สำนักงานวิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย

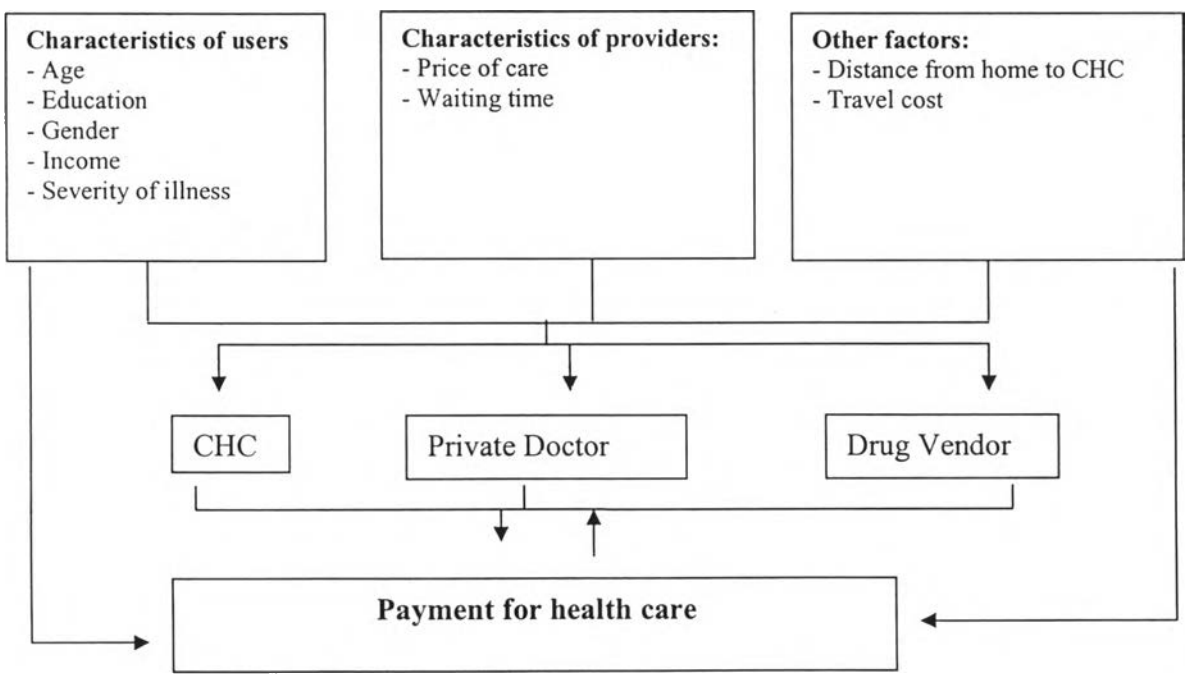
CHAPTER 4
RESEARCH METHODOLOGY

The aim of this study is to analyze the health seeking behavior and payment for health care in Hungha District, Thaibinh province, Vietnam. It compares the payment for heath care between income groups and identifies the main factors that influence the health seeking behavior for each health services and payment for health care among patients in Hungha district, Thaibinh province, Vietnam.

This chapter will construct models to compare the payment for health care and identify the factors influencing the behavior of patient.

4.1 Conceptual Framework

Figure 4.1 Conceptual Framework



From Figure 4.1, factors will be divided in two main groups: factors of patients and factors of providers. Other factors are distance and travel cost. Most factors will affect on health seeking behavior of patients. They can choose one kind of health services (Commune health center, private doctor or drug vender) that are the primary level of health care service. And these factors also affect on payment for health care of patients as well as the kind of health service that patients use.

Dependent variables

Using of health care services: Number of patient who chosen the health care services when they were ill. Health services include: Commune health center (CHC), private doctor, drug vendor or pharmacist

Independent variables

- Education: The number of years that patient got education, calculates in number of years.
- Income: The monthly average income of patient's family
- Waiting time: The time that patient has to wait to examine. Calculate in minutes
- Distance: the distant from patient's home to health care services, calculate in km.
- Payment: amount of money that patient has to pay for health care, including: pay for service, drug and travel cost.

4.2 Data Analysis

4.2.1 Describe the situations of utilization in Commune Health Center, Private Doctor and Drug vendor

Calculate the proportion of patients using each kind of health services and compare the utilization with characteristics of patients: Income groups, education level, severity of illness, age, gender and so on. Mathematical analysis was simple descriptive; such as mean, frequency ... and tables in order to compare among characteristics with Chi-square test ...

4.2.2 Identify the Factors Influencing the Patient Choose each Kind of Health Care Services.

Logit analysis presents a unique complement to multiple regression in its ability to utilize a binary dependent variable. Logit analysis does not predict just whether an event occurred or not, but instead predicts the probability of an event. In this manner, the dependent variable can be any value between one and zero. This means that , the predicted value must be bound to fall within the range of zero and one.

The logit model is based on the cumulative logistic probability function:

$$P_i = f(Y) = f(\beta_0 + \beta_i * X_i) = \frac{1}{1+e^{-y}} = \frac{1}{1+e^{-(\beta 0+\beta i Xi)}} \quad (1)$$

Where

P_i = the probability that an individual will make a certain choice, given knowledge if X_i with $i = 1, 2, \ldots, n$

F = a cumulative probability function.

Xi = the individual's attributes which give information to interpret the dependent variables into the logit model.

β_0 = constant and β_i is coefficient of X_i.

e = exponential, e = 2.7182

From equation (1), another equation can be written:

$(1+e^{-y})\ P_i = 1$

Then, $e^{-y} = \frac{1-p_i}{p_i}$

and $e^{y} = \frac{p_i}{1-p_i}$

Taking logarithm of both sides of above equation,

$$Y = \ln \frac{p_i}{1-p_i}$$

Thus, final equation is:

$$\ln \frac{p_i}{1-p_i} = \beta_0 + \beta_i * X_i$$

Multiple logit regression analysis was applied to identify the factors might influence the patient choose the available health services. The dependent variable was using CHC, Private practitioner, and drug vendor. For example, if the dependent variable was using CHC, patients who used CHC were denoted by 1, and the others was denoted by 0

The explanatory variables were: income, education, distance from patient's home to CHC, waiting time and payment for health care.

The logit regression is:

$$\text{Ln} \frac{p_i}{1-p_i} = \beta_0 + \beta_1 TIME + \beta_2 INCOME + \beta_3 DIST + \beta 4 EDUC + \beta_5 PAY$$

Where TIME = the time that patient had to wait.

INCOME = patient's monthly average income (Currency is VND)

EDUC = number of years that patient got education.

DIST = distance from patient's home to health services, calculate in Km

PAY = the amount of money that patients have to pay for health care

Purpose of this equation was to identify the factor affecting the probability of patient using in kind of services. That means, three equations were run, each equation for one of health services, these were commune health center, private doctor and drug vendor. The meaning of this factor was that, for example, if household monthly income increases, probability of patient using commune health center will increase or decrease. Using this result, policy-

makers might predict the change pattern of health care services in Hungha district in the near future and can set up appropriate policies to improve the real situations of health services.

Reasons for including the above factors as dependent variable were as following

Education level: Higher levels of education may lead to increased efficiency in a family's purchase and use of medical services. A greater among of education in the household may enable a family to recognize the early symptoms of illness, resulting in a greater willingness to seek treatment. Such families are likely to spend more for preventive services and less for more acute illness later. Years of education in a household may be proxy measure for a greater awareness of the need for medical care, for different attitudes toward seeking care, and for greater efficiency in its purchase and production. Different in education among families are expected to result in different in use and expenditure for medical services.

Family income: A number of studies have examined the relationship between family income and expenditures on medical care and also the effect of income on use of medical care. When the studies are based on survey data, it is often found that families with higher incomes have greater expenditure for medical care, although the percentage of income spent on medical care decline as income increases. Other studies also found that, when family income increases, patient prefer using better quality of heath services to lower quality of these. In theory, a good is normal good if people consume it increases when income increases and is inferior good if the number of people using it decreases if income increases.

Price of health care services: The price of a service and the use of that service are, according to economic theory, inversely related; as the price is reduce, purchase or use of the service will increase. Knowledge of price elasticity of demand for medical services is therefore of greater important for public policy.

Waiting time or time cost and distant from patient's home to heath services: When time costs are high, people will substitute purchase services for their own time. Since time has an opportunity costs, it is also scare and should be viewed as one of resource constrains facing

the consumer. If either time or money costs of a service decreases, the quality of demand would be expected to increase. In medical care, time is used in traveling to a provider and in waiting to be treated. Three policies implications follow from the finding that time costs have an important effect on the demand for medical services. First, as out-of-pocket prices to patients decreases, demand for medical care becomes more responsive to the cost of time. Second, society has determined that certain population groups should receive an increase in their use of medical services. Although the money prices to these groups have been reduced, it may be desirable to reduce their time costs to increase further their use of services. Locating clinics closer to population groups will lower their travel costs and increase utilization. Third, when planners determine the number and size of hospitals, the patient's time costs should be considers, together with institutional costs, as the relevant costs for planner to minimize. Consumers are willing to pay higher money prices to decreases their time cost.

Table 4.1 Expected Sign of Coefficients of Logit Regression for each Health Services

Variable	Coefficient		
	Commune health center	Private doctor	Drug vendor
TIME	-	-	-
INCOME	+	+	-
DIST	-	-	-
EDUC	+	+	-
PAY	-	-	-

Test for significant of each factor.

For the model analysis, the hypothesis that a coefficient was different from zero by using Wald test and p value to access the significant of each coefficient. The hypothesis was:

H_0: $\beta_i = 0$

H_1: $\beta_i \neq 0$ i = 1, 2, 3, ...

The meaning of this test is if the null hypothesis H_0 is true, the corresponding independent variable is not related with dependent variable, and its value is useless. On the other hand, if alternative H_1 is true, it implies that there is a relationship between that variable and dependent variable.

To determine significantly, following process was used: Conduct a test with confident significant 95% and calculate the Wald test.

$$z = \frac{\beta_i}{se(\beta_i)}$$ SE is standard error

Reject H_0 if the computed z value is less than critical z value or $p < 0.05$

Accept H_0 if otherwise

Predict a change of probability of independent variable when dependent variable is changing

Suppose, a logit regression is

$$\text{Ln}\frac{p}{1-p} = \beta_0 + \beta_1 X$$

When X increases to X+a, that means X change is ΔX a change of right equation is:

$$\Delta \text{Ln}\frac{p}{1-p} = \beta_1 \Delta X$$

To simplify, we utilize the fact that for any continuous variable X, $\Delta \log X \approx \Delta X/X$, and the fact that log (x/y) = log x - log y. Then

$$\Delta \text{Ln}\frac{p}{1-p} \approx \left(\frac{1}{p} + \frac{1}{1-p}\right)\Delta p = \frac{1}{p(1-p)}\Delta p$$

So we get

$$\frac{1}{p(1-p)}\Delta p = \beta_1 \Delta X \text{ and}$$

$$\mathbf{\Delta p \approx \beta_1 [p(1-p)]\, \Delta X}$$

I 21620295

When X changes, we will calculate a change of p and predict the probability of independent variable.

Applying this theory to this research, for example, we predict the probability of patient using Commune Health Center when income increases. We get a formula

$\Delta p \approx \beta_1[p(1-p)]\ \Delta INCOME$

When INCOME increases 100,000 VND, a change of probability is

$\Delta p \approx \beta_1[p(1-p)]^* 100$

4.2.3 Compare the Payment for Health Care between Income Groups, Kind of Diseases.

The payment for health care of patients will be divided into 3 groups:

Group 1 (Poor group): Monthly average of family income was less and equal to 400,000 VND

Group 2 (Average group): Monthly average of family income was more than 400,000 to less and equal to 700,000 VND

Group 3 (Rich group): Monthly average of family income was more than 700,000 VND

The data was compared the mean of payment for each kind of health services and kind of diseases and also among income group.

4.2.4 Identify the Sources of Finance that Patients Use to Pay for Health Care

The sources of finance that patients use to pay for health care were household saving, borrow money, sell assets, and sell rice and livestock.

These sources were compared between income groups and kind of diseases.

Four mainly common diseases chosen were:

Respiratory Diseases

Digestive Diseases

Dental Diseases

Transport and Labor Accidents

Other diseases

4.3 Research Design

- This study is a cross-sectional study.
- Target population: Patient or head of family that had a person who got sick in 4 week prior before interviewing.
  Patient had to use one kind of health services: CHC, private doctor, drug vendor inside the district
- Study areas: Hungha district, Thaibinh province, Vietnam.
- Time study: from February 13, 2004 to February 29, 2004
- Data will be collected based on questionnaires

Sample Design

Hungha is district in rural Thaibinh including 34 communes. The sample will be distributed in both town and village.

The sample can be chosen by using the formula:

$$n = \frac{Z^2_{\alpha} \cdot p \cdot q}{d^2}$$

Where:

n = desired sample size.

Z_{α} = the degree of confidence which is required to be within the specified range. With 95% confidence interval, Z will be 1.96

p = the variability of characteristic to be measured in the population. In this study, it is an estimation of the proportion of the person who will go to available health services if they get sick.

Nguyen Van Hoa (2002) showed that, the proportion of patients used CHC, private doctor and drug vendor is 86%. Therefore, p = 0.86

q = 1-p is a pessimistic value. In this case, p= 0.86 so q = 0.14.

d is the degree of accuracy required. This is the maximum error for the sample. This is the largest acceptable percentage of difference between the estimated value from the sample and the true population value. In this study, d will be 5%

So the sample size can be calculated following the formula:

$$n = \frac{(1.96)^2 \cdot 0.86 \cdot 0.14}{(0.05)^2} = 185.$$

The sample size is 250

Sample Selection

34 communes will be allocated in 3 clusters. The cluster is determined based on the distance from center of district to the longest commune in the HH district. Randomizing from 3 clusters choose three communes, each communes interviews 83 or 84 patients.

Figure 3.2 Sample Distribution