CHAPTER 2

ECONOMIC THEORIES AND LITERATURE REVIEW

2.1 Economic Theories

2.1.1 The Solow Growth Model

Assumptions of the model, in the closed economy Savings and investment are exogenous variables. Factor accumulation and technological growth are also exogenous variables.

Production function, with physical capital K, labor L and level of technology is A:

$$Y(t) = F(K(t), A(t)L(t)) \tag{2.1}$$

Thus following from Robert Solow the economic growth will depend on

1. An increase in saving and investment
2. An increase in capabilities of Human capital
3. An increase in the level of technology, or new technology innovation

Solow try to explain that in each country has the different way to drive their economic growth by there is no impact from the economic size. The main factors lead economic grow further more are capital, labor, and technology in the production.

- **Capital**

The economic growth (Y, GDP) when the condition that an investment per capita (I/L) is in the high rate.

$$Y = F(K, L) \rightarrow Y/L = F(K/L, L/L) \tag{2.2}$$

People spend their income for consumption and save their proportionate of income

Y = C + S Recall that the equilibrium of investment and saving, so that

Y = C + I

Assume that the capital stock has depreciation (D) over the period of time. If the rate of investment is less than the increase in depreciation of capital stock means that the investment cannot compensate the loss from depreciation at that time. Thus (I/L) is declining. If the rate of depreciation is fixed at k, the depreciation is equal to kD.

At the Steady – State, the investment per capita equate to depreciation and this point is relate to the level of output per capita (Y/L). Whenever investment per capita is less than depreciation leads to a decrease in Y/L. Therefore, arises in investment $(I_1/L \rightarrow I_2/L)$ will increase capital $(K_1/L \rightarrow K_2/L)$ and national income $(Y_1/L \rightarrow Y_2/L)$.

In case that, there is a population growth while the level of output still be at the same level. It will effect investment per capita (I/L) and the output per capita (Y/L) decrease.Therefore, an increase in population is the same as an increase in the depreciation (kD).

$$Y = F(K, L) \rightarrow \frac{Y}{\Delta L} = F(\frac{K}{\Delta L}, \frac{\Delta L}{\Delta L}) \tag{2.3}$$

Therefore,

$$kD \rightarrow (k + \Delta L)D \tag{2.4}$$

To maintain the level of output to be stable, the investment rate should not less than the rate of population growth.

- **Technological Progress**

We can increase the level of output per capita by increase a potential in labor force or increase capabilities, knowledge and the important thing is technological progress.

$$Y = F(K/L) \rightarrow zY/L = F(zK/L, zL/L). \tag{2.5}$$

2.1.2 The Neoclassical Growth Model

The neoclassical growth model, which originated with the work of Solow and Swan, consists of the following relationships: a production function,

$$Y_t = f(h_t, k_t) \tag{2.6}$$

Where Y_t is output, h_t is labor, and k_t is capital at date t = 0, 1, 2…, plus a law of motion for the capital stock,

$$k_{t+1} = (1-\delta)k_t + \sigma Y_t \tag{2.7}$$

where $\delta \in (0,1)$ is the depreciation rate and $\sigma \in (0,1)$ is the saving rate.

We assume that f is homogeneous of degree 1, increasing, concave, and twice continuously differentiable. And now let h_t is constant and normalize $h_t = 1$, and $F(k) = f(1,k)$. Then $F' > 0$, $F'' < 0$. We further assume that $F(0) = 0$, $F'(0) = \infty$ and $F'(\infty) = 0$. The initial capital stock, k_0, is given exogenously.

Set, a fixed number of hours $h_t = 1$,

saves or invests $i_t = \sigma Y_t$,

consumes $c_t = (1-\sigma)Y_t$, each period.

Substituting the production function into the law of motion for capital yields a first order difference equation in k_t,

$$k_{t+1} = (1-\delta)k_t + \sigma F(k_t) \equiv g(k). \tag{2.8}$$

Together with the initial condition k_0, (2.1) completely determines the entire time path of the capital stock. A steady state of the system is a solution to $k = g(k)$.

Our assumptions imply that, as shown, $g(0) = 0, g'(0) > 1$, and there is a unique k*>0 such that $k^* = g(k^*)$. So, we get 2 steady state, k=0 and k=k*. At k*, we have $\sigma F(k^*) = \delta k^*$, which implies that savings just offsets depreciation and the capital – output ratio is $\frac{k}{y} = \frac{\sigma}{\delta}$ and also that

$$c^* = y^* - \delta k^* \tag{2.9}$$

It's clear that k* is increasing in $\frac{\sigma}{\delta}$. Moreover, c* is first increasing and then decreasing in σ. The saving rate that maximizes steady state consumption can easily be shown to satisfy

$$F'(k^*) = \delta \tag{2.10}$$

This is Phelps' so-called "golden rule" of capital accumulation. Notice that the basic neoclassic growth model not exhibit long run growth: for all $k_0 > 0$, as t increase, k_t converges to k*

2.1.3 The New Growth Theory of Externalities

Romer (1986) stated that the definition of capital should include the following result from investigate that capital stock. For instance, the new innovation of product, leads to create the new capital in the form of knowledge that can be accumulate from the past. The accumulate of capital stock including knowledge cause the externalities (positive externalities) to the economic growth.

Romer assumes that the production function at j unit is following this equation:

$$Y_j = \bar{A}K_j^{-1-\alpha}N_j^{\alpha} \tag{2.11}$$

When K_j and N_j is the capital stock and labor at j unit and $\bar{A}$ is technological progress, by assuming $\bar{A} = AK^{\beta}$ means that the technological progress is involving in this model or mention to capital accumulate in the economy (K) and A is constant term. We will see that knowledge is included in this capital. Let $\bar{A}$ into the production equation we will get,

$$Y_j = AK^{\beta}K_j^{-1-\alpha}N_j^{\alpha} \tag{2.12}$$

While K is national capital stock; K is determined by explicit factors not the firms decisions resulting an increasing in capital gain even though, firms are facing the constant rate of return.

In order to set up the model as mentioned earlier to be simultaneous to the equilibrium of perfect competitive market. On the other words, no surplus profitable occur one firm but continuous growth will never change on this model even though, there is no change from external technology.

$1-\alpha+\beta \geq 1$ This is a growth rate that the externalities in macro level has great enough for compensating the declining return on investment in each producer/firm.

Assume that the size of each firm is equal and same as other. The total capital stock is $K = nK_j$ if every firm has the same ability to access the same level of available technology and get the factors of production in the same price, they will produce by using the same proportion of factors of production.

So, the equation of aggregate production function is:

$$Y = An^{\beta} K^{-1-\alpha+\beta} N^{\alpha} \tag{2.13}$$

For the supply of labor, is fixed to be constant for firm and assume that the proportion of saving is constant with the rate of production output. Thus, the capital growth rate will equal to:

$$\frac{K_{t+1}}{K_t} - 1 = 3AK_t^{-\alpha+\beta} N^{\alpha} n^{\beta} - \delta \tag{2.14}$$

In the case that, $1-\alpha+\beta > 1$ the economy will growth in accelerate rate. It is hard to become possible. Nowadays, for $1-\alpha+\beta = 1$ the economy grow in the constant rate. This is simultaneous to the long – run equilibrium that output and capital growth constant by assuming there is an constant rate of saving.

If we want to determine the level of saving from initial the model or determine by economic representative agent. The problem is the amount of consumption at any time period under the condition that shows in this equation:

$$\max \int_0^{\infty} u(C_t) e^{-pt} dt \tag{2.15}$$

condition $\qquad \dot{K}_t = \bar{A} K_t^{-1-\alpha} - c$

and $\qquad \dot{K}_t \geq 0$

While $\bar{A}$ is explicit determination and the depreciation rate is zero.

To prove mathematically we use Current-value Hamiltonian (H) value based on year.

$$H = u(c) + \lambda[\bar{A} K^{-1-\alpha} - c] \tag{2.16}$$

When λ is the shadow price of investment measured by the unit of utils based on current-value prices.

The necessary conditions are following:

$$\frac{\partial H}{\partial c} = 0: \qquad u'(c) = \lambda \Rightarrow c^{-\varepsilon} = \lambda \tag{2.17}$$

$$\frac{\partial H}{\partial K} = \rho\lambda - \dot{\lambda}: \qquad \lambda\overline{A}(1-\alpha)K^{-\alpha} = \rho\lambda - \dot{\lambda} \tag{2.18}$$

instead $\lambda = c^{-\varepsilon}$ in (2) we will get,

$$\overline{A}(1-\alpha)K^{-\alpha} = \rho - \left(-\varepsilon\frac{\dot{c}}{c}\right) \tag{2.19}$$

Assume that $\overline{A} = AK^{-\beta}$ so,

$$\frac{\dot{c}}{c} = \frac{1}{\varepsilon}\left[A(1-\alpha)K^{-\alpha+\beta} - \rho\right] \tag{2.20}$$

If $1-\alpha+\beta = 1$ the return is constant returns to scale, the economy can maintain growth in stability. Since, diminishing returns to capital is compensated by the change in technology that is the externality impact ($\overline{A}$).

If $1-\alpha+\beta > 1$, the increasing returns to capital leads to economic growth unlimited.

If $1-\alpha+\beta < 1$, the economy will stuck in the long – run term.

- **The New Growth Theory from Human Capital**

According to Lucas model (1988), mentioned that the change in technological progress is the result in Human Capital. Advancement of Knowledge is hard to separate from human capital. Lucas interpreted that human capital as a average knowledge in the labor that necessary in the production function. Lucas (1988) the technological accumulate of human capital can be separated in 2 characteristics. The first one, There is no externalities from human capital accumulation. Second, there is the externalities impact from human capital accumulation.

Therefore, the firm's production not only depends on human capital but also depends on the average level of human capital per capita in the economy. In case, Technology of human capital accumulation has the impact from externalities, the investment on education will be less than the level of Socially Optimum.

The first model Lucas, the production equation is:

$$Y_t = BK_t^{-1-\alpha}(NH_t)^{\alpha} \tag{2.21}$$

While H_t is Human Capital of representative agent in the economy system

Assume that human capital accumulation (H_t) has the same characteristic as (K_t)

$$H_{t+1} = I_t^h + (1-\delta)H_t \tag{2.22}$$

$$K_{t+1} = I_t^h + (1-\delta)K_t \tag{2.23}$$

and $I_t = I_t^h + I_t^h$

$Y_t = C_t + I_t$

In the long – run equilibrium the economy will accumulate of both capitals until the surplus output from both of capital are equal.

$$(1-\alpha)BK_t^{-\alpha}(NH_t)^{\alpha} = \alpha BK_t^{-1-\alpha}(NH_t)^{\alpha-1}N \tag{2.24}$$

Reform the equation;

$$\frac{K_t}{H_t} = \frac{(1-\alpha)}{\alpha} \tag{2.25}$$

plug $H_t = \frac{\alpha}{1-\alpha}K_t$ in production equation

$$Y_t = BK_t^{-1-\alpha}N^{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}K_t^{\alpha}$$

$$= BN^{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{\alpha}K_t$$

$$= AK_t \qquad \text{while } A \equiv BN^{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{\alpha} \tag{2.26}$$

Assume that the investment is holding constant (δ) of level of output. If we want to hold $\left(\frac{K}{H}\right)$ to be constant, the economies have to separate the physical investment as $(1-\alpha)$ or equal to $\delta(1-\alpha)$ of the level of output.

The economic growth at Steady – State

$$\frac{Y_{t+1}}{Y_t} = \frac{K_{t+1}}{K_t} \tag{2.27}$$

$$= \frac{\delta(1-\alpha)AK_t + (1-\delta)K_t}{K_t}$$

$$= \delta(1-\alpha)A + (1-\delta)$$

or $$\frac{\Delta Y_t}{Y_t} = \delta(1-\alpha)A - \delta \tag{2.28}$$

According from Lucas model, an increase in saving will affect the economic growth in the positive way. Meanwhile, The Neo – classical stated that an increase in saving leads to an increase only in the level of income (level effect).

2.1.4 The Theory of Human Capital

- General Issues

One of the most important ideas in labor economics is to think of the set of marketable skills of workers as a form of capital in which workers make a variety of investments. This perspective is important in understanding both investment incentives, and the structure of wages and earnings.

Human capital corresponds to any stock of knowledge or characteristics the worker has (either innate or acquired) that contributes to his or her productivity. This definition is broad, and this has both advantages and disadvantages. The advantages are clear: it enables us to think of not only the years of schooling, but also of a variety of other characteristics as part of human capital investments. These include school quality, training, attitudes towards work, etc. Using this type of reasoning, we can make some progress towards understanding some of the differences in earnings across workers that are not accounted by schooling differences alone.

The disadvantages are also related. At some level, we can push this notion of human capital too far, and think of every difference in remuneration that we observe in the labor market as due to human capital. For example, a worker who is paid less than Ph.D. there must be because of lower skills in some other dimension that’s not being measured by my years of schooling. This is the famous unobserved heterogeneity issue. The presumption that all pay differences are related to skills (even if these skills are unobserved to the economists in the standard data sets) is not a bad place to start when we want to impose a conceptual structure on empirical wage distributions, but there are many notable exceptions, some of which will be discussed later. Here it is useful to mention three:

1. Compensating differentials: a worker may be paid less in money, because he is receiving part of his compensation in terms of other (hard-to-observe) characteristics of the job, which may include lower effort requirements, more pleasant working conditions, better amenities etc.

2. Labor market imperfections: two workers with the same human capital may be paid different wages because jobs differ in terms of their productivity and pay, and one of them ended up matching with the high productivity job, while the other has matched with the low productivity one.

3. Taste-based discrimination: employers may pay a lower wage to a worker because of the worker's gender or race due to their prejudices.

In interpreting wage differences, and therefore in thinking of human capital investments and the incentives for investment, it is important to strike the right balance between assigning earning differences to unobserved heterogeneity, compensating wage differentials and labor market imperfections.

- Human Capital Investments and The Separation Theorem

According with the partial equilibrium schooling decisions and establish a simple general result, sometimes referred to as a "separation theorem" for human capital investments. There is the basic model in continuous time for simplicity. Consider the schooling decision of a single individual facing exogenously given prices for human capital. Throughout, assume that there are perfect capital markets. The separation theorem referred to in the title of this section will show that, with perfect capital markets, schooling decisions will maximize the net present discounted value of the individual. More specifically, consider an individual with an instantaneous utility function $u(c)$ that satisfies the standard neoclassical assumptions. In particular, it is strictly increasing and strictly concave. Suppose that the individual has a planning horizon of T, discount the future of rate $\rho > 0$ and faces a constant flow rate of death equal to $\nu \geq 0$. Standard arguments imply that the objective function of this individual at time t = 0 is

$$\max \int_0^T \exp(-(\rho+\nu)t)u(c(t))dt. \tag{2.29}$$

Suppose that this individual is born with some human capital $h(0) \geq 0.$ Suppose also that his human capital evolves over time according to the differential equation

$$\dot{h}(t) = G(t, h(t), s(t)), \tag{2.30}$$

where $s(t) \in [0,1]$ and may be useful to model constraints of the form $s(t) \in \{0,1\}$, which would correspond to the restriction that schooling must be full-time (or other such restrictions on human capital investments).

The individual is assumed to face an exogenous sequence of wage per unit of human capital given by $[w(t)]_{t=0}^{T}$, so that his labor earnings at time t are

$$W(t) = w(t)[1-s(t)][h(t)+w(t)], \tag{2.31}$$

where $1-s(t)$ is the fraction of time spent supplying labor to the market and $w(t)$ is non-human capital labor that the individual may be supplying to the market at time t.

The sequence of non-human capital labor that the individual can supply to the market, $[w(t)]_{t=0}^{T}$, is exogenous. This formulation assumes that the only margin of choice is between market work and schooling.

Finally, let us assume that the individual faces a constant (flow) interest rate equal to r on savings. Using the equation for labor earnings, the lifetime budget constraint of the individual can be written as

$$\int_0^T \exp(-rt)c(t)dt \leq \int_0^T \exp(-rt)w(t)[1-s(t)][h(t)+w(t)]dt. \tag{2.32}$$

The Separation Theorem, which is the subject of this section, can be stated as follows:

Suppose that the instantaneous utility function $u(\cdot)$ is strictly increasing. Then the result $\left[\hat{c}(t), \hat{s}(t)\hat{h}(t)\right]_{t=0}^{T}$ is a solution to maximize of equation (2.29) subject to (2.30), (2.31), and (2.32) if and only if $\left[\hat{s}(t), \hat{h}(t)\right]_{t=0}^{T}$ maximizes

$$\int_0^T \exp(-rt)w(t)[1-s(t)][h(t)+w(t)]dt \tag{2.33}$$

subject to (2.27), (2.28) and $[\hat{c}(t)]_{t=0}^{T}$ maximizes equation 1 subject to (4) given $[\hat{s}(t),\hat{h}(t)]_{t=0}^{T}$.

That is, human capital accumulation and supply decisions can be separated from consumption decisions.

2.1.5 The Pure Theory of Public Expenditure (Paul A. Samuelson)

Assume that there are two categories of goods: Private consumption goods $(X_1,...,X_n)$ which can be distributed among different individuals (1, 2,…1,…, s) according to the relations $X_j = \sum_1^s X_j^i$;

The collective consumption goods $(X_{n+1},...,X_{n+m})$ which each individual's consumption does not reduce other's consumption of that good, so that $X_{n+j} = X_{n+j}^i$ simultaneously for each and every i^{th} individual and each collective consumption goods.

Moreover, assume that each individual has a consistent set of ordinal preferences with respect to consumption of all goods (collective as well as private) which can be summarized by a regularly smooth and convex utility index

$$u^i = u^i(X_1^i,...,X_{n+m}^i) \tag{2.34}$$

Follow the convention of the partial derivative of any function with respect to its j^{th} argument by a j subscript, thus

$$u_j^i = \partial u_i / \partial X_j^i, \tag{2.35}$$

The amount of economic quantities can be divided into two groups,

1. Outputs which everyone want to maximize,
2. Inputs which everyone want to minimize.

In this convention we are sure that $u_j^i > 0$ always.

In order to keep the production assumptions at the minimum level of simplicity, here is a regularly convex and smooth production-possibility schedule relate to the totals of outputs, as following $F(X_1,...,X_{n+m})=0$, with $F_j>0$ and ratios F_j/F_n determinate and subject to the laws of diminishing returns.

There is a maximal utility frontier representing the Pareto-optimal points which there are an (s-1)fold infinity, with the property that from such a frontier point you can make someone better off by making some other worse off.

Then, consider in the sense of a social welfare function which represents a consistent set of ethical preferences among all the possible states of the system. The restriction that placed on the social welfare function is that it will be increase or decrease when any one person's ordinal preference increase or decrease, all others staying on their same indifference levels, for mathematically any one of its indexes can be written,

$$U=U(u^1,...,u^s) \text{ with } U_j>0.$$

- The Optimal Conditions

There is a best state of the world which is defined mathematically in simple regular cases by the marginal conditions.

$$\frac{u_j^i}{u_r}=\frac{F_j}{F_r} \qquad \begin{array}{l}(i=1,2,...,s;r,\ j=1,...,n)or\\(i=1,2,...,s;r=1;\ j=2,...,n)\end{array} \tag{2.36}$$

$$\sum_{i=1}^{s}\frac{u_{n+j}^i}{u_r^i}=\frac{F_{n+j}}{F_r} \qquad \begin{array}{l}(j=1,...,m;r=1,...,n)or\\(j=1,...,m;r=1)\end{array} \tag{2.37}$$

$$\frac{U_i u_k^i}{U_q u_k^q}=1 \qquad \begin{array}{l}(i,q=1,...,s;k=1,...,n)or\\(q=1;i=2,...,s;k=1).\end{array} \tag{2.38}$$

Equation (2.36) represents the subset of relations which indicates the Pareto – optimal utility frontier and represents the "new welfare economics". The new element added as shown in the set (2.37), which constitutes a pure theory of government expenditure on collective consumption goods. Furthermore, (2.36) and (2.37) themselves define the (s-1)fold infinity of utility frontier points. And when a set of interpersonal normative conditions equivalent to (2.38) is supplied are we able to define an unambiguously "best" state.

- Impossibility of decentralized spontaneous solution

So much involved in the optimizing equation that an omniscient calculating machine could theoretically solve if fed the postulated functions. An analogue calculating machine can be provided by competitive market pricing,

(a) So long as the production functions satisfy the neoclassical assumptions of constant returns to scale and generalized diminishing returns and,

(b) So long as the individuals' indifference curves have regular convexity and, then add

(c) So long as all goods are private, then insert between the right and left hand sides of equation (1) the equality with uniform market prices p_j / p_r and adjoin the budget equations by following,

$$p_1 X_1^i + p_2 X_2^i + \ldots + p_n X_n^i = L^i \quad (i = 1,2,\ldots,s), \tag{2.39}$$

L^i is a lump – sum tax for each individual. If there were no collective goods, then (2.36) and (2.39) have their solution huge simplified. One of reasons is the perfect competition among productive enterprises have to ensure that the goods are produced at the minimum costs and are sold at proper marginal costs. And others is each individual, in seeking as a competitive buyer to get to the highest level of indifference subject to given prices and tax, would be led as if by an invisible hand handle the grand solution of the social maximum point. A necessary thing is political decision making, but a computationally minimum type: namely, algebraic taxes and transfer $(L^1,\ldots,L^s)$ would have to be varied until society is jumped to the ethical observer's optimum. In terms of communication theory and game terminology, each person is motivated to do the signaling of the tastes needed to define and reach the attainable – bliss point. Now all of the above remains valid even if collective consumption is not zero but is instead explicitly set at its optimum values as determined by (2.36), (2.37) and (2.38). However no decentralized pricing system can serve to determine optimally levels of collective consumption.

2.1.6 Theory of Benefit Taxation (David G. Duff)

- Defining benefit taxes and user fees

Although it is possible to discern a benefit rationale for almost any tax, including taxes on income, consumption or wealth (Tomas Hobbes, 1962), tax scholars and policy makers have found it useful to distinguish between general taxes on broad measures of each taxpayer's economic capacity and more targeted charges. The basic general definition can be defined as mandatory levies that are not related to any specific benefit or government service (Richard M. Bird, 1997), whereas benefit taxes and user fees constitute mandatory or voluntary levies imposed on persons deriving particular benefits from specific categories of publicly provided goods and services. The former of benefit taxes and user fees have been defined as compulsory levies applied to individuals or institution who are assumed to benefit as a group from certain government services. The definitions of actual taxes make it possible to categorize many government levies as being of one type or another. For other taxes and levies, however, classification as a general tax, a benefit tax, or a user fee is less clear. Ontario's Employer Health tax suggests categorization as a benefit tax imposed on employers in relation to government- financed health services provided to employees.

- Argument for and against benefit taxes and user fees

Benefit taxation is generally favoured on these three grounds:

(1) benefit taxes and user fees advance economic efficiency by make sure that scarce resources are allocated to their most highly valued uses, both within the public sector and between the public and private sectors (Harry M. Kitchen, 2003); (2) these levies also enhance the accountability of the public sector, make it more responsive to differing preferences and changes in the demand for publicly provided goods and services; (3) benefit taxation embodies a basic principle of fairness, since taxpayers pay only for those publicly provided goods and services that they use. In order to determine the role of benefit taxes and user fees as a method of increasing revenue, this section review argument for and against benefit taxation as well as the expected implications of these levies for government revenues and their political viability.

- Efficiency

Economic efficiency is one of the goals of tax policy and a central objective of public policy more generally. In normative principle, efficiency promotes the allocation of scare resources for the most valued uses in order to maximize aggregate welfare. Applied to the public sector, the suggestion is that the taxes should be imposed only where the value of public goods and services are provided which are financed by taxes (determined by taxpayers' willingness to pay) exceeds of the goods and services that taxpayer could obtain as in the private sector. Knut Wicksell argued that an efficient outcome could be arrived at through sequential voting on packages of expenditures and taxes util the point where near unanimity is achieved. Others have posited bargaining solutions among groups of taxpayers who enter into mutual agreements over the total level of public expenditures and distribution of the taxes needed to finance these expenditures. Furthermore, these theoretical solutions, benefit taxes and user fees represent a more practical way to ensure that the level and mix of public expenditures and the distribution of the taxes to finance these expenditures approach an economically efficient result. More illustrate shown in Figure 1.12;

FIGURE 1

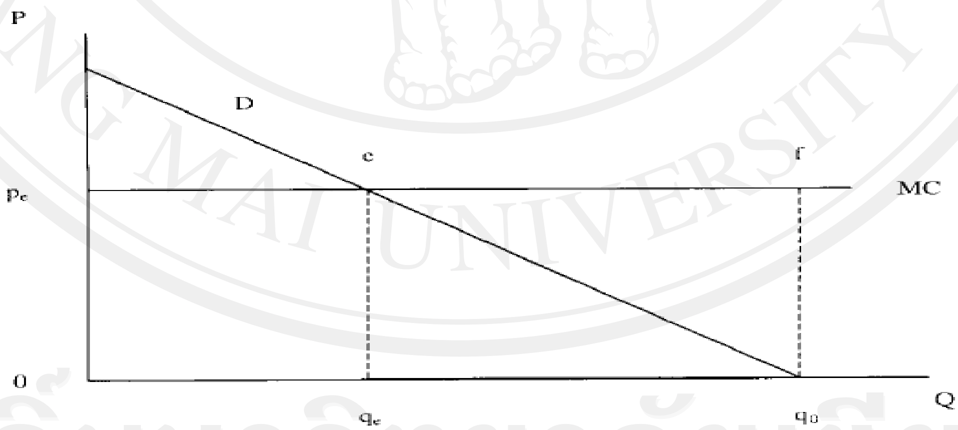

Figure 2.1: The level and mix of public expenditures and the distribution of the taxes

Source: Benefit taxes and user fees in theory and practice

Assume that a constant marginal cost of production (MC) and decreasing demand as the price (p) of the publicly provided good or service increases (D), economic efficiency is achieved at the quantity of the good or service (q_e) that is demanded at a price equal to its marginal cost of production (p = MC). Where the price charged for the good or service exceeds its marginal cost (p > MC), on the other hand, the amount that taxpayers are willing to pay for the good or service exceeds the cost of additional output, suggesting that scarce resources could efficiently be devoted to increased production of the good or service.

Figure 1 also illustrates one of the painful effects of government finance through general taxes. Since the general taxes that individuals pay are not directly linked to their use of publicly provided goods and services, the price that must be paid in order to obtain additional units of a publicly provided good or service is effectively zero, causing individuals to demand more of these goods or services (q_0) than they would be prepared to pay for if they were required to bear the marginal costs of their production directly. The resulting inefficiency, denoting resources that could be employed more efficiently for other purposes, is illustrated by the area ef q_0 .

There are three drawbacks of general taxes compared to benefit taxes and user fees mention to this inefficiency. First, he increasing demand in public goods and services that is attributable to this method of government finance is apt to manifest itself as shortages in the supply of these goods and services, government finance through general taxes can also lead to inefficiently large investments in the government's long-term capacity to supply public goods and services. Second, while benefit taxes and user fees are allocated these limit supplies of public goods and services according to the price that users are willing to pay, reliance on general tax financing is typically associated with rationing by queues, which themselves involve economic costs.

Finally, since general taxes are apt to discourage otherwise efficient economic activities, while benefit taxes and user fees encourage economically efficient 30ongevit, the economic impact of general tax financing on the demand for publicly provided goods and services is compounded by its impact on the supply and demand of privately produced goods and services.

- Fairness

Like economic efficiency, fairness is an established goal of tax policy and a central objective of public policy more generally. In much tax policy analysis, fairness is understood in terms of horizontal equity (the principle that taxpayers with similar levels of economic well-being should pay similar taxes) and vertical equity (the principle that taxpayers who enjoy higher levels of economic well-being should pay appropriately higher taxes). For the most part, these equity concepts are applied to general taxes, with horizontal equity determining the base of the tax (e.g., income, consumption, or wealth) and vertical equity defining applicable rates.

- Distributional impact

As a general rule, lower-income households devote a larger proportion of their incomes than higher-income households to consumption, while higher-income households are able to devote a larger proportion of their incomes than lower-income households to personal saving. To the extent that benefit taxes and user fees apply to the consumption of publicly provided goods and services, therefore, it follows that these levies may impose a heavier financial burden on lower-income households than on higher-income households. For this reason, one of the main criticisms directed at benefit taxes and user fees is that they are likely to be regressive.

- Revenue

In addition to these arguments for and against benefit taxes and use fees, attitudes toward these levies may also turn on their anticipated impact on government revenues. While those advocating smaller government might be expected to welcome benefit taxes and user fees that reduce aggregate revenues collected by government bodies, this support may turn to opposition where these levies facilitate the expansion or introduction of new publicly provided goods or services. Conversely, those who are favourably inclined toward the public sector might support benefit taxes and user fees that finance new or expanded goods or services while rejecting levies that decrease total government revenues. In theory, the impact of benefit taxation on government revenues is wholly indeterminate. On the one hand, as Bird explains, earmarking revenues from benefit taxes and user fees to specific programs ‘may induce the public to support new or increased taxes on which they otherwise could not agree by linking the taxes with the expansion of some government activity which they desire.’ Indeed, these levies are often seen as a way to raise additional revenue from populations increasingly resistant to general taxes. On the other hand, economic analysis suggests that the introduction of benefit taxes and user fees to pay for publicly provided goods and services previously financed by general taxes is apt to reduce public demand for these goods and services at least where this demand is elastic and, hence, the total revenues that are devoted to their production.

2.1.7 Corruption

As it is said in the World Bank Report (1997) that corruption originates at the junction of public sector and private sector, and it will spread on condition of abusive power of policy or management system together with an ineffective binding force. In order to take advantage of corruption and rent-seeking, private interests at home and abroad will exert their influence on it, while the public institution will yield to those corruption roots if the binding force is trustless. Vito Tanzi (1999) believes that the direct cause of corruption comes from the governmental control of social resources, such as the examination and approve of managerial permit, the subjective disposition made by the revenuer and the decision-making authority of the public investment item. Meanwhile, the indirect cause may be the shortcomings of civil servant evaluation system, (such as to employ or to prompt without grading will aggravate the corruption crisis.) the extremely low salary standard (for it will gives the chance to bribe as a compensation), the imperfect punishment mechanism (other things be unaltered, the increasing of punishment will reduce the occurrence of corruption), deficiency of an effective in-house system (such as lack for an honest and efficient supervisor, well auditing bodies, well-designed and clear laws and regulations, and a judicial department that could avoid any influence from different political factions), the senior leader that cannot set a good example to the subordinates or be involved in corruption dealings, or to abide his or her relations' corrupt behavior. However, from the basic theoretical level, the occurrence of corruption originates from these three factors.

- Humanity weakness

The saying that "Gold cannot be pure and man cannot be perfect" is a right attitude and standard of a private. In fact, the whole mankind does also it. From a view of morality, human common weakness is selfish trend, which is born and coexisted with a person’s life. Well, it maybe seems a little panic, but truth sometimes is most cruel. Abraham Maslow, famous person-based psychologist, pointed out that need is human nature, what’s more, needs are divided into a few levels and gradually further forward. Firstly people want to meet physiological needs such as food, housing and sexual desires. Then, on the basis of this, people shall continue to pursue spirit, including the feelings of safety, achievement, respect and so on. Must say, Maslow’s demand level theory has certain scientific nature is also easily accepted by most people. In fact, change an angle, people’s demands are the basic reasons of self-interest tendency’s existence. Now, when mention to Surveillance system construction’s guide-idea or principle, many scholars point out ‘humanity hypothesis’ theory. Namely, human are assumed to be villains. That is the starting point and foothold of surveillance system construction. In fact, this is a self-deceived statement also is an obscure theory. People have self-interest tendency, not hypothesis. It is fact. But ‘self-interest tendency’ is not ‘actually villain’ .If continued to strengthen our self-activation, ‘selfish tendencies’ may be a noble person. If the extend supervisory system is strictly. The selfish tendencies may be law-abiding people. Let All rights are very easily lead to abuse of power. This is unchanged experience. The people have power use it until meet ending.

- Government’s control of social resource

Humanity’s weakness only provides intrinsic motivations or conditions for corruption. External conditions must also be true. Government’s control of social resource is a necessary condition like that. Government is a public agency.

It is to maintain social order and basic tools to protect the rights of all citizens. To achieve its functions, the government must control the part of social resources, such as tax, financial investment, business license or even direct management of large state-owned enterprises. This control ability is public power in another point. The corrupt persons no matter they are officials or other participants. Their main objective is to seek a final material or spiritual interests. However, the corrupt persons have found that benefit from legitimate market must pay the same amount of labor. And though the illegal use of government’s power and government controlled resource, it will have over or even non-pay profit seeking.

- System defects

System defects including the defects of power operation system and the defeats of power supervise system. People have greedy heart. The rights of government's control of social resources provide possibility of corruption. But whether it will eventually be put into action or not depend on opportunities and risks. In this regard, economics Gary Becker, in his "Economic analysis of human" point out-All human behavior can use economics cost-benefit way to be studied and explained, that is none other than to try to use of small cost for the largest gains .Corruption is actually a kind of economic movement. Corrupt elements have their own cost-benefit analysis. If the system has serious defect, the risk of being discovered and punished is low.

2.2 Econometric Theories and Estimation

First of all data must be tested for the stationary by using Panel Unit Root Test. The technical details of unit root and stationarity tests are kept to a minimum. Excellent technical treatments of nonstationary time series may be found in Hamilton (1994), Hatanaka (1995), Fuller (1996) and the many papers by Peter Phillips. Useful surveys on issues associated with unit root testing are given in Stock (1994), Maddala and Kim (1998) and Phillips and Xiao (1998).

2.2.1 Panel Unit Root Test

- Testing for Nonstationarity and Stationarity

To understand the econometric issues associated with unit root and stationarity tests, consider the stylized trend-cycle decomposition of a time series y_t:

$$y_t = TD_t + z_t \tag{2.40}$$

$$TD_t = k + \delta t \tag{2.41}$$

$$z_t = \phi z_{t-1} + \varepsilon_t, \varepsilon_t \sim WN(0, \sigma^2) \tag{2.42}$$

where TD_t is a deterministic linear trend and z_t is an AR(1) process. If $|\phi| < 1$ then y_t is I(0) about the deterministic trend TD_t. If $|\phi| = 1$, then $z_t = z_{t-1} + \varepsilon_t = z_0 + \sum_{j=1}^{t} \varepsilon_j$, a stochastic trend and y_t is I(1) with drift. Simulated I(1) and I(0) data with k=5 and $\delta = 0.1$. The I(0) data with trend follows the trend $TD_t = 5 + 0.1t$ very closely and exhibits trend reversion. In contrast, the I(1) data follows an upward drift but does not necessarily revert to TD_t.

Autoregressive unit root tests are based on testing the null hypothesis that $\phi = 1$ (difference stationary) against the alternative hypothesis that $\phi < 1$ (trend stationary). They are called unit root tests because under the null hypothesis the autoregressive polynomial of $z_t, \phi(z) = (1 - \varphi z) = 0,$ has a unit root equal to unity.

Stationary tests take the null hypothesis that y_t is trend stationary. If y_t is then first differenced it becomes

$$y_t = \delta + \Delta z_t \tag{2.43}$$

$$\Delta z_t = \phi \Delta z_{t-1} + \varepsilon_t - \varepsilon_{t-1} \tag{2.44}$$

Unit root and stationarity test statistics have nonstandard and nonnormal asymptotic distributions under their respective null hypotheses. To complicate matters further, the limiting distributions of the test statistics are affected by the inclusion of deterministic terms in the test regressions.

These distributions are functions of standard Brownian motion (Wiener process), and critical values must be tabulated by simulation techniques. MacKinnon (1996) provides response surface algorithms for determining these critical values, and various S+FinMetrics functions use these algorithms for computing critical values and p-values.

1. Levin-Lin-Chu Test

Individual unit root test have limited power. The power of a test is the probability of rejecting the null hypothesis is unit root. Levin-Lin-Chu Test (LLC) suggest the following hypotheses

H_0: each time series contains a unit root

H_1: each time series is stationary

where the lag order ρ is permitted to vary across individuals. The procedure works as follows:

First, run augmented Dickey – Fuller (ADF) for each cross-section on the equation:

$$\Delta y_{it} = \rho_i y_{i,t-1} + \sum_{L=1}^{\rho_i} \theta_{iL} \Delta y_{it-L} + \alpha_{mi} d_{mt} + \varepsilon_{it} \tag{2.45}$$

Second, run two auxiliary regressions:

1. Δy_{it} on $\Delta y_{i,t-L}$ and d_{mt} to obtain the residuals $\hat{e}_{it}$ and
2. $y_{i,t-1}$ on $\Delta y_{i,t-L}$ and d_{mt} to get residuals $\hat{v}_{i,t-1}$.

The third step is involves standardization of the residuals by performing

$$\tilde{e}_{it} = \hat{e}_{it} / \hat{\sigma}_{\varepsilon i} \tag{2.46}$$

$$\tilde{v}_{i,t-1} = \hat{v}_{it} / \hat{\sigma}_{\varepsilon i} \tag{2.47}$$

where $\sigma_{\varepsilon i}$ denotes the standard error from each ADF.

Finally, run the pooled OLS regression

$$\tilde{e}_{it} = \rho \tilde{v}_{i,t-1} + \tilde{\varepsilon}_{it} \tag{2.48}$$

The null hypothesis is $\rho = 0$. Notice that the standard deviation for t-statistics has to be adjusted (Levin et al., 2002). The necessary condition for the Levin-Lin-Chu test is $\sqrt{N_T} / T \to 0$, while sufficient condition would be $N_T / T \to 0$ and $N_T / T \to k$.

If T is very small, the test is undersized and has low power. The null hypothesis that all cross – sections have a unit root is very restrictive. That is, it does not allow the intermediate case, where some individuals are subject to a unit root and some are not. If T is very large, then suggest individual unit root time-series tests. If N is very large, usual panel data procedures can be applied.

2. Im, Pesaran and Shin Test

The Im-Pesaran-Shin (IPS) test is not as restrictive as the Levin-Lin-Chu test, since it allows for heterogeneous coefficients. The null hypothesis is that all individuals follow a unit root process:

$$H_0 : \rho_i = 0 \forall i$$

The alternative hypothesis allows some of individuals to have unit roots:

$$H_1 : \begin{cases} \rho_i < 0 ... fori = 1,2,..., N_1 \\ \rho_i = 0 ... fori = N_1 + 1,..., N \end{cases}$$

When $t_{\rho i}$ is the individual t-statistic for testing the null hypothesis: $\rho_i = 0$ for all I, then the test is based on averaging individual unit root test $\bar{t} = \frac{1}{N}\sum_{i=1}^{N} t\rho_i$. If this statistic is properly standardized, it is asymptotically $N(0,1)$ distributed. Im-Pesaran-Shin requires $N/T \to 0$ for $N \to \infty$. If either N is small or if N is large relative to T, then both Im-Pesaran-Shin and Levin-Lin-Chu show size distortions.

3. Breitung's Test

The procedure of the Breitung's test can be described as follows. The first step is same as Levin-Lin-Chu test, except that not include deterministic terms. To regress Δy_{it} on $\Delta y_{i,t-L}$ and obtain the residuals $\hat{e}_{it}$. And also run $y_{i,t-1}$ on $\Delta y_{i,t-L}$ and obtain the residuals $\hat{v}_{i,t-1}$. Afterwards, forward orthogonalization transformation is applied to the residuals $\hat{e}_{it}$ such that to obtain $e*_{it}$. Finally, run the pooled regression $e*_{it} = \rho v*_{i,t-1} + \varepsilon *_{it}$, which is asymptotically $N(0,1)$ distributed.

4. Combining p-Value Tests

Fisher – type Test

The Fisher – type test uses p-values from unit root tests for each cross-section i.

The formula of the test looks as follows:

$$P = -2\sum_{i=1}^{N} \ln p_i$$

The test is asymptotically chi-square distributed with $2N$ degrees of freedom ($T_i \to \infty$ for finite N). This test can handle unbalanced panels. Furthermore, the lag lengths of the individual augmented Dickey-Fuller tests are allowed to differ.

5. Further Tests and Properties

There are several other combining p-values tests, in particular, the inverse normal test Z (standard normal distribution), the logit test L (logistic distribution), and the modified Fisher – type test when N is large. All of these tests share certain advantages:

- The number of cross-section observations N can be finite or infinite.
- Each individual allows for different types of non-stochastic and stochastic components.
- The time dimension T can vary for each individual.
- The power is superior to both the Levin-Lin-Chu test and Im-Perasan-Shin test.

6. Residual – Based LM Test

Hadri (2000) proposes a test builds on the Kwaitkowski-Phillips-Schmidt-Shin test (KPSS) from time series testing. KPSS tests are used for testing the null hypothesis that an observable time series is stationary around a deterministic trend. The Hadri test is based on OLS residuals obtained from regressing y_{it} on a constant or a constant plus trend.

The null hypothesis is that there is no unit root in any series. The alternative hypothesis is that the panel has a unit root.

$$y_{it} = \gamma_{it} + \varepsilon_{it} \tag{2.49}$$

$$\gamma_{it} = \gamma_{i,t-1} + u_{it} \tag{2.50}$$

$$H_0 : \sigma_u^2 = 0 \tag{2.51}$$

If variance u_{it} is zero, then γ_{it} becomes a constant and thus y_{it} is stationary. The Hadri test allows for heteroskedasticity adjustments. Its empirical size is close to its normal size if N and T are large.

2.2.2 Simultaneous Equations Models

The emphasis in simultaneous equation models is on situations where two or more variables are jointly determined by a system of equations. Nevertheless, the population model, the identification analysis, and the estimation methods apply to a much broader range of problems.

A system of linear simultaneous equations for the population can be written as:

$$
\begin{aligned}
&y_1 = y_1\gamma_1 + z_1\delta_1 + u_1 \\
&\vdots \\
&y_G = y_G\gamma_G + z_G\delta_G + u_G
\end{aligned}
\qquad (2.52)
$$

where y_h is $1\times G_h$, γ_h is $G_h \times 1$, z_h is $1\times M_h$, and δ_h is $M_h \times 1$, $h = 1,2,...,G.$ These are structural equations for endogenous variables $y_1, y_2,..., y_G$.

The vector y_h denotes endogenous variables that appear on the right-hand side of the hth structural equation. By convention, y_h can contain any of the endogenous variables $y_1, y_2,..., y_G$ except for y_h. The variables z_h are the exogenous variables appearing in equation h. Usually there is some overlap in the exogenous variables across different equations. The restrictions imposed in system (2.52) are called exclusion restrictions because certain endogenous and exogenous variables are excluded from some equations. The $1\times M$ vector of all exogenous variables z is assumed to satisfy

$$E(z'u_g) = 0, \qquad g = 1,2,...,G \qquad (2.53)$$

When all of the equations in system (1) are truly structural, assume that

$$E(u_g|z) = 0, \qquad g = 1,2,...,G \qquad (2.54)$$

However, the assumption (2.53) is sufficient for consistent estimation. Sometimes, especially in omitted variables and measurement error applications, one or more of equations in system (2.52) will simply represent a linear projection onto exogenous variables.

Assumption (2.53) implies that the exogenous variables appearing anywhere in the system are orthogonal to all the structural errors. If some elements in z_1 do not appear in the second equation, then assuming that they do not enter the structural equation for y_2. If there are no reasonable exclusion restrictions in an SEM, it may be that the system fails the autonomy requirement.

Generally, in the system (2.52), the error u_g in equation g will be correlated with y_g, and so OLS and GLS will be inconsist.

2.2.3 Generalized Method of Moments Estimator (GMM Estimator)

Definition of the GMM estimator of δ in

$$y_t = z_t'\delta_0 + \varepsilon_t, \qquad t = 1,\dots,n \tag{2.55}$$

is constructed by exploiting the orthogonality conditions

$$E[g_t(w_t,\delta_0)] = E[x_t\varepsilon_t] = E[x_t(y_t - z_t'\delta_0)] = 0. \tag{2.56}$$

The idea is to create a set of estimating equations for δ by making sample moments match the population moments defined by (2.56). the sample moments based on (2.56) for an arbitrary value are

$$g_n()\delta = \frac{1}{n}\sum_{t=1}^{n} g(w_t,\delta) = \frac{1}{n}\sum_{t=1}^{n} x_t(y - z_t'\delta)$$

$$= \begin{pmatrix} \frac{1}{n}\sum_{t=1}^{n} x_{1t}(y - z_t'\delta) \\ \vdots \\ \frac{1}{n}\sum_{t=1}^{n} x_{Kt}(y - z_t'\delta) \end{pmatrix}$$

These moment conditions are a set of K linear equations in L unknowns. Equating these sample moments to the population moment $E[x_t\varepsilon_t] = 0$ gives the estimating equations

$$S_{xy} - S_{xy}\delta = 0 \tag{2.57}$$

Where $S_{xy} = n^{-1}\sum_{t=1}^{n} x_t y_t$ and $S_{xz} = n^{-1}\sum_{t=1}^{n} x_t z_t'$ are the sample moments If $K = L$ (δ_0 is just identified) and S_{xy} is invertible then the GMM estimator of δ is

$$\hat{\delta} = S_{xz}{}^{-1} S_{xy}$$

which is also known as the indirect least squares estimator. If $K > L$ then there may not be a solution to the estimating equations (2.57). In this case, the idea is to try to find δ that makes $S_{xy} - S_{xz}\delta$ as close to zero as possible. To do this, let $\hat{W}$

denote a $K \times K$ symmetric and positive definite weight matrix, possibly dependent on the data, such that $\hat{W} \xrightarrow{p} W$ as $n \to \infty$ with W symmetric and p.d then the GMM estimator of δ, denoted $\hat{\delta}(\hat{W})$, is defined as

$$\hat{\delta}(\hat{W}) = \arg\min_{\delta} J(\delta, \hat{W})$$

where

$$\begin{aligned} J(\delta, \hat{W}) &= n g_n(\delta)' \hat{W} g_n(\delta) \\ &= n(S_{xy} - S_{xz}\delta)' \hat{W} (S_{xy} - S_{xz}\delta) \end{aligned} \qquad (2.58)$$

Since $J(\delta, \hat{W})$ is a simple quadratic form in δ, straightforward calculus may be used to determine the analytic solution for $\hat{\delta}(\hat{W})$:

$$\hat{\delta}(\hat{W}) = (S'_{xz} \hat{W} S_{xz})^{-1} S'_{xz} \hat{W} S_{xy} \qquad (2.59)$$

Asymptotic Properties

Under standard regularity conditions, it can be shown that

$$\hat{\delta}(\hat{W}) \xrightarrow{p} \delta_0$$

$$\sqrt{n}(\hat{\delta}(\hat{W}) - \delta_0) \xrightarrow{d} N(0, a\,\mathrm{var}(\hat{\delta}(\hat{w})))$$

where

$$a\,\mathrm{var}(\hat{\delta}(\hat{W})) = (\Sigma'_{xz} W \Sigma_{xz})^{-1} \Sigma'_{xz} WSW \Sigma_{xz} (\Sigma'_{xz} W \Sigma_{xz})^{-1} \qquad (2.60)$$

A consistent estimate of $a\,\mathrm{var}(\hat{\delta}(\hat{W}))$, denoted $a\widehat{\mathrm{var}}(\hat{\delta}(\hat{W}))$, may be computed using

$$a\,\widehat{\mathrm{var}}(\hat{\delta}(\hat{W})) = (S'_{xz} \hat{W} S_{xz})^{-1} S'_{xz} \hat{W} \hat{S} \hat{W} S_{xz} (S'_{xz} \hat{W} S_{xz})^{-1} \qquad (2.61)$$

where $\hat{S}$ is a consistent estimate for $S = a\,\mathrm{var}(\bar{g})$.

- Multiple – Equations GMM Estimator

Suppose there are m equations, the system can be written as follows:

$$y_{1i} = X'_{1,i}\delta_1 + u_{1i},$$

$$\ldots$$

$$y_{mi} = X'_{m,i}\delta_m + u_{mi}$$

where for all $j = 1,...,m, \delta_j \in R^{k_j}, k_j = m_j + l_j$, and the random $l - vectorZ_t$ is such that

$$rank(EZ_i X'_{j,i}) = k_j,$$

$$EZ_i u_{ji} = 0.$$

Equivalently, the system can be re-written in the matrix notation as

$$y_1 = X_1\delta_1 + u_1,$$

$$\ldots$$

$$y_m = X_m\delta_m + u_m,$$

where, for $j = 1,...,m, X_j$ collects the n observations on the right – hand side variables in the j - th equation:

$$X_j = \begin{pmatrix} X'_{j,1} \\ \vdots \\ X'_{j,n} \end{pmatrix},$$

y_i collects the n observations on the left – hand side variable in the j - th equation:

$$y_i = \begin{pmatrix} y_{j1} \\ \vdots \\ y_{jn} \end{pmatrix},$$

and u_j is defined similarly. Further, define

$$X = \begin{pmatrix} X_1 & \cdots & 0 \\ 0 & & X_m \end{pmatrix},$$

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

$$U = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix},$$

$$\delta = \begin{pmatrix} \delta_1 \\ \vdots \\ \delta_m \end{pmatrix}.$$

Note that X is $(nm) \times k$, where $k = k_1 + ... + k_m$, Y and U are $(nm) \times 1$, and δ is $k \times 1$. The system can now be compactly written as $Y = X\delta + U$.

In this system the ml population moment conditions as follows:

$$E\begin{pmatrix} Z_i u_{1i} \\ \vdots \\ Z_i u_{mi} \end{pmatrix} = E(I_m \otimes Z_i)U_i = 0,$$

where $U_i = (u_{1i}, ..., u_{mi})'$, and $A \otimes B$ denotes the Kronecker product of A and B. To define sample moment conditions that can be used for estimation, consider the ml-vector of sample correlations between the exogenous variables and errors:

$$(I_m \otimes Z)' U = (I_m \otimes Z)' (Y - X\delta),$$

where Z denotes the $n \times 1$ matrix of observations on the exogenous variables. Let A_n be and $(ml) \times (ml)$ weight matrix. The system or multiple – equation GMM estimator is obtained by solving

$$\min_{d \in R^k} (Y - Xd)'(I_m \otimes Z) A_n' A_n (I_m \otimes Z)'(Y - Xd).$$

Thus, the system GMM estimator is given by:

$$\hat{\delta} = (X'(I_m \otimes Z)A_n'A_n(I_m \otimes Z)'X)^{-1} X'(I_m \otimes Z)A_n'A_n(I_m \otimes Z)'Y.$$

Define

$$W_n = A_n'A_n,$$

and introduce the partition

$$W_n = \begin{pmatrix} W_{11,n} & \cdots & W_{1m,n} \\ \cdots & \cdots & \cdots \\ W_{m1,n} & \cdots & W_{mm,n} \end{pmatrix},$$

Where each $W_{ij,n}$ is an $l \times l$ systematic matrix. The system GMM estimators for the m equations can be written as

$$\begin{pmatrix} \hat{\delta}_1 \\ \vdots \\ \hat{\delta}_m \end{pmatrix} = \left(\begin{pmatrix} X_1'Z & \cdots & 0 \\ 0 & & X_m'Z \end{pmatrix} \begin{pmatrix} W_{11,n} & \cdots & W_{1m,n} \\ \cdots & \cdots & \cdots \\ W_{m1,n} & \cdots & W_{mm,n} \end{pmatrix} \begin{pmatrix} Z'X_1 & \cdots & 0 \\ 0 & & Z'X_m \end{pmatrix} \right)^{-1} \quad (2.62)$$

$$\times \begin{pmatrix} X_1'Z & \cdots & 0 \\ 0 & & X_m'Z \end{pmatrix} \begin{pmatrix} W_{11,n} & \cdots & W_{1m,n} \\ \cdots & \cdots & \cdots \\ W_{m1,n} & \cdots & W_{mm,n} \end{pmatrix} \begin{pmatrix} Z'y_1 \\ \vdots \\ Z'y_m \end{pmatrix}$$

$$= \begin{pmatrix} X_1'ZW_{11,n}Z'X_1 & \cdots & X_1'ZW_{1m,n}Z'X_m \\ \cdots & \cdots & \cdots \\ X_m'ZW_{m1,n}Z'X_1 & \cdots & X_m'ZW_{mm,n}Z'X_m \end{pmatrix}^{-1} \begin{pmatrix} X_1'ZW_{11,n}Z'y_1 + \ldots + X_1'ZW_{1m,n}Z'y_m \\ \cdots \\ X_m'ZW_{m1,n}Z'y_1 + \ldots + X_m'ZW_{mm,n}Z'y_m \end{pmatrix}.$$

We can compare the above expression with that for equation-by-equation GMM:

$$\begin{pmatrix} \tilde{\delta}_1 \\ \vdots \\ \tilde{\delta}_m \end{pmatrix} = \begin{pmatrix} X_1'ZA_{1n}'A_{1n}Z'X_1 & \cdots & 0 \\ 0 & & X_m'ZA_{mn}'A_{mn}Z'X_m \end{pmatrix}^{-1} \begin{pmatrix} X_1'ZA_{1n}'A_{1n}Z'y_1 \\ \cdots \\ x_M'ZA_{mn}'A_{mn}Z'y_m \end{pmatrix}.$$

From the comparison, it is apparent that the equation-by-equation GMM estimator is a particular case of the system GMM estimator with weighting matrices $W_{ij,n} = 0$ for $i \neq j$.

- Large-sample properties of the multiple – equation GMM estimator

From (3.34), can re-write as follows:

$$\begin{pmatrix} \hat{\delta}_1-\delta_1 \\ \vdots \\ \hat{\delta}_m-\delta_m \end{pmatrix} = \left(\begin{pmatrix} \sum_{i=1}^{n} X_{1,i}Z_i' & \cdots & 0 \\ 0 & & \sum_{i=1}^{n} X_{m,i}Z_i' \end{pmatrix} \begin{pmatrix} W_{11,n} & \cdots & W_{1m,n} \\ \cdots & \cdots & \cdots \\ W_{m1,n} & \cdots & W_{mm,n} \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} Z_i X_{1,i}' & \cdots & 0 \\ 0 & & \sum_{i=1}^{n} Z_i X_{m,i}' \end{pmatrix} \right)^{-1}$$

$$\times \begin{pmatrix} \sum_{i=1}^{n} X_{1,i}Z_i' & \cdots & 0 \\ 0 & & \sum_{i=1}^{n} X_{m,i}Z_i' \end{pmatrix} \begin{pmatrix} W_{11,n} & \cdots & W_{1m,n} \\ \cdots & \cdots & \cdots \\ W_{m1,n} & \cdots & W_{mm,n} \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n} Z_i u_{1i} \\ \vdots \\ \sum_{i=1}^{n} Z_i u_{mi} \end{pmatrix}. \quad (2.63)$$

As the assumptions

- $\{(Y_i', Z_i') : i \geq 1\}$ are iid.
- $W_n \to_p W$ positive definite (and symmetric).
- The elements of $U_i = (u_{1i}, ..., u_{mi})'$ and Z_i have finite second moments which together with the reduced form equations implies that $EZ_i X_{j,i}'$ is finite for all $j = 1,...,m$.

Under these assumptions there is consistency of the system GMM estimator: $\hat{\delta}_j \to_p \delta_j$ for all $j = 1,...,m$. Next, for asymptotic normality we also assume that:

- The elements of U_i and Z_i have finite fourth moments which implies that the elements of Y_i have finite fourth moments as well.

Under these assumptions,

$$n^{-1/2} \begin{pmatrix} \sum_{i=1}^{n} Z_i u_{1i} \\ \vdots \\ \sum_{i=1}^{n} Z_i u_{mi} \end{pmatrix} \to_d N(0, \Omega),$$

where

$$\Omega = E\begin{pmatrix} Z_i u_{1i} \\ \vdots \\ Z_i u_{mi} \end{pmatrix}\begin{pmatrix} Z_i u_{1i} \\ \vdots \\ Z_i u_{mi} \end{pmatrix}'$$

$$= E(U_i \otimes Z_i)(U_i \otimes Z_i)'$$
$$= E(U_i U_i' \otimes Z_i Z_i').$$

Then, the form of equation will be

$$n^{1/2}\begin{pmatrix} \hat{\delta}_1 - \delta_1 \\ \vdots \\ \hat{\delta}_m - \delta_m \end{pmatrix} \rightarrow_d N(0, V(W)),$$

where

$$V(W) = (C'WC)^{-1} C'W\Omega WC(C'WC)^{-1},$$

$$C = \begin{pmatrix} Q_1 & \cdots & 0 \\ 0 & & Q_m \end{pmatrix},$$

$$Q_j = EZ_i X'_{j,i}, \text{ for } j = 1,\ldots,m.$$

Assume that

- Ω is positive definite.

As usual, the efficient GMM estimator corresponds to W_n that satisfies

$$W_n \rightarrow_p \Omega^{-1}.$$

$$W_n = \hat{\Omega}^{-1}{}_n$$

$$= \left(n^{-1} \sum_{i=1}^{n} (\hat{U}_i \hat{U}_i' \otimes Z_i Z_i') \right)^{-1},$$

where $\hat{U}_i$ is constructed using some preliminary consistent estimators of δ_j's,

The asymptotic variance of the efficient GMM estimator is given by

$$V\left(\Omega^{-1}\right) = (C'\Omega^{-1} C)^{-1}.$$

- The J – Statistic

The J – Statistic, introduced in Hansan (1982), refers to the value of the GMM objective function evaluated using an efficient GMM estimator:

$$J = J(\hat{\delta}(\hat{S}^{-1}), \hat{S}^{-1}) = ng_n(\hat{\delta}(\hat{S}^{-1}))'\hat{S}^{-1}g_n(\hat{\delta}(\hat{S}^{-1})) \qquad (2.64)$$

where $\hat{\delta}(\hat{S}^{-1})$ denotes any efficient GMM estimator of δ and $\hat{S}$ is a consistent estimate of S. If $K = L$ then $J = 0$, and if $K > L$ then $J > 0$. Under regularity condition and if the moment conditions are valid, then as $n \to \infty$

$$J \xrightarrow{d} \chi^2(K - L)$$

Hence, in a well specified overidentified model with valid moment conditions the J – Statistic behaves like a chi – square random variable with degrees of freedom equal to the number of overidentifying restrictions. If model is mis-specified and or some of the moment conditions do not hold then the J – Statistic will be large relative to a chi-square random variable with $K - L$ degrees of freedom.

The J – Statistic acts as an omnibus test statistic for model mis-specification. A large J – Statistic indicates a mis-specified model.

- Normalized Moments

If the model is rejected by the J – statistic, to aid in the diagnosis of model failure, the magnitudes of the individual elements of the normalized moments $\sqrt{n}g_n(\hat{\delta}(\hat{S}^{-1}))$. Under the null hypothesis that the model is correct and orthogonality conditions are valid, the normalized moments satisfy

$$\sqrt{n}g_n(\hat{\delta}(\hat{S}^{-1})) \xrightarrow{d} N(0, S - \Sigma_{xz}[\Sigma'_{xz} S^{-1} \Sigma_{xz}]^{-1} \Sigma'_{xz})$$

As a result, for a well specified model the individual moment t-ratios

$$t_i = g_n(\hat{\delta}(\hat{S}^{-1}))_i / SE(g_n(\hat{\delta}(\hat{S}^{-1}))_i), \quad i = 1, \ldots, K \qquad (2.65)$$

where

$$SE(g_n(\hat{\delta}(\hat{S}^{-1}))_i = \left(\left[\hat{S} - \hat{\Sigma}_{xz} [\hat{\Sigma}'_{xz}\, \hat{S}^{-1}\, \hat{\Sigma}_{xz}]^{-1}\, \hat{\Sigma}'_{xz} \right] / T \right)_{ii}^{1/2}$$

are asymptotically standard normal. When the model is rejected using the J – Statistic, a large value of t_i indicates mis-specification with respect to the i th moment condition. Since the rank of $S - \sum_{xz}[\sum'_{xz} S^{-1} \sum_{xz}]^{-1} \sum'_{xz}$ is $K - L$, the interpretation of the moment t-ratios (2.64). in particular, if $K - L = 1$ then $t_1 = ... = t_k$.

2.3 Definitions

2.3.1 Human Development Index and Economics Variables

There are three basic parts which are the components for human development; health, education, and income. The measurement of human development launched by UNDP is the Human Development Index (HDI). This index comprises these three parts to indicate the quality of life. Here we will look at the relevance of human development and the individual components of the HDI

- Human Development Index

The HDI is computed by from data representing the measurement of health, education, and income. For the income component a log transformation is applied, in effect discounting higher incomes due to supposed diminishing marginal utility. For the educational component the transformed variable consists of two-thirds of the percentage rate of literate adults among all adults and one-third of the combined first, second, and third-level educational gross enrolment ratio in percentage. The health or longevity component is directly measured by life expectancy at birth (Neumayer, 2001).

Then the formulate of HDI is computed as below:

$$Xindex = \frac{(actualvalue - \min value)}{(\max value - mixvalue)},$$

X=(Income, Longevity, Education).

This index is calculated for each variable. Until 1994 the minimum and maximum values were not absolutely fixed but derived from the minimum and maximum achieved values worldwide each year. Then the methodology was changed and the minimum and maximum values became fixed in absolute terms. However, since the maximum values are chosen such that they are higher than or equal to the actual value a country can possibly achieve every country's index for each variable lies between zero and one. A country's HDI is then simply the arithmetic average of its three indexes:

$$HDI = \frac{1}{3}(Incomeindex) + (ongivityindex) + (Educationindex)$$

The closer the HDI is to zero, the less developed the country and the lower its population's standard of living will be.

- Health (Longevity)

Healthier people can enjoy longer and better lives. The advances in medicine and progress in technology lead to reductions in infant mortality rates. This has resulted in reducing the risk of losing human capital, a main factor for growth.

One conclusion from previous studies is improvement in health leads to economic growth, for example reductions in serious diseases such as malaria increase the growth rate per capita (Bloom and Sachs 1998, WHO 2001, Canning 2005, and Wacziarg 2005)

One variable that can represent health is life expectancy at birth. This is the expected number of years of life that an average newborn infant would enjoy to live in the future.

Figure 2.2: Trend of Life expectancy at birth (years)

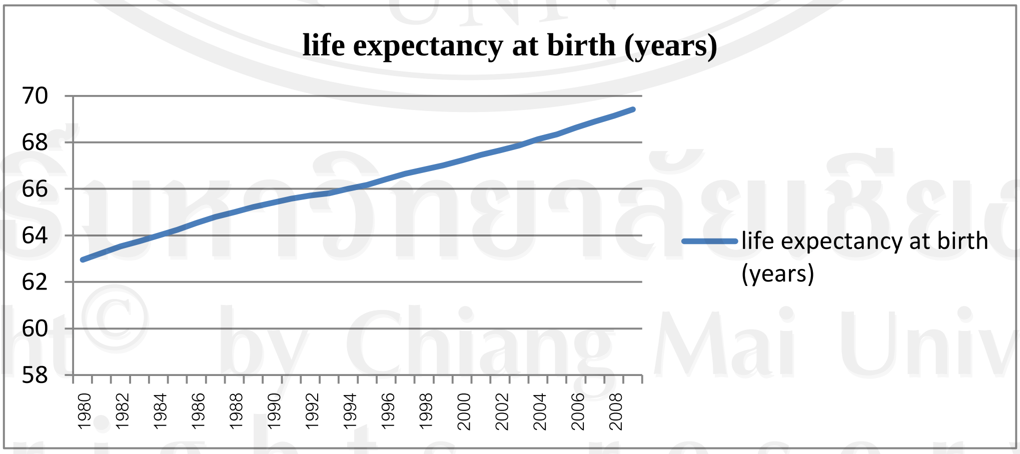

Figure 2.2: Trend of Life expectancy at birth (years)

Source: World Bank (2010)

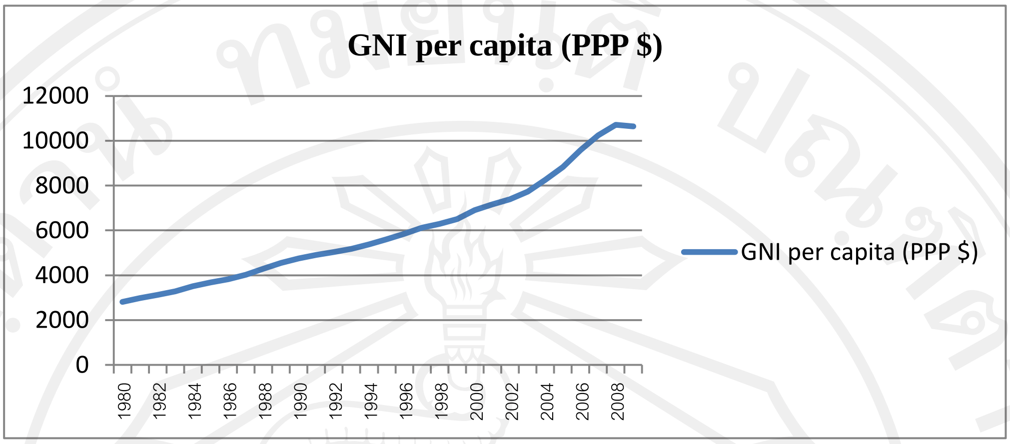

Figure 2.3: Trend of Gross National Income per capita (PPP $)

Source: World Bank (2010)

According to World Bank data, trends of life expectancy at birth and GNI per capita are growing simultaneously. Many studies provide evidence that higher life expectancy relates to higher economic growth and has a positive effect on capital accumulation (World Bank 1998, Bhargava et al. 2001, Chakraborty 2004). But when considering the impact of life expectancy on economic growth with a demographic transition it was found that an increase in pre – transitional countries life expectancies or reductions in their mortality rate tends to reduce per capita income (Cervellati and Sunde 2009).

- Education

As education is one component of human development, the basic concept of human development is enlarging people's choices. A person who has a higher level of education will have an advantage getting a better job. This can imply that an investment in education as knowledge accumulation will give a higher rate of return in the future in order to have a better life in the future with good job and high rate of income (Todaro et. Al).

Rate of educational attainment can indicate investment in education and knowledge accumulation in human capital. As human capital is a determinant factor of economic growth so, more knowledge accumulated in human capital will be an advantage. Many studies found that improving the quality of education and increasing educational attainment contribute to economic growth (Romer 1986 and 1990, Lucas 1988, Barro 1997, Deniz and Dogruel 2008).

- Income per capita (PPP $)

Per capita income is one measurement of economic growth, rises in per capita income persistently represents growth. There is a link between economic growth and human development. Human development and economic growth have a two – way relationship (Rains, Stewart 2000).

One way economic growth effects human development is when a person's income increases they can afford more goods and services and attains some important improvements to their lives such as higher education.

Likewise human development can effect economic growth, as a well educated highly skilled workforce can stimulate economy growth as human capital is an important factor of the production process. As demonstrated in many theories of human capital, human capital accumulation and knowledge accumulated in human capital leads to economic growth through technological progress (Solow's model, New growth theory).

- Tax revenue (percentage of GDP)

A huge part of a government's revenue comes from taxation. There are several kinds of taxes but the main propose of taxation is to increase government revenue in order to finance the government expenditures and finance government projects (Guff D., Van Den Berg 2001). Even though some studies find that an increase in tax rate will reduce the annual growth rate and distort economic activity (Myles D., Lee and Gordon 2004, Feldstein 2006). But if the amount of tax revenue is used in an efficient way to promote social welfare and long term development projects, there will be a positive effect the future trend.

There is an interesting relationship between foreign aid and economic growth. Foreign aid has an impact on growth, the evidences from the past such as Burnside and Dollar 2000, Easterly, Levine and Roodman 2004, Doucouliagos and Paldam 2005, that aid transferred effects the growth of recipient countries. But aid effectiveness depends on some conditions, such as the recipient countries must be democratic (Bearce 2008), have good policy and a high quality government. Countries meeting these conditions will see a positive effect on economic growth.

2.3.2 Corruption

Corruption, commonly defined as illegal activities or a behavior of rent (seeking for own benefit). Corruption can occur in private firms, government, bureaucrats, and international organizations.

Nowadays, many countries around the world have a policy to reduce their corruption rate, but the problem of high corruption rates still remain.

Corruption is a chronic problem as abarrier for economic growth. Many studies find that corruption has a negative impact on growth and corruption can reduce investment growth and also reduce public return of investments (Tanzi and Davoodi 1997, Batra et al. 2003, Aidt, Toke, Dutta, Jayasri, and Sena 2007). Thus, in terms of corruption on public investment and government projects which propose to raise the standard of living through public spending. These will not effective if there is corruption behind them. This will lead to a reduction in investment growth and the reduction in rate of return will erode the economic growth as well.

- Corruption Perception Index

What is Corruption Perception Index? Transparency International (TI) defines corruption as the abuse of entrusted power for private gain. This definition encompasses corrupt practices in both the public and private sectors. The Corruption Perceptions Index (CPI) ranks countries according to perception of corruption in the public sector. The CPI is an aggregate indicator that combines different sources of information about corruption, making it possible to compare countries.

The CPI draws on different assessments and business opinion surveys carried out by independent and reputable institutions. It captures information about the administrative and political aspects of corruption. Broadly speaking, the surveys and assessments used to compile the index include questions relating to bribery of public officials, kickbacks in public procurement, embezzlement of public funds, and questions that probe the strength and effectiveness of public sector anti-corruption efforts.

For a country or territory to be included in the index a minimum of three of the sources that TI uses must assess that country. Thus inclusion in the index depends solely on the availability of information.

Perceptions are used because corruption, whether frequency or amount, is to a great extent a hidden activity that is difficult to measure. Over time, perceptions have proved to be a reliable estimate of corruption. Measuring scandals, investigations or prosecutions, while offering non-perception data, reflect less on the prevalence of corruption in a country and more on other factors, such as freedom of the press or the efficiency of the judicial system. TI considers it of critical importance to measure both corruption and integrity, and to do so in the public and private sectors at global, national and local levels. The CPI is therefore one of many TI measurement tools that serve the fight against corruption.

The CPI values between zero to ten. Zero being very high corruption and ten being very low corruption. As shown in the figure 3 below,

Figure 2.4: Example of CPI

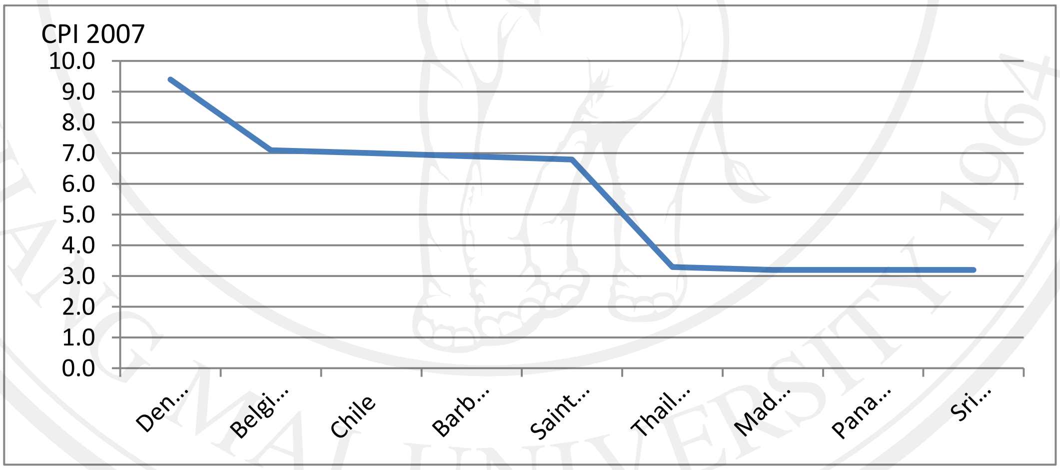

Figure 2.4: Example of CPI

Source: Transparency International (2007)

In these countries, for the sample of CPI, Denmark has the least corruption, while Sri Lanka has the highest corruption rate.

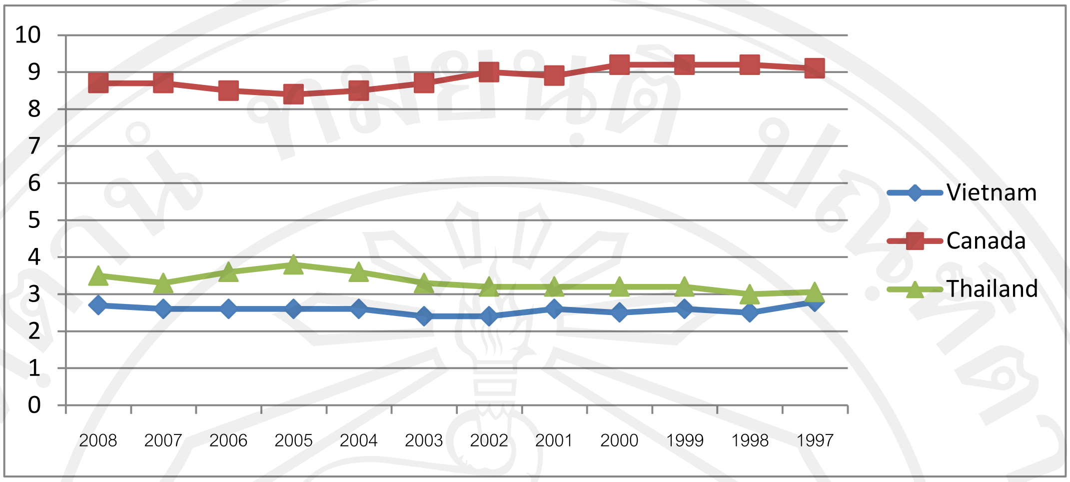

Figure 2.5: Trends of CPI in Vietnam, Canada, and Thailand (1997-2008)

Source: Transparency International

As shown in the graph trends of CPI are varying over time. This may imply that in some countries corruption has reduced, even though there have been many anti-corruption campaigns.

2.4 Literature Review

MauRo (1995) this paper analyzes about assembled data sets consisting of subjective indices of corruption, red tape, the efficiency of the judicial system and various categories of political stability for a cross section of countries.

Jones (1996) have studied about the human capital considered from the variables such as school enrolment. Then he interpreted that the investment rate was higher than capital accumulation. His study adopted Regression estimation to analyse his assumption and model. As Barro and Lee (1993) explained that gross school enrolment indicates human capital accumulation or can be used as a measurement with human capital. Human capital which has improved in skill and practical work relates to important capital in order to drive economic grow further.

Tanzi and Davoodi (1997) they studied about corruption and public investment effect on economic growth by using cross - country data and regression analysis. The corruption problem distorts the decision - making process relating to public investment projects. His hypothesis is that corruption increases the number of projects in the country. And he found that corruption will increase in the share of public investment, make a fall in average productivity of the investment, and make a reduction in some other of public spending. These effects of corrupt slow down economic growth. The conclusion of his study, firstly corruption can reduce growth by increasing public investment while reducing productivity of the investment. Secondly corruption reduces growth by increasing public investment that is not accompanied by its current expenditure. Thirdly corruption reduces the quality of existing infrastructure and finally corruption lowers government revenue used to finance productive projects.

Barro (1998) has studied economic growth and investment by using panel data of 100 countries observed from 1960 to 1995 and holding fixed measures of government policies and institutions and the character of the national population. His hypothesis is growth has positive related to the starting level of average years of school attainment of adult males at the secondary and higher levels. He found that human capital would be important determinant of growth and his study confirms this link. The effect of the school - attainment variable on the growth rate is significant when holding constant the investment ratio and the fertility rate. Countries that start with a higher level of education attainment grow faster than others. The channels of effect involve the positive effect of human capital on physical investment, the negative effect of human capital on fertility, and an additional positive effect on growth for given values of investment and fertility.

Dearden (2000) he studied about the economic costs of corruption. One approach of the study emphasized cost of diversion of resources from productive activity to avoid predation. The alternative approach has focused on the impact of bureaucratic inefficiency upon investment and growth by following the endogenous growth models which in efficiency could reduce investment or could lead to the misallocation of resources. The framework of his study is corruption can either increase or decrease the production costs faced by a firm or individual. The result of corruption in tax evasion reduce government tax revenues but may also lead to increased government costs through the allocation of contracts to higher priced contractors. Moreover, economic costs may be even more substantial as projects that can be undertaken merely to generate bribes rather than as a reflection of economic development.

Wilhelm and Fiestas (2005) this is the working paper of The World Bank Institute. They studied about the impact of public spending on growth and poverty reduction. This study mentioned that government spending is driven by the objective to positively affect growth and poverty reduction result from improving provision of social services. The public non – spending interventions affects the relationship between spending and outcomes which can affect the level of service provisions to the poor. From the hypothesis growth and spending has a positive two-way relationship they explained that growth and spending can occur in both directions that are the higher growth leads to better sectoral outcomes and the improved sectoral services and goods leads to poverty reduction, but higher poverty reduction may require the improvement in sectoral outcomes. Furthermore, the impact of public expenditures on income growth is subject to lags as outcomes may be achieved in a direct or indirect way. Anyway, the study is based on the lessons from the 1990s where public expenditures of central government decreased slightly. They had explored the data set in nine countries and found three countries where government expenditures fell significantly. They used the variables to analyse the median share of education expenditures, health expenditures as a share of total expenditures, agriculture and social security expenditures as a share of total expenditures. In summary they found that overall spending levels declined over the analyzed period.

Haque (2007) his study is about the growth effects in the role of public spending which has the corruption behind it. He assumed that, corruption increases public investment, and corruption reduces the returns to public investment and has a negative impact on economic growth.

He used the methodology of three stage least squares and gives the interdependency of corruption, public investment and growth. Analysis using the simultaneous equation system is valueable from the policy perspective and focus on medium growth impacts of public investment and corruption. The data set has 192 observations covering 58 countries including Europe, Americas, Africa, Asia and Oceania. The results in regression found that there is an interesting relationship between public investment and corruption affect on investment rate. The significant crowding out of private investment by public investment is at the 10 percent level. Implying that the public sector increase the price of goods and services that effectively turn marginal private investment unprofitable. On the other hand, corruption tends to raise the level of private investment in the economy.

Aidt (2008) he studied the role of political accountability determining corruption and economic growth. The model identifies two governance regimes by showing the quality of political institutions and showing the relationship between corruption and growth. He used a threshold model to estimate the impact of corruption on growth where corruption is an endogenous variable. He found that in a regime with high quality political institutions, corruption has a negative impact on economic growth. In a regime with low quality institutions, corruption has no impact on growth. Moreover, corruption is insignificant in all specifications. All of the results are similarly with Mauro (1995) who also finds that corruption is weakly related to growth in a linear growth model.