CHAPTER III

RESEARCH METHODOLOGY

3.1 Study design

This study is a retrospective study made by using secondary data to analyze the determinants of average household health care expenditure in the year of 1998 and 2000. Since we use data of health care expenditure from NSO's survey, health care expenditure is defined as out-of-pocket of household expenditure for health care here. This study will focus on the explanation what determines the household consumption on health care by using econometric instruments. We adopt multiple regression analysis in estimation of model. Misspecification test is conducted for heteroskedasticity by White's general heteroskedasticity test and against functional misspesification by RESET-test. We use "Eviews 4.1" for regression analysis.

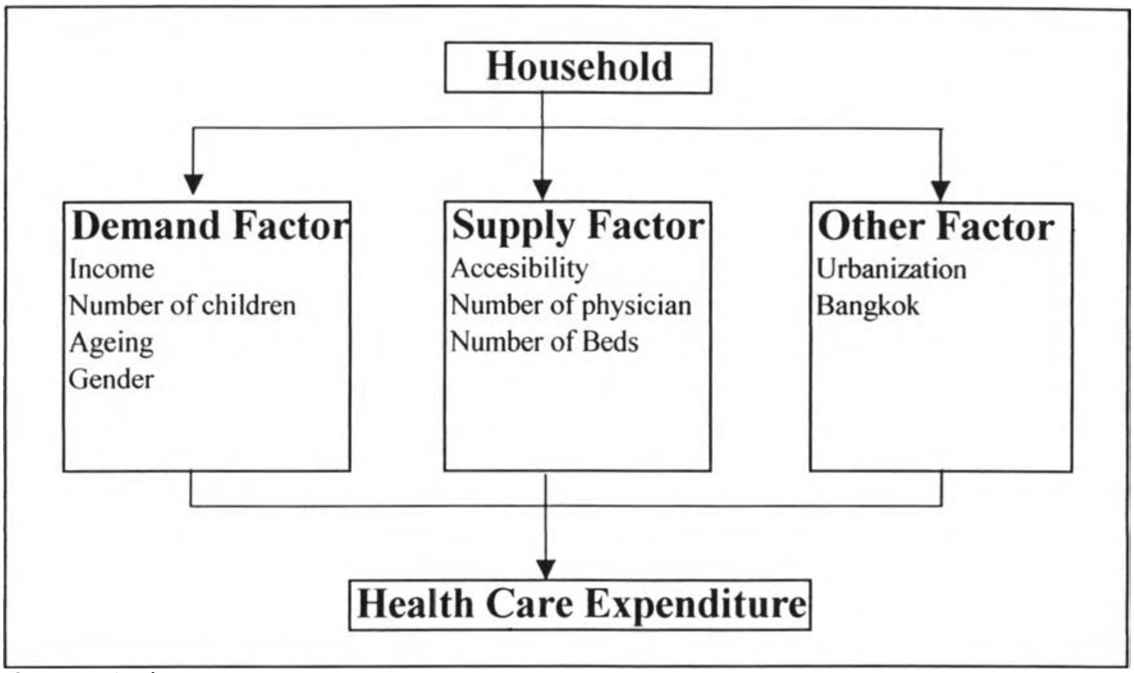
3.2 Study population

The population of this study is whole provinces in Thailand, 76 provinces including Bangkok. The data in this study is the secondary data from household survey conducted by NSO, Ministry of Information, communication and technology Thailand and MoPH.

3.3 Conceptual Framework

Although there is no clear theory what factor determine health care expenditure, we will estimate the relationship between health care expenditure and some factors, such as the condition of health provider, income, education and demographic structure.

Figure3-1 Conceptual Framework



Source: Author

The model in this study includes these variables as follows;

Dependent variables: Health care expenditure, which is measured as average monthly household health care expenditure (including expenditure for medication) by province. This health car expenditure data is available by NSO's survey. Therefore, health care expenditure means out-of-pocket of household expenditure for health care in this study.

Independent variable[7]: Income, Number of population under 14, Number of population over 60, Gender, Urbanization, Number of physician, Number of beds, Accessibility to hospital, Dummy variable (Bangkok). These variables are separated into 3 groups, demand factor, supply factor and other factor.

Demand factors

The variable of income uses the data of average monthly household income by province. According to past works, this variable is expected to increase health care expenditure at household level. In economic theory, the goods, which quantity

---

[7] In order to alleviate multicollinearity problem, we checked partial correlation between each variables by using statistic program, STATA6.0. We didn't find any variables with high partial correlation with more than 0.7 .

demanded increase when income goes up, are called "normal goods". On the contrary, the goods which quantity demanded decrease when income goes up, are called "inferior goods". In other words, normal goods have positive income elasticity and inferior goods have negative income elasticity. Furthermore, the goods with income elasticity less than 1 are called "necessity goods" and above 1 are called a "luxury goods".

Gbesemete and Gerdtham (1992) states that "the age structure of the population may be of prime importance" to determine people's health care expenditure. Besides children under 15 years of age use health care service more than other age. Then, number of children is prospected to have positive effects on health care expenditure.

Ageing is measured by number of population over 60 and also has positive effects to health care expenditure. We know that people doesn't consume health care flatly each year over their life and consume more when they get older.[8]

The variable of Gender is indicated by number of female per 100,000 population. This variable is expected positive or negative sign from regression result. Because female at reproductive age maybe tend to use more hospital service related to childbirth. Conversely, female use less hospital services if male has better access to the services(Preker and Carrin(2004)).

Supply factors

Number of physicians is expected to increase health care expenditure , it is well-known as physician-induced demand. But it is very difficult to identify the induced demand by physicians and consumers, respectively. Carlsen and Grytten (1998) concluded that many studies in 1970s support physician induced demand hypothesis using aggregated data. But recent studies didn't find inducement from their results and reject the hypothesis. Therefore, this study will consider only whether health care expenditure and the number of physicians are related each other or not. Moreover, waiting time for patient would decrease when number of physician increase., then, latent price of health care for patients would drop.

Number of Beds represents medical capacity. This variable is expected to have positive effect on health care expenditure. Because health providers have

---

[8] Tokita(2004) presents Japan's example. Ministry of Welfare estimated that average medical care expenditure per capita in a lifetime was about 17 million Yen (approximately 5.7 million Thai Baht, if 1 Baht equal to 3 Yen). In the total, they spent around 62% after 60 years of age.

incentive to increase health care expenditure for recovering their investment and also pull health care expenditure up. In general, this is called as "Roemer effect".

Accessibility to hospital is measured by the number of hospitals per 1000 square kilometers. If access to hospital is improved, then, health care expenditure is expected to increase. Because access cost falls down due to decrease waiting time and travel cost and as a result, latent cost for patients drop as well. The falling of latent cost would make people go to hospital easily and increase health care expenditure.

Other Factors

Urbanization is percentage of population in municipal area to rural area. As Gbesemete and Gerdtham(1992) insisted, urbanization was brought with industrialization, then people living in urban area faces health risk, such as air pollution or traffic accident. In some cases, patient's travel cost and time cost of hospital in urban area is cheaper than that in rural area (Leu 1986). There is , however, no theoretical agreement of effect on health care expenditure.

Bangkok Dummy is a dummy variable. We use dummy variable because Bangkok, is the capital of Thailand, has biggest population equal to 10% of total population and has possibility to have specific characteristic. More over, the vicinities of Bangkok, Samut Prakan, Nonthaburi, Pathum Thani, Nakhon Pathom and Samut Sakhon, are also included in this dummy variable.

Definition, data source, and expected sign of estimate of each variable are summarized in Table 3-1 and Table 3-2.

Table3-1 Definition and source of variables

Variables	Definition	Source
Devendent variable		
HE	Average monthly household health care expenditure by province	NSO (2000, 2001)
Independent variables		
INC	Average monthly household income by province	NSO (2000, 2001)
Child	Number of population under 14 years old per 100,000 population	NSO (2002)
Elder	Number of population over 60 years old per 100,000 population	NSO (2002)
Fe	Number of female per 100,000 population	NSO(2002)
Acc	Number of hospitals in each 1,000 km^2	MoPH (2003)
Phy	Number of physician per 100,000 population	MoPH (2003)
Bed	Number of bed per 100,000 population	MoPH (2003)
Urban	Number of people living in urban area per 100,000 population	NSO (2002)
BKK	Bangkok Dummy	-

Table3-2 Expected sign of coefficients

INC	Child	Elder	Fe	Acc	Phy	Bed	Urban	BKK
+	+	+	+/-	+	+	+	+/-	+/-

Model 1 use 76 provinces as sample for regression, while sample size of Model 2 and Model 3 are 38. Model 2 estimate for higher income provinces and Model 3 estimate for lower income provinces. The provinces are ranked by income level as of 1998, and then we separate into 2 groups, higher income group and lower income group. Higher income group consists of the ranking 1^{st} to 37^{th} provinces and lower income group consists of the ranking 38^{th} to 76^{th} provinces. This study separate samples, so that it can observe changes among 2 groups between 1998 and 2000. Furthermore, Bangkok dummy variable is excluded in regression analysis of Model 2 and Model 3. Because Bangkok and the vicinity are included in higher income group for both 1998 and 2000, so that the analysis might have problems when compare those results between higher income group and lower income group. Higher income provinces and lower income provinces are as follows;

Higher income group:

Chiangmai, Sukhothai, Lopburi, Pattani, Singburi, Angthong, Chainat, Phichit, Nakhon Si Thammarat, Khon Kaen, Samutsongkhram, Trat, Nakhonnayok, Surat Thani, Krabi, Phitsanulok, Patthalung, Suphanburi, Rayong, Nong Khai, Ratchaburi, Ayutthaya, Trang, Kanchanaburi, Saraburi, Chachoengsao, Chonburi, Petchaburi, Lampang, Songkhla, Nakhonpathom, Chanthaburi, Samutsakhon, Samutprakarn, Phuket, Pathumthani, Nonthaburi, Bangkok.

Lower income group :

Mae Hong Son, Roi Et, Chaiyaphum, Narathiwat, Surin, Kalasin, Sakaeo, Nan, Yasothon, Phayao, Yala, Phangnga, Sakon Nakhon, Si Sa Ket, Udon Thani, Uthai Thani, Nakhon Phanom, Phetchabun, Amnat Charoen, Buri Ram, Nong Bua Lamphu, Maha Sarakham, Phrae, Lamphun, Chiangrai, Satun, Ranong, Nakhon Ratchasima, Tak, Nakhonsawan, Kamphaengphet, Phrachinburi, Mukdahan, Loei, Prachuapkhirikhan, Uttaradit, Chumphon, Ubon Ratchathani.

When we consider representative household at each province as *i*, the model can be written as follows:

Model 1(Sample 76, all provinces)

$$HE_i = f(INC_i, Child_i, Elder_i, Fe_i, Acc_i, Phy_i, Bed_i, Urban_i, BKK)$$

or in a multiplicative form:

$$HE_i = \beta_0 \times INC_i^{\beta 1} \times Child_i^{\beta 2} \times Elder_i^{\beta 3} \times Fe_i^{\beta 4} \times Acc_i^{\beta 5} \times Phy_i^{\beta 6} \times Bed_i^{\beta 7} \times Urban_i^{\beta 8} \times BKK^{\beta 9}$$

Model 2(Sample 38, the provinces with higher income)

$$HE_i = f(INC_i, Child_i, Elder_i, Fe_i, Acc_i, Phy_i, Bed_i, Urban_i)$$

or in a multiplicative form:

$$HE_i = \beta_0 \times INC_i^{\beta 1} \times Child_i^{\beta 2} \times Elder_i^{\beta 3} \times Fe_i^{\beta 4} \times Acc_i^{\beta 5} \times Phy_i^{\beta 6} \times Bed_i^{\beta 7} \times Urban_i^{\beta 8}$$

Model 3(Sample 38, the provinces with lower income)

$$HE_i = f(INC_i, Child_i, Elder_i, Fe_i, Acc_i, Phy_i, Bed_i, Urban_i)$$

or in a multiplicative form:

$$HE_i = \beta_0 \times INC_i^{\beta 1} \times Child_i^{\beta 2} \times Elder_i^{\beta 3} \times Fe_i^{\beta 4} \times Acc_i^{\beta 5} \times Phy_i^{\beta 6} \times Bed_i^{\beta 7} \times Urban_i^{\beta 8}$$

with $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9 > 0$
,where

HE = average monthly household health care expenditure
INC = average monthly household income
Child = number of population under 14 years old per 100,000 people
Elder = number of population over 60 years old per 100,000 people
Fe = number of female per 100,000 people
Acc = number of hospitals in each 1,000km^2
Phy = number of physician per 100,000 people
Bed = number of hospital beds per 100,000 people
Urban = number of people living in urban area per 100,000 people
BKK = Bangkok Dummy

As Gbesemete and Gerdtham(1992) pointed out, linear multiplicative functional form was considered suitable for estimation. Therefore, we adopt this idea and transfer our model to log linear form. More importantly, this study can interpret coefficient of estimation equation as elasticity, that is, how much the factor affect to increase(or decrease) health care expenditure when the factor increase 1%.

Then, our estimation equation can be written as follows:

Estimation equation 1

$$\ln HE_i = \ln \beta_0^* + \beta_1 ln\mathrm{INC}_i + \beta_2 ln\mathrm{Child}_i + \beta_3 ln\mathrm{Elder}_i + \beta_4 ln\mathrm{Fe}_i + \beta_5 ln\mathrm{Acc}_i + \beta_6 ln\mathrm{Phy}_i + \beta_7 ln\mathrm{Bed}_i + \beta_8 ln\mathrm{Urban}_i + \beta_9^* \mathrm{BKK} + \mu_i$$

Estimation equation 2

$$\ln HE_i = \ln \beta_0^* + \beta_1 ln\mathrm{INC}_i + \beta_2 ln\mathrm{Child}_i + \beta_3 ln\mathrm{Elder}_i + \beta_4 ln\mathrm{Fe}_i + \beta_5 ln\mathrm{Acc}_i + \beta_6 ln\mathrm{Phy}_i + \beta_7 ln\mathrm{Bed}_i + \beta_8 ln\mathrm{Urban}_i + \mu_i$$

Estimation equation 3

$$\ln HE_i = \ln \beta_0^* + \beta_1 ln\mathrm{INC}_i + \beta_2 ln\mathrm{Child}_i + \beta_3 ln\mathrm{Elder}_i + \beta_4 ln\mathrm{Fe}_i + \beta_5 ln\mathrm{Acc}_i + \beta_6 ln\mathrm{Phy}_i + \beta_7 ln\mathrm{Bed}_i + \beta_8 ln\mathrm{Urban}_i + \mu_i$$