CHAPTER IV

METHODOLOGY

This chapter is composed of two sections: models in the study, and variables and sources of data.

4.1 Models in the study

De Gregorio, and Lee (2002) used cross-country data covering a number of countries for the period from 1960 to 1990 to examine the relationship among income inequality, the level and dispersion of education, and the level of income across countries. They estimated the following regression.

$$G_{j,t} = a_0 + a_1\sigma_{j,t} + a_2E_{j,t} + a_3\log Y_{j,t} + a_4[\log Y_{j,t}]^2 + a_D\mathbf{D}_j + \epsilon_{j,t} \quad (4.1.1)$$

where G is the Gini coefficient

σ is dispersion of educational attainment in the population

E is the average years of school attainment for the population aged 15 and over

Y is GDP per capita

D is a set of dummy variables that distinguish certain characteristics and region to which countries belong.

The subscript j and t index countries and periods, respectively.

In terms of this paper, it focuses on the issue of whether educational expansion can reduce income inequality and whether higher educational inequality increases the inequality of income. Quality of education is also added as an explanatory variable to capture the effect of educational quality on income inequality. This is because the higher quality of education might raise more on productivity of low income students, and thus narrow the gap of income inequality. For more explanation, the low income students might gain more benefits from the higher educational quality, such as higher educational expenditure and lower pupil-teacher ratio, so their productivities increase a lot. While the high income students may have higher productivities, so improvement in quality of education has less impact on their productivities. Income differences between these two groups of people are lower finally. Thus, higher quality of education may lead to more equal of income.

The hypothesis mentioned above is tested by two methods. Firstly, it is tested by cross section technique in the context of Thailand by using provincial data for four periods, 1996, 1998, 2000 and 2002 instead of over period data like most research. Thus, there are four regression equations. Each regression is for each period. Secondly, provincial data are grouped into thirteen groups by educational service areas and the hypothesis is tested by pooled least square technique.

Adapting the above regression equation to fit the hypothesis and data, this study estimates the following regression.

The first regression equation is

$$YI_i = a_0 + a_1E_i + a_2EI_i + a_3EX_i + a_4PT_i + a_5\log Y_i + a_6[\log Y_i]^2 + \epsilon_i \quad (4.1.2)$$

The second regression equation is

$$YI_{i,t} = b_{0i} + b_1E_{i,t} + b_2EI_{i,t} + b_3EX_{i,t} + b_4PT_{i,t}\, b_5\log Y_{i,t} + b_6[\log Y_{i,t}]^2 + \epsilon_{i,t} \quad (4.1.3)$$

where

YI is income inequality

E is educational attainment

EI is educational inequality

EX is expenditure on eduation

PT is pupil-teacher ratio

logY and $[logY]^2$ are logarithm of income and square of logarithm of income respectively. They are added into the regression equation to test Kuznets hypothesis. The square term is added to represent the quadratic form of inverted-U shape.

The subscript i indexes educational service areas. For the first regression i is 1, 2,..., 76 and for the first regression i is 1, 2,..., 13.

The subscript t indexes periods. t = 1996,1998, 2000, 2002

In order to test the hypothesis by pooled least square method, provincial data are grouped to represent different areas of educational administration in Thailand under hypothesis that provinces in the same educational service areas might have the same educational management structure contributing to similar pattern of income distribution, but they might differ from provinces in different educational service areas.

Educational service areas, taking into consideration the number of educational institutions, the number of population, geography, and cultural background as the main criteria as well as other appropriate conditions, are divided into thirteen groups as follows.

Central Educational Service Area: Bangkok

Educational Service Area 1: Nonthaburi, Samut Prakan, Patum Thani, Samut Sakhon, and Nakhon Pathom

Educational Service Area 2: Yala, Pattani, Narathiwat, and Satun

Educational Service Area 3: Songkhla, Nakhon Si Thammarat, Patthalung, Surat Thani, and Chumphon

Educational Service Area 4: Phuket, Trang, Krabi, Phang-nga, and Ranong

Educational Service Area 5: Ratchaburi, Phetchaburi, Prachuapkhirikhan, Suphan Buri, Kanchanaburi, and Samut Songkhram

Educational Service Area 6: Lop Buri, Phra Nakhon Si Ayutthaya, Ang Thong, Saraburi, Sing Buri, Chainat, and Uthai Thani

Educational Service Area 7: Phitsanulok, Nakhon Sawan, Phichit, Kampang Phet, Uttaradit, Tak, Sukhothai, and Phetchabun

Educational Service Area 8: Chiang Mai, Chiang Rai, Lampang, Nan, Phrae, Lamphun, Mae Hong Son, and Phayao

Educational Service Area 9: Udon Thani, Nong Khai, Loei, Khon Kaen, Sakon Nakhon, and Nong Bua Lamphu

Educational Service Area 10: Ubon Ratchathani, Roi Et, Maha Sarakham, Nakhon Phanom, Kalasin, Yasothon, Mukdahan, and Amnart Charoen

Educational Service Area 11: Nakhonratchasima, Chaiyaphum, Buriram, Surin, and Si Sa Ket

Educational Service Area 12: Chachoensao, Prachin Buri, Chonburi, Chanthaburi, Rayong, Trat, Nakhon Nayok, and Sa Kaeo

4.2 Variables and Sources of Data

4.2.1 Income Inequality

Gini coefficient is used to measure an inequality of income because of the availability of data. Its value ranges between zero and one. If Gini coefficient is equal to zero, income is absolutely equal. It implies that everybody has the same level of income. In contrast, if Gini coefficient is equal to one, income is absolutely unequal. It means that the total income is possessed by one person. This study uses provincial Gini coefficient during 1996 to 2002 reported by National Economic and Social Development Board. This value is computed every two year. For the second regression, Gini coefficient for each educational service area is the weighted average of provincial Gini coefficient which is weighted by the number of population.

4.2.2 Educational Attainment

There are two indicators widely used as educational attainment. In the early research, the used variable is school enrollment ratio which is the ratio of number of students enrolled at a grade level (primary, secondary, or tertiary) to total population of corresponding age group. The most commonly used are primary and secondary enrollment ratios. One problem of using this variable is that enrollment ratio only measures the flow of population's education or access to education. It does not show the cumulated educational attainment. Psacharopoulos and Arriagada (1986) suggested that the proper indicator is the stock of educational attainment defined as average year of schooling. Thus, this study uses average year of schooling of workers to represent educational level. This indicator can be constructed from the educational data of employed labor at the age of thirteen and over from labor force survey by province in the third quarter reported by National Statistical Office. It is computed as follows:

No education is the status of a person who never enters schools or does not receive education. It is computed as zero year of schooling.

Lower than lower elementary level is the status a person who undergoes lower than four years of elementary education. It is computed as two years of schooling.

Lower elementary level is the status of a person who undergoes at least four years of elementary education, but does not finish the higher level. It is computed as four years of schooling.

Upper elementary level is the status of a person who undergoes at least six years of elementary education, but does not finish the higher level. It is computed as six years of schooling.

Lower secondary level is the status of a person who has completed elementary education and undergoes at least three years of secondary education, but does not finish the higher level. It is computed as nine years of schooling.

Upper secondary level is the status of a person who has completed elementary education and undergoes at least six years of secondary education, but does not finish the higher level. It is computed as twelve years of schooling.

Vocational level is the status of a person who has completed lower secondary education and undergoes at least a three-year course in vocational and technical colleges, but does not finish the higher level. It is computed as twelve years of schooling.

Higher Educational level is divided into two streams.

University level is the status of a person who has completed upper secondary education and attends four to six-year courses for bachelor degree, including a person who obtains graduate diploma, master and doctoral degree level. It is computed as sixteen years of schooling.[1]

[1]The data from labor force survey do not classify the number of people who graduate university level into bachelor, master, or doctoral degree. Thus, people who complete university level are given the same number of year which is sixteen years of schooling.

Higher Vocational level is the status of a person who has completed upper secondary education or vocational education and undergoes at least a two-year vocational course from vocational institutions. It is computed as fourteen years of schooling.

Teacher Training is the status of a person who obtains a certificate or diploma from teacher training colleges. It is computed as fourteen years of schooling.

Other level is computed as two years of schooling.[2]

In the year 2002, the labor force survey changed responders from employed labor at the age of thirteen and over to fifteen and over. Some definitions of level of education also change.

No education is the status of a person who never enters schools or does not receive education. It is computed as zero year of schooling.

Lower than elementary level is the status of a person who undergoes lower than six years of elementary education. It is computed as three years of schooling.

Elementary level is the status of a person who undergoes at least six years of elementary education, but does not finish the higher level. It is computed as six years of schooling.

Lower secondary level is the status of a person who has completed elementary education and undergoes at least three years of secondary education, but does not finish the higher level. It is computed as nine years of schooling.

Upper secondary level is divided into three streams.

---

[2]Other level is the status people who do not qualify any categories. They might not receive any certificates of education. Therefore, I give them two years of schooling which is equal to people who graduate lower than lower elementary level since they also do not receive any certificates.

Academic stream is the status of a person who has completed elementary education and undergoes at least six years of secondary education, but does not finish the higher level. It is computed as twelve years of schooling.

Vocational stream is the status of a person who has completed lower secondary education and undergoes at least a three-year course in vocational and technical colleges, but does not finish the higher level. It is computed as twelve years of schooling.

Teacher training stream is the status of a person who has completed lower secondary education and undergoes at least a three-year course in teacher training colleges, but does not finish the higher level. It is computed as twelve years of schooling.

Higher Educational level is divided into three streams.

Academic level is the status of a person who has completed upper secondary education and attends four to six-year courses for bachelor degree, including a person who obtains graduate diploma, master and doctoral degree level. It is computed as sixteen years of schooling.

Occupational stream is the status of a person who has completed upper secondary education or vocational education and undergoes at least two years of an occupational course. It is computed as fourteen years of schooling.

Teacher Training is the status of a person who obtains a certificate or diploma from teacher training colleges. It is computed as fourteen years of schooling.

Other level is computed as two years of schooling.

The average year of schooling of labor force in each province is equal to the summation of the multiplication of the number of workers who graduate at each level of education in that province and the number of years for graduating at each

level and is divided by the total number of labor forces in that province. It can be written in the formula form shown below.

$$\text{Average year of schooling for each province} = \frac{\sum_{i=1}^{n} W_i N_i}{T}$$

where

W is the number of workers who graduate at each level of education in each province.

N is the number of years for graduating at each level of education. For example, it is equal to four for lower primary level.

T is the total number of labor force in each province.

n is the number of educational level which is equal to eleven.

i = 1, 2,..., n

After obtaining average year of schooling for each province, group them by educational service area and weight each values by the proportion of labor force in that province to total labor force in that educational service area. Following this process, one can obtain average year of schooling for each educational service area.

4.2.3 Educational Inequality

A standard deviation of educational attainment is constructed to measure educational inequality. The formula to compute a standard deviation is

$$\text{S.D. for each province} = \sqrt{\frac{\sum_{i=1}^{n} Wi\left(N_i - \mu\right)^2}{T}}$$

where

μ is the average year of schooling in each province.

The standard deviation of educational attainment for each educational service area can be calculated by using the same weight as average year of schooling.

4.2.4 Quality of Education

Quantity alone is not enough to measure educational achievement. Quality must be taken into consideration. Two approaches to measure the quality of education are input approach, and output approach. The former uses expenditure on education, pupil-teacher ratio, and repetition rate, while the latter measures cognitive skills of individuals. This paper uses the first approach because of the unavailability of data on the second one. The quality of education used in this paper is composed of educational expenditure, and pupil-teacher ratio. The repetition rate is omitted because most provinces have the same repetition rate, and from observation, their differences are only little.

4.2.4.1 Pupil-Teacher Ratio

The lower pupil-teacher ratio represents the higher quality of education since teachers can pay attention to their students thoroughly.

$$\text{Pupil-Teacher Ratio for each province} = \frac{S}{T}$$

where

S is the total number of students in each province.

T is the total number of teachers in each province.

The number of students and teachers obtained from the data set of Ministry of Education. This provincial data set is available only during 1996-2003. Pupil-teacher ratio for each educational service area is an average value of pupil-teacher ratio by province.

4.2.4.2 Educational Expenditure

More educational expenditure helps to raise quality of education by improving education instruments and encouraging teachers to do the best of their ability via higher salaries.

The data of educational expenditure by province obtained from National Account Office is in the form of current price, so it is adjusted by using consumer price index reported by Bureau of Trade and Economic Indices, Ministry of Commerce in order to be constant price. After adjusting, it is divided by the number of population in order to change it into educational expenditure per capita.

However, the data of educational expenditure by province collected from National Account Office is available only during 1999 and 2002. As a result, there is a need to estimate the educational expenditure of the year 1996 and 1998. Since only educational budget of Thailand collected from the Ministry of Finance is available during 1996 to 1998, this study estimates educational expenditure from educational budget by nonlinear estimation under the assumption that they have the same growth rate.

After obtaining the estimated value of the educational expenditure per capita of Thailand in the year of 1996 and 1998, the provincial educational expenditure per capita is calculated by assuming that the proportion of educational expenditure in each province is constant. When using the number of population as a weight and calculating weighted average of educational expenditure by province for

each educational service area, expenditure on education by educational service area is achieved.

4.2.5 Economic Development

Log of Gross Provincial Product (GPP) per capita at constant 1988 price and square of it are used to measure economic development and test inverted-U shape hypothesis. The data of this variable can be collected from National Economic and Social Development Board. For each educational service area, log of GPP per capita is equal to log of the ratio between summation of GPP per capita for all provinces in that educational service area and total population in that educational service area. The log form is used instead of the level form because changes in log of GPP represent change in growth rate of GPP. The square term is also added in order to represent a quadratic form. If the sign of level term is positive, while that of square term is negative, it will support inverted-U shape hypothesis that income inequality rises in the early stage of economic development, but falls in the later stage. In contrast, if the sign of level is negative, but that of square term is positive, the results represent U shape instead of inverted-U shape. That means income firstly becomes more equal, but it is worse later.