.0599156	0.0600203	0	1
Number of years of schooli	6.794909	6.439662	6.228891	0	16
Health status - good or fairl good	0.8563636	0.7409283	0.7375381	0	1
Member of Voluntary Healt Insurance	0.1970909	0.2033755	0.1709054	0	1

Table 2 below compares the average out-of-pocket health care expenditure between insured and uninsured using consumption quintiles.

Table 2. **Average out of pocket expenditure during last three months (based on consumption quintiles)**

Consumption Quinti	Insured ('000 VND)	Uninsured ('000 VND)	Total ('000 VND)
Quintile 1 (Poorest)	357.05	132.46	230.33
Quintile 2	328.98	120.25	190.08
Quintile 3	66.36	177.08	179.91
Quintile 4	186.11	240.89	213.79
Quintile 5	292.92	270.79	241.06

Average	264.98	222.79	301.47

The above analysis shows that in the middle two quintiles insured on an average spent substantially more than the uninsured, while the differences between the two groups is not prominent in the poorest and richest quintiles. The average difference between the groups across five quintiles is not very huge. Jowett et al (2003) conducted similar analysis using the same dataset and found different results from us, i.e. the insured in their analysis on an average spent less than uninsured. The difference in the findings is attributed to difference in the equivalence scales used to account for the heterogeneity in household sizes. Our understanding is that Jowett et al (2003) have calculated equivalent per capita annual consumption expenditure by giving equal consumption weight to children and adults in the household, whereas we have followed the approach suggested by Aronson et al (1994) and Wagstaff et al (1999).

One possible explanation of higher average health care expenditure among the insured is that they may have higher health care needs than the uninsured as partly evident from the descriptive analysis discussed above. In order to further understand that impact of health insurance on health care expenditure, we use econometric models that would attempt to control for the effects of covariates. It is important to mention here that in our study the ordinary regression methods may suffer from endogeneity and sample selection bias. We use econometric models to test and correct for endogeneity and sample selection biases, if found. These models are discussed in detail in the following section.

ECONOMETRIC MODELS

We use econometric models to estimate and correct for care-seeking and insurance-seeking self-selection biases.

Model 1 (Ordinary Least Squares model for cost of healthcare)

The general OLS model for cost of healthcare can be expressed as a function of insurance and a set of exogenous explanatory variables:

$$Y_i = \alpha_r + \beta_r x_i^{'} + \delta_r M_i + \varepsilon_{ir} \qquad (1)$$

where, $Y_i =$ cost of seeking healthcare faced by sick individuals

$x_i^{'} =$ **[vector of exogenous variables]**, and

M_i = binary variable for health insurance membership status

As discussed earlier, model 1 can suffer from self-selection biases. The first selection bias is attributed to care-seeking decision as cost of healthcare is only observed if an individual decides to seek care given his/her illness and other exogenous factors. Therefore,

Y_i *is observed if* $z_i = 1$

Y_i *is not observed if* $z_i = 0$

where, z_i = binary variable for care seeking given illness

If this realisation of the cost of seeking healthcare is distributed in a non-random fashion due to the unobserved factors that may influence both the decision to seek healthcare ($z_1 i$) and the cost of healthcare ($Y_1 i$), then equation (1) will suffer from self-selection bias attributed to non-random care-seeking behaviour. This care-seeking decision can be estimated using a probit function which models the

latent tendency or propensity of z_i to be equal to one. This tendency, denoted by z_i^*, is a continuous function represented in equation (2).

$$z_i^* = \alpha_z + w_i'\beta_z + M_i\delta_z + \varepsilon_{iz} \qquad (2)$$

where, w_i' = [vector of exogenous variables] and

M_i = Membership status of voluntary health insurance programme

Therefore,

$$z_i = 0 \; if \; Z_i^* \leq 0 \;\; \rightarrow Y_i = 0$$

$$z_i = 1 \; if \; Z_i^* > 0 \rightarrow Y_i = Y_i$$

The second selection bias occurs due to the endogenous behaviour of insurance membership variable M_i which affects both care-seeking behaviour and cost of healthcare equations. Since insurance-seeking decision is also a binary variable, the propensity to purchase insurance can be modelled as a continuous variable using a probit model:

$$M_i^* = \alpha_m + v_i'\beta_m + \varepsilon_{im} \qquad (3)$$

Therefore,

$$M_i = 0 \; if \; M^* \leq 0;$$

$$M_i = 1 \; if \; M^* > 0$$

where, v_i' = [vector of exogenous variables]

The two self-selection biases can be tested and corrected simultaneously using sample selection models proposed by Heckman (1974, 1979). Here we propose four econometric models besides an OLS model (model 1) to get unbiased estimates of the impact of health insurance on cost of seeking healthcare (see figure 5).

Model 2 (Bivariate probit followed by Heckman's Treatment Effects model) and Model 3 (Independent corrected probit followed by Heckman's Treatment Effects model)

As discussed earlier, model for cost of healthcare may be biased due to care-seeking self-selection and/or insurance endogeneity, while care-seeking self-selection may itself be biased due to insurance self-selection. To estimate and correct for these biases, we use two part models (models 2 and 3 both have two parts). The first part of these models estimates care-seeking self-selection decision with correction for insurance endogeneity bias. This is done using a bivariate probit function (model 2) or two independent probit schedules for care-seeking and insurance-seeking decisions (model 3). Following Heckman (1979), we use the outcome of this first part to generate instantaneous probability of not seeking care (IMR or λ_c) to be employed in the second part of the model (Heckman's treatment effects) which estimates cost of healthcare. Inclusion of λ_c in the second stage corrects for care-seeking self-selection bias, when present. Heckman model itself is a two step model, with first step being a probit function for insurance-seeking decision and second step modelling the final outcome of interest, i.e. cost of seeking healthcare. First step of Heckman model generates a separate IMR (λ_h) for non-participation in insurance programme, which is included in the outcome step which also includes λ_c previously estimated. By including the two IMRs (λ_c and λ_h) in the outcome step of Heckman's model, we simultaneously control for care-seeking and insurance-seeking self-selection biases. The two parts of models 2 and 3 are discussed in detail below.

Part 1 – model for care-seeking decision with correction for insurance endogeneity: Models 2 and 3 differ from each other in part 1. First part of model 2 is a bivariate probit function (equation 4), which simultaneously estimates the two binary decisions of care-seeking and insurance-seeking, in turn allowing for a correlation ρ (rho) between the error terms (ε_m and ε_z) of equations (2) and (3) (Jones, 2007).

$$\left.\begin{array}{c} z_i^* = \alpha_z + w_i'\beta_z + M_i\delta_z + \varepsilon_{iz} \\ M_i^* = \alpha_m + v_i'\beta_m + \varepsilon_{im} \\ \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \sim N\left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right\} \end{array}\right\} \quad (4)$$

If $\rho \neq 0$, then the two probit equations are not independent of each other and the bivariate probit stage is appropriate to model the two decisions as one system of equations. If ρ is not significantly different from zero, then the two choice-based decisions can be modelled independently. If this is the case, then we use model 3 that starts with independent estimation of insurance-seeking decision (equation 3) which is subsequently used to generate an inverse mills ratio (IMR or λ_h) to be employed separately in care-seeking probit model as an additional regressor to correct for insurance endogeneity in care-seeking decision. IMR for health insurance (λ_h) is the instantaneous probability of not purchasing insurance given health status and other explanatory factors and can be expressed as:

$$\lambda_h = \frac{f(\theta_h)}{1 - F(\theta_h)} \quad (5)$$

where $f(\theta_h)$ and $F(\theta_h)$ are probability density and cumulative density functions of (θ_h) respectively, and θ_h is the inverse of the predicted value of purchasing health insurance (equation 6) obtained from equation (3).

$$\theta_h = -\frac{(\alpha_m + v_i'\beta_m)}{(\varepsilon_{bi})^{1/2}} \quad (6)$$

Therefore, the first part of both models 2 and 3 estimate care-seeking self-selection decision with correction for insurance endogeneity bias either through a bivariate probit or through independent probit models. The final outcome from the first stage of models 2 and 3 is a separate IMR (λ_c), as mentioned earlier, to be used in the second part of the model. Here λ_c is the instantaneous probability of not seeking care (corrected for insurance endogeneity), given illness and other exogenous factors and can be expressed as:

$$\lambda_c = \frac{f(\theta_c)}{1 - F(\theta_c)},$$

where θ_c is the inverse probability of seeking care given illness, insurance status and exogenous factors.

Second part – modelling cost of healthcare with correction for care-seeking and insurance-seeking self-selection biases: Second part of models 2 and 3 is the same, i.e. Heckman's treatment effects model, which itself is a two-step model. First step of the model is a probit function that models insurance-seeking decision (equation 3) and predicts IMR for insurance self-selection ($\lambda_1 h$). The second step in Heckman's model is the outcome schedule that models cost of healthcare using λ_h for insurance-seeking self-selection (from first step of Heckman's model), λ_c for care-seeking self-selection (generated in part 1 of the model) and a set of exogenous variables. Hence, by including λ_c and λ_h, Heckman's treatment effects model estimates the impact of insurance membership on cost of seeking healthcare after simultaneously correcting for both care-seeking and insurance-seeking self-selection biases (equation 7).

$$Y_i = \alpha_t + x_i'\beta_t + \lambda_{ci}\,\tau_t + \lambda_{ih}\delta_t + \varepsilon_{it} \qquad (7)$$

where subscript *t* is used for Heckman treatment effects model and subscript *i* is for individual.

Model 4 (Independent uncorrected probit followed by Heckman's Treatment Effects model) and Model 5 (Heckman's Treatment Effects model only)

Model 4 is used if bivariate probit and independent corrected probit estimates in model 2 and 3 suggest that insurance endogeneity is not a source of bias when modelling care-seeking decision, i.e. model 2 has statistically non-significant rho and model 3 shows a non-significant λ_h (lambda). Therefore, in the first stage of model 4, care-seeking decision is now modelled without correction for insurance self-selection (same as equation 2). The outcome of the first stage of model 4 is an IMR (λ_p), which is the similar to λ_c, with the difference that λ_p is calculated after an uncorrected probit rather than insurance corrected as in models 2 and 3.

This λ_p is then used in the second stage of model 4, i.e. Heckman's treatment effects model, which estimates λ_h and uses both λ_p and λ_h lambdas in the outcome schedule (equation 8).

$$Y_i = \alpha_u + x_i'\beta_u + \lambda_p\delta_u + \lambda_h\delta_u + \varepsilon_{ui} \qquad (8)$$

The difference between equation (7) and (8) is that the former estimates λ for care-seeking using a bivariate probit or independent corrected probit, while latter uses uncorrected probit for its estimation.

Model 5 only corrects for insurance-seeking self-selection bias in estimation of cost of healthcare, and ignores care-seeking selection bias. This model is the same as the second stage of models 2 to 4, without any correction for care-seeking selection bias (equation 9).

$$Y_i = \alpha_s + x_i'\beta_s + \lambda_h\delta_s + \varepsilon_{si} \qquad (9)$$

Table 3: Econometric models for estimation of cost of healthcare function

Models/Stages	Model 1	Model 2	Model 3	Model 4	Model 5
Part 1	OLS model for cost of healthcare without correction of self-selection bias	Bivariate Probit model for care-seeking and insurance-seeking self-selection decisions	Independent Probit model for insurance-seeking self-selection generates IMR ($\lambda_1 h$) for non-participation in insurance programme → λ_h used in separate probit model for care-seeking self-selection	Independent Probit for care-seeking self-selection decision without correction for insurance endogeneity	Heckman's Treatment effects model for cost of healthcare, correcting for insurance-seeking bias only
Outcome from part 1	-	Inverse Mills Ratio (IMR - λ_c) for non-participation in care-seeking	IMR ($\lambda_1 c$) for non-participation in care-seeking generated from part 1, to be used in	IMR ($\lambda_1 c$) for non-participation in care-seeking generated from	-

		generated from part 1, to be used in part 2 of the model	part 2 of the model	part 1, to be used in part 2 of the model	
Part 2		Heckman's Treatment effects model for cost of healthcare. Outcome part of Heckman uses λ_c from part 1 and λ_h generated within Heckman's model	Heckman's Treatment effects model for cost of healthcare. Outcome part of Heckman uses λ_c from part 1 and λ_h generated within Heckman's model	Heckman's Treatment effects model for cost of healthcare. Outcome part of Heckman uses λ_c from part 1 and λ_h generated within Heckman's model	

Identification of self-selection decisions using instrumental variables

Heckman's treatment effects model requires us to use an instrumental variable (IV) that is uniquely associated with decision to self-select into insurance but is not associated with decision to seek care or cost of healthcare. In the models discussed above, we use a binary variable for 'medium to high level of worry about future health' as IV. From Waters (1999), it follows that the appropriateness of an instrumental variable can be tested using three methods. 1) IV should have a significant non-zero coefficient when regressed on the suspected endogenous variable, 2) IV should not have a significant non-zero coefficient when regressed against the cost of seeking healthcare, and finally 3) Likelihood ratio test should show a greater log likelihood value when IV is substituted for the endogenous variable in cost of healthcare equation, compared with when the predicted value of the endogenous variable (calculated with the same instrumental variable) is used in cost of healthcare function. Our proposed instrumental variable fulfils all three tests and is therefore an appropriate candidate to be an instrumental variable.

Heckman's sample selection model can also rely on identification using instrumental variables but in practice it is often difficult to find such explanatory factors that uniquely determine care-seeking selection. Therefore sample selection model is often estimated using the same regressors as in both parts of the model. Therefore, identification relies on the non-linearity of IMR (Jones, 2007).

It should also be mentioned that since insurance was purchased more than three months ago and cost of healthcare was incurred within last three months, certain variables are naturally unique in cost of healthcare equation due to the difference in the time of event occurrence. Therefore variables like hospital inpatient stay (binary) and number of illnesses within last three months are not used when modelling insurance-seeking decision.

RESULTS

This section presents and compares findings of the econometric models discussed in the previous section. Analysis was carried out using STATA software, version 9.2. Unit of analysis was an individual for whom the questionnaire was completed. Socioeconomic status was captured using log of annual equivalent household expenditure. Need for healthcare was recorded using self-assessed health indicator

and a variable for 'number of illnesses in last three months'. Utilisation was measured using 'number of health visits' and 'whether or not hospital admission was required'.

OLS analysis is carried out using only the individuals who were sick and had non-missing values on cost of healthcare in last three months. Therefore eligible observations were 982, but 22 were dropped from regression due to missing values on explanatory variables. The OLS model passed Ramsey RESET test with test score F (3, 938) = 0.98 and p > F = 0.40. The explanatory power of the model, i.e. R^2, is 0.34. OLS results show a significant negative effect of insurance membership on log of healthcare expenditure, which implies that compared to the uninsured individuals, insured spent significantly less given their level of need and utilisation. Regression results also show that socioeconomic status of an individual is positively related to their cost of seeking healthcare, after controlling for observed need and utilisation. Cost of healthcare was also highly sensitive to inpatient admission, which is hardly surprising. Residuals from the OLS regression were found to be normally distributed (figure 4). This was further confirmed using skewness test in STATA.

OLS results may suffer from care-seeking and insurance-seeking self-selection biases. Also care-seeking decision may suffer from endogeneity bias due to insurance self-selection. Subsequent models estimate and correct for these biases. Model 2 estimates insurance endogeneity effect on care-seeking decision in the bivariate probit stage and finds that the estimated correlation coefficient (rho=0.57) of the two decisions is not statistically significant and the Wald test for null hypothesis of rho=0 can not be rejected. Model 3 estimates an independent probit function for insurance-seeking decision and uses IMR (λ_h) from this probit model as an omitted variable in the probit model for care-seeking decision. The results show that the coefficient on λ_h generated from insurance-seeking decision model is not statistically significant, suggesting that insurance endogeneity is not a problem in care-seeking decision model (table 5). Therefore, we use model 4, which estimates probit function for care-seeking decision without correcting for insurance endogeneity bias. All models of care-seeking decision, with or without correction for endogeneity bias, suggest that socioeconomic status is not a significant variable in the decision to seek care given illness. Models also show that individuals who are generally in better long-term health are less likely to seek care when ill. Care-seeking models also suggest that residents of Hai Phong and Ninh Binh provinces are less likely to seek care when ill compared to the reference province Dong Thap.

We generate inverse mills ratio λ_c for care-seeking self-selection decision in models 2, 3 and 4, which is subsequently used in the second part of these models, i.e. Heckman's treatment effects regression. As discussed earlier, Heckman's model is itself a two-step model, with first step being a probit equation for insurance self-selection decision which generates an IMR λ_h for insurance decision subsequently used in the second step of Heckman model alongside λ_c generated in first part of models 2, 3 and 4. Coefficient on λ_h in the second step of Heckman model is significantly positive, suggesting that insurance-seeking self-selection bias does exists in the estimation of cost of healthcare model (table 6). Including λ_h as an additional regressor in the second step corrects for this bias.

All three models (2, 3 and 4) showed statistically significant coefficients on λ_c, suggesting strong evidence of care-seeking self-selection bias in the analysis of cost of healthcare faced by individuals (table 6). From these results, one can infer that the observed cost of seeking healthcare is biased due to non-random care-seeking decision of individuals, which was subsequently corrected in our models 2, 3 and 4 along with correction of insurance-seeking self-selection bias. Negative sign of the coefficient on λ_c suggests that higher probability of non-selection into care-seeking is associated with lower cost of healthcare.

The final model used in our analysis is Heckman's treatment effects model which estimates and corrects only for insurance-seeking self-selection bias in cost of healthcare model. Model 5 has been used in some studies recently, but it ignores care-seeking self-selection bias. Estimates from the model show that IMR

$\hat{E}(\lambda)_h$, which estimates insurance self-selection bias in the first step of Heckman's model, is statistically significant, suggesting evidence of insurance self-selection bias even without simultaneous correction of care-seeking selection.

Table 6 shows that the size of impact of membership in voluntary health insurance on cost of healthcare more than doubled after care-seeking and insurance-seeking self-selection biases were corrected. This shows that the OLS model had underestimated the impact of insurance membership on cost of seeking healthcare. Table 6 also shows that the positive association between socioeconomic status (i.e. log of equivalent annual household expenditure) and cost of care increases only marginally in its magnitude after correction for self-selection. It is also observed that residents of Hai Phong are likely to spend than the residents of Ninh Binh and Dong Thap provinces, given their need and level of utilisation. Magnitude of the coefficients on health status and inpatient admission also reduced marginally, though their qualitative effects remained the same.

From the results, it was also revealing to note that insurance membership does not clearly seem to influence the decision to seek care given illness, though it significantly reduces the cost of care once treatment is sought. We also observed that socioeconomic status was not significantly associated with decision to purchase insurance, whereas rural residence only showed positive association with insurance seeking in model 3. Our analysis also shows that years of schooling is an important predictor of insurance-seeking behaviour.

DISCUSSION OF RESULTS

In the paper we analysed the impact of voluntary health insurance programme in Vietnam after correcting for care-seeking and insurance-seeking self-selection biases. After applying econometric models to individual level survey data from Vietnam, the study revealed that, given the level of need and utilisation, health insurance has a significantly negative relationship with cost of healthcare. Though descriptive analysis in table 2 suggest that, on an average, insured individuals tend to spend more on health care than uninsured in the sample. But by using econometric models we were able to control for observed and unobserved factors to suggest the strong negative relationship. The apparent higher average expenditure among the insured can be attributed to adverse selection into insurance of those with greater health care needs, which is a common observation across developing countries (Liu, 2004; Ranson, 2002). The study also shows that the actual impact of insurance membership on the cost of healthcare faced by individuals is significantly higher after correction of self-selection biases. This is due to non-random distribution of care-seeking and insurance-seeking behaviour which directly influences the cost of healthcare.

While policy makers in developing countries face with the challenge of establishing most equitable and accessible health care system, analytical studies on health financing options are crucial in providing evidence for policy making. Our empirical analysis shows that though health insurance reduces average health care expenditure, the uptake of insurance is not significantly associated with socioeconomic status of an individual. If the goal of the health care system is to provide affordable health care with direct targeting of those who are economically marginalised, then policy makers will have to do more to improve access to health insurance programme to the poor.

Acknowledgements

The authors are grateful to Mark Dusheiko, University of York, who provided his invaluable time to discuss econometric models used in the study. We would also like to thank Matthew Jowett, who kindly allowed us to use the data used in this study.

REFERENCES:

Ahuja, R and Jutting, J. 2004. 'Are the Poor Too Poor to Demand Health Insurance?' *Journal of Microfinance* 6(1)

Aronson, J.R., Johnson, P., Lambert, P.J., (1994). 'Redistributive effect and unequal tax treatment'. *Economic Journal*, 104, 262–270.

Avi, D., Joseph, S. and David, W. B.(2006). 'The Effect of Private Insurance on the Health of Older, Working Age Adults: Evidence from the Health and Retirement Study'. *Health Services Research*, 41 (3p1), 759–787.

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* 2007, **5:**6

Ensor, T. (1995). 'Introducing health insurance in Vietnam'. *Health Policy and Planning*, 10(2); 154-163.

Ensor, T., and San, P. (1996). 'Access and payment for health by the poor in Northern Vietnam'. *International Journal of Health Planning and Management*, 11(1); 69-84. Jowett and Thompson 1999

Fabricant, S.J., Kamara, C.W., and Mills, A. (1999). 'Why the poor pay more: household curative expenditures in rural Sierra Leone'. Int J Health Plann Manag *14:* 179.199

Friedman, M.(1956). 'The Quantity Theory of Money—a Restatement.' In Studies in the Quantity Theory of Money, edited by M. Friedman, Chicago: *Univ. Chicago Press*

Gronau, R. (1974). 'Wage comparisons: a selectivity bias'. *Journal of Political Economy*, 82: 1119-1155.

Gruber, J.(2000). 'Medicaid.' *NBER*, Working Paper, 7829

Hadley, J. (2003). 'Sicker and Poorer--The Consequences of Being Uninsured: A Review of the Research on the Relationship between Health Insurance, Medical Care Use, Health, Work, and Income.' *Medical Care Research and Review*, 60 (2): 3-76.

Heckman, J.J. (1974). 'Shadow prices market wages and labor supply'. *Econometrica*, 42: 679-694

Heckman, J.J. (1976). 'The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator of such models'. *The Annals of Economic and Social Measurement* 5:475-92.

Heckman, J.J. (1979). 'Sample Selection Bias as a Specification Error'. *Econometrica* 47 (1):153-161.
Heckman, J.J. (1990). 'Varieties of selection bias'. *American Economics Review*, 80:313-18.
Hendrik, P., van Dalen (2006). 'When Health Care Insurance Does Not Make A Difference – The Case of Health Care; 'Made in China' TI 2006-091/1 Tinbergen Institute Discussion Paper

Isaac, O. Demand for voluntary health insurance by the poor in developing countries: Evidence from rural Ghana Center for Development Research (ZEFb) http://www.chronicpoverty.org/pdfs/2003conferencepapers/Osei-Akoto.pdf

Jonathan, M., and Harvey, S. R. (2003). 'Insurance and the utilization of medical services'. *National Bureau of Economic Research.* Working Paper 9812

Jowett, M., Cantoyannis ,P., and Vinh, N.D. (2003). 'The impact of voluntary health insurance on private health expenditure in Vietnam'. *Social Science and Medicine*, Jan 56(2): 333-42

Jutting, J. (2003).' Do community-based health insurance schemes improve poor people's access to health care? Evidence from Senegal'. *OECD developmen*t, Paris.

Neelam, S. and William, S. (2005). 'Private health insurance: implications for developing countries'. *Bulletin of World health Organisation*, 2005; 83: 127-134.

O'Donnell, O., F. Rosati, and E Van Doorslaer. (2005). 'Health effects of child work: evidence from rural Vietnam'. *Journal of Population Economics*, 18: 1-31.

Palmer N., Mueller D. H., Gilson, L., Mills, A., & Haines, A. (2004). 'Health financing to promote access in low income settings – how much do we know?'. *The Lancet*, 364, 1365–1370.

Pia, S. and Kara, H. (2006). 'Horizontal equity in utilisation of care and fairness of health financing: a comparison of micro-health insurance and user fees in Rwanda'. *Health Econ*, 15: 19–31

Ramesh B. and Nishant, J. (2006-07). 'Factoring affecting the demand for health insurance in a micro-insurance scheme'. *Indian Institute of Management,* W.P. No. 2006-07-02

Ranson, M. K. (2002). 'Reduction of catastrophic health care expenditures by a community-based health insurance scheme in Gujarat, India: current experiences and challenges'. *Bull World Health Organ,* vol.80 no.8 Genebra Aug. 2002

Smith, P., and Witter, S. (2001). 'Risk Pooling in Health Care Finance. Report Prepared for the World bank Workshop Resource Allocation and Purchasing in health: Value for Money, Reaching the Poor.' York/Washington D.C.

Solon, O., & Tien, T. (1997). The Challenges of the Vietnam Health Insurance Programme

Tuma, Nancy B. (1982) 'Nonparametric and partially parametric approaches to event-history analysis.' in Samuel Leinhardt (ed.), *Sociological Methodology,* San Francisco.
UNDP. (2006). *Human Development Report,* 2006

Waddington, C.J., Enyimayew, K.A. (1989). 'Price to pay: the impact of user charges in Ashanti-Akim District, Ghana'. *International Journal of Health Planning and Management,* Vol. 4 No.1, pp.17-47.

Wagstaff A, van Doorslaer E, van der Burg H, et al.(1999). 'Equity in the finance of health care: some further international comparisons'. *J Health Econ'* 1999;18(3):263-90.

Waters, Hugh (1999). 'Measuring the Impact of Health Insurance with a Correction for Selection Bias – A Case Study of Ecuador'. *Health Economics,* Vol. 8, No. 5, pp. 473-483.

World Bank. (1993). *World development report.* New York: Oxford University Press.

Yip, W., and Berman, P. (2001). 'Targeted health insurance in a low income country and its impact on access and equity in access: Egypt's school health insurance.' *Health Economics,* 10, 207–220.

Yuanli L. 'Development of the rural health insurance system in China'.*Health Policy and Planning*; 19(3): 159–165

APPENDICES

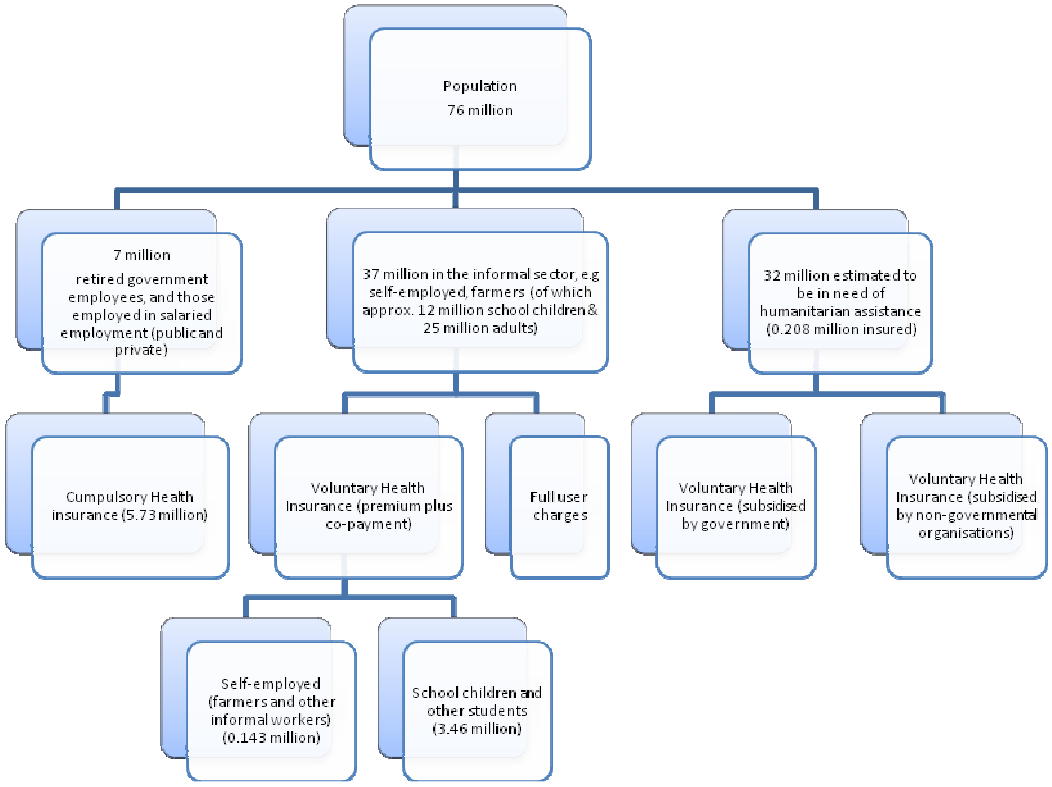


Figure 1: Structure of the Vietnamese Health Insurance programme

Table 4: Econometric models for self-selection into voluntary health insurance programme in Vietnam

Depedent Variable: Whether or not bought insurance membership, given illness	**Model 2 - Biprobit estimation of care-seeking and insurance-seeking decisions**	**Model 3 - independent model of insurance-seeking decision**	**Model 4 - insurance decision modelled in second step of Heckman Treatment effects model**
Log of equivalent annual household expenditur	0.338	0.284	0.162
	(-0.27)	(0.20)	(-0.1)
Age	-0.130**	-0.125**	-0.0165
	(-0.052)	(0.052)	(-0.027)
Age-squared	0.00143***	0.00136***	0.000246
	(-0.00051)	(0.00052)	(-0.0003)
Female	0.19	0.0984	-0.14
	(-0.43)	(0.26)	(-0.13)
Rural resident	0.474	0.580**	0.0966
	(-0.39)	(0.26)	(-0.17)
Resident of Hai Phong	-1.112**	-1.240***	0.883***
	(-0.51)	(0.29)	(-0.18)

Resident of Ninh Binh	-1.343**	-1.382**	-1.461***
	(-0.57)	(0.54)	(-0.2)
Occupation - service	0.514*	0.479	0.913**
	(-0.29)	(0.30)	(-0.38)
Occupation - farmer	0.760**	0.734**	0.672**
	(-0.33)	(0.33)	(-0.33)
Occupation - government employee	2.181***	2.177***	2.354***
	(-0.39)	(0.38)	(-0.35)
Occupation - student	1.551***	1.582***	0.998**
	(-0.44)	(0.45)	(-0.41)
Occupation - retired	0.936*	1.031**	0.649
	(-0.52)	(0.44)	(-0.45)
Number of years of schooling	0.208***	0.214***	0.0899***
	(-0.08)	(0.056)	(-0.024)
Health status - good or fairly good	-0.557	-0.602	-0.114
	(-0.5)	(0.37)	(-0.17)
Medium to high level of worry about future hea of self and family	1.745***	1.772***	0.788***
	(-0.26)	(0.21)	-0.21
Constant	-6.638***	-6.289***	-4.168***
	(-2.55)	(2.13)	-1.04
Observations	1184	1184	960

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5: Econometric models for care-seeking self-selection

Dependent Variable: Whether or not sought care, given illness	**Model 2 - Care-seeking and insurance-seeking biases corrected**	**Model 3 - Care-seeking and insurance-seeking biases corrected**	**Model 4 - Care-seeking and insurance-seeking biases corrected**
Member of Voluntary Health Insurance scheme	-1.345	-0.631	-0.381
	(-2.55)	(-0.77)	(-0.40)
Log of equivalent annual household expenditure	0.126	0.122	0.11
	(-0.12)	(-0.13)	(-0.13)
Age	0.0203	-0.0189	0.0205
	(-0.029)	(-0.029)	(-0.029)
Age-squared	-0.000228	-0.000218	-0.000237
	(-0.00031)	(-0.00031)	(-0.00031)
Female	0.223	0.223	0.216
	(-0.16)	(-0.16)	(-0.16)
Rural resident	0.306	0.292	0.288
	(-0.21)	(-0.2)	(-0.2)
Resident of Hai Phong	-1.288***	-1.262***	-1.235***

	(-0.31)	(-0.29)	(-0.28)
Resident of Ninh Binh	-0.570***	-0.549***	-0.523***
	(-0.2)	(-0.2)	(-0.19)
Occupation - service	0.566	0.577	0.582
	(-0.39)	(-0.39)	(-0.39)
Occupation - farmer	-0.372	-0.36	-0.357
	(-0.26)	(-0.26)	(-0.26)
Occupation - government employee	-0.0876	-0.125	-0.115
	(-0.62)	(-0.63)	(-0.63)
Occupation - student	0.149	0.0992	0.0637
	(-0.4)	(-0.39)	(-0.37)
Occupation - retired	-0.41	-0.417	-0.41
	(-0.32)	(-0.31)	(-0.31)
Number of years of schooling	0.0384	0.033	0.0301
	(-0.035)	(-0.031)	(-0.031)
Health status - good or fairly good	-0.390**	-0.379**	-0.375**
	(-0.18)	(-0.17)	(-0.17)
Mills ratio - for insurance membership	-	0.168	-
	-	-0.3	-
Rho	0.58	-	-
	-1.25	-	-
Constant	0.0256	0.113	0.193
	(-1.21)	(-1.24)	(-1.23)
Observations	1184	1184	1184

Table 6: Econometric models for cost of healthcare faced by individuals

Dependent variable: log of cost of healthcare faced by individual	**Model 1: OLS - no correction for care or insurance-seeking biases**	**Model 2 - Care-seeking and insurance-seeking biases corrected**	**Model 3 - Care-seeking and insurance-seeking biases corrected**	**Model 4 - Care seeking and insurance-seeking biases corrected**	**Model 5: Insurance-seeking bias corrected only**
Member of Voluntary Healt Insurance	-0.786*** (-0.27)	-1.632*** (-0.53)	-1.700*** (-0.53)	-1.719*** (-0.53)	-1.758*** (-0.53)
Log of equivalent annual household expenditure	0.458*** (-0.13)	0.469*** (-0.09)	0.469*** (-0.09)	0.470*** (-0.09)	0.479*** (-0.09)
Age	0.00201 (-0.023)	0.0135 (-0.019)	0.0133 (-0.019)	0.0135 (-0.019)	0.0154 (-0.019)
Age-squared	0.0000329 (-0.00026)	-0.0000983 (-0.00021)	-0.0000958 (-0.00021)	-0.0000978 (-0.00021)	-0.00012 (-0.00021)
Female	0.0512 (-0.14)	-0.028 (-0.1)	-0.0261 (-0.1)	-0.0273 (-0.1)	-0.011 (-0.1)
Rural resident	0.234	0.104	-0.113	0.109	0.125

	(-0.19)	(-0.13)	(-0.13)	(-0.13)	(-0.13)
Resident of Hai Phong	0.344	1.511***	1.460***	1.487***	1.357***
	(-0.26)	(-0.26)	(-0.26)	(-0.26)	(-0.25)
Resident of Ninh Binh	0.206	0.22	0.226	0.223	0.169
	(-0.17)	(-0.16)	(-0.15)	(-0.15)	(-0.15)
Occupation - service	0.336	0.0382	0.0335	0.0361	0.0404
	(-0.28)	(-0.23)	(-0.22)	(-0.23)	(-0.23)
Occupation - farmer	-0.0293	-0.0818	-0.0883	-0.0849	-0.128
	(-0.21)	(-0.18)	(-0.18)	(-0.18)	(-0.18)
Occupation - government employee	-0.352	0.175	0.152	0.169	0.148
	(-0.48)	(-0.39)	(-0.39)	(-0.39)	(-0.39)
Occupation - student	-0.182	-0.107	-0.106	-0.102	-0.111
	(-0.32)	(-0.25)	(-0.25)	(-0.25)	(-0.25)
Occupation - retired	-0.219	-0.196	-0.196	-0.195	-0.247
	(-0.3)	(-0.25)	(-0.25)	(-0.25)	(-0.25)
Number of years of schooli	0.0209	0.00359	0.00425	0.00417	0.00654
	(-0.027)	(-0.02)	(-0.02)	(-0.02)	(-0.02)
Health status - good or fairl good	-0.887***	-0.777***	-0.774***	-0.776***	-0.801***
	(-0.16)	(-0.13)	(-0.13)	(-0.13)	(-0.13)
No. of illnesses in last 3 months	0.0531	-0.0137	-0.0169	-0.0166	0.028
	(-0.064)	(-0.053)	(-0.053)	(-0.053)	(-0.05)
No. of health visits in last 3 months	-0.0431	0.0477	0.0527	0.0519	-0.00793
	(-0.064)	(-0.055)	(-0.055)	(-0.055)	(-0.049)
Used in-patient services	2.578***	2.170***	2.172***	2.173***	2.117***
	(-0.23)	(-0.17)	(-0.17)	(-0.17)	(-0.17)
Mills ratio from care-seekin	-	-0.244**	(-0.265**)	-0.263**	-
decision model	-	(-0.11)	-0.11	(-0.11)	-
Lambda for insurance decis - from Heckman		0.791***	0.753**	0.776***	0.776***
treatment effects model		(-0.3)	(-0.3)	(-0.3)	(-0.3)
Constant	-0.484	-0.466	-0.482	-0.486	-0.546
	(-1.18)	(-0.85)	(-0.85)	(-0.85)	(-0.85)
Observations	960	960	960	960	960

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1