Human Capital Accumulation, Positive Externality and Optimal Taxation

Yulei Peng
Department of Economics
Texas A&M University

Abstract

I consider the optimal taxation issue in a dynamic economy with human capital accumulation, and agents’ ability is heterogeneous and private information. Moreover, the agents with higher ability have positive external effects on others. By using the two-sector endogenous model, this paper finds out that it is optimal to impose different income and capital income taxes on people with different abilities. Specially, positive marginal income tax is adopted for people with lower ability while no tax is imposed for people with higher ability; on the other hand, marginal capital income tax is zero whether the agent is low ability person or not. Furthermore, as for people using the capital and labor for human capital accumulation, the government should subsidize them whatever their ability is.

Key words: Optimal taxation, Human capital accumulation, Knowledge spillovers
1. Introduction

Lucas (2004) has pointed out that the origins of the modern economic world can be seen, in part, as a result of human and physical capital accumulation, in the transition from a traditional agricultural society to a society of industrialization.

Retrospecting the history of economic development, we could found out the following facts: the process of industrialization is accompanied with urbanization. During the proceeding, the rate of population living in cities is increasing and the share of the workforce in agriculture is decreasing dramatically. In 1850, the farmers made up 64% of labor force in United States, and the proportion fell to 38% in 1900, to 12.2% in 1950, and then to 3.6% in 1980, while the similar story also happened in Britain.

As we could see from the table 1, when the economy took off during the periods, accordingly, the share of rural population in that country declined significantly. In South Korea, as a typical example of Newly Industrial Economics (NIEs), the growth rate maintained exceptionally high between the early 1960s and 1990s, meanwhile, the share of rural population fell from 72.3% in 1960 to 20.4% in 2000. During the last twenty years of the twentieth century, some Asian countries have achieved great progress in economic development, such as China, Indonesia, Malaysia and Philippines, the shares of rural population also decreased remarkably. Since China started the “Reform and Open” policy in the end of 1970s, the rate of rural population has been declined from 80.4% in 1980 to 56.9% in 2008, as a country with over one billion people, there are millions of people crowd into the cities and adapt to the new life there. Just shown in the table 1, Argentina and Brazil have been undergoing the similar transitions between 1960 and 2008.

Without double, physical capital accumulation is the key factor to determine the growth rate of the economy especially during the period of industrial revolution. As for human capital, as Galor and Moav (2004) point out human capital accumulation has replaced the physical accumulation as a prime engine of economic growth in the transition from the Industrial Revolution to modern growth. Since the labor’s ability is the vital point which influences the speed of transition, the labor forces from rural areas have the incentive to invest their resources to improve their personal ability which will increase the intensity of competition in the cities. Accordingly, the original skilled workers, in the urban areas, need to increase their investment in
Table 1

<table>
<thead>
<tr>
<th>Country</th>
<th>Share of rural population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>26.4</td>
</tr>
<tr>
<td>Brazil</td>
<td>55.1</td>
</tr>
<tr>
<td>China</td>
<td>84</td>
</tr>
<tr>
<td>Indonesia</td>
<td>85.4</td>
</tr>
<tr>
<td>Malaysia</td>
<td>73.4</td>
</tr>
<tr>
<td>Mexico</td>
<td>49.2</td>
</tr>
<tr>
<td>Philippines</td>
<td>69.7</td>
</tr>
<tr>
<td>South Korea</td>
<td>72.3</td>
</tr>
<tr>
<td>Thailand</td>
<td>80.3</td>
</tr>
</tbody>
</table>

Source: World Bank

order to enhance their skill levels, which would create benefit for other agents from the rural areas. Acemoglu (1996) states that the social increasing returns in human capital accumulation are empirically important.

Since the effect of human capital on aggregate income is of central importance to both policymakers and economists, and while taxation policies are important instruments for government to improve the economic efficiency and social equity, especially when individuals’ ability is heterogeneous and private information, it is meaningful to consider how to constitute proper public policies to accelerate human capital and substance capital accumulation with regarding the distortion.

1.1. Literature Review

The optimal taxation stems from the work of Ramsey (1927), who considered the problem of choosing an optimal tax structure in an economy with a representative agent when only distorting taxes are available. This approach studies the problem of choosing taxes within a given set of available tax instruments and makes the conclusion that in raising infinitesimal revenue by
proportionate taxes on given commodities the taxes should be such as to diminish in the same proportion the production of each commodity taxed. I have to mention that another assumption crucial to this approach is that all activities of agents are observable and do not consider the situation with private information.

Following the Ramsey’s work, Atkinson and Stiglitz (1972) points out that if the utility function satisfies certain separable and homothetic assumptions, commodity taxation is uniform, which means the optimal taxes are equated across consumption goods. Atkinson and Stiglitz (1980) states that tax rates depend on income elasticities, with necessities taxed more than luxuries. Moreover, the familiar intuition from partial equilibrium that goods with low price elasticities should be taxed heavily does not necessarily apply in a general equilibrium setting. As for the intermediate goods, the taxes on them should be zero, and there are taxes only on final goods (Chair and Keheo(1999)). In these papers, the only inputs for goods production is labor and do not need capital and technology.

As considering the taxing capital income, in a model of infinitely lived households, The Chamley-Judd result states that the capital income could tax at initially high rates, but then the taxing rates should drop to zero in the long run (Judd (1985); Chamley (1986)). Here, in view of the capital’s significance for goods production, they made the above conclusion and did not consider the capital was so important for human capital production. Jones, Manuelli and Rossi (1997) extend the applicability of the Chamley-Judd result by showing that the return to human capital should not be taxed in the long run. However, in this model, it did not consider the situation with heterogeneous agents, and made the conclusion about the optimal capital income tax when their ability was uncertain in the future.

The Mirrlees approach to optimal taxation is built on a different foundation from Ramsey taxation. Rather than starting an ad-hoc restricted set of tax instruments as in Ramsey taxation, Mirrlees (1971) assumed that an informational friction endogenously restricted the set of taxes that implement the optimal allocation. This set-up allows arbitrary nonlinear taxes with heterogeneous agents and private information. The central problem of the Mirrlees approach is to find the optimal incentive-insurance tradeoff: how to provide the best insurance against adverse events while providing incentives for the agents to reveal their types. The vital result is that the consumption-leisure margin of an agent with the highest skill is undistorted (Mirrlees (1971), Diamond (1978)).
In recent years, lots of literature starting with Golosov, Kocherlakota and Tsyvinski (2003) and Werning (2001) extends the static Mirrlees (1971) framework to dynamic settings. Implementation of dynamic Mirrlees models is more complicated than implementation of either static Mirrlees models or Ramsey models (Golosov, Tsyvinski and Werning (2006)). In the paper of Golosov, Kocherlakota and Tsyvinski (2003), they consider the problem of optimal taxation when individual skills are unobservable and evolve stochastically over time. Given additive separability of preferences between consumption and labor, they have the different conclusion that although the uniform commodity taxation theorem is generally valid, but the zero capital income taxation theorem is generally not. Moreover, Kocherlakota (2004) did further study and pointed out that the marginal tax rate on saving may be negative or positive which was depending on the realization of individual labor income. Kocherlakota (2005) allows for a general process for skill shocks and derives an implementation with linear taxes on wealth and arbitrarily nonlinear taxes on the history of effective labor. Depending on the above study, Golosov and Tsyvinski (2006) try to find a tax system that implements the optimal allocation. They propose a tax system implementing the optimum: an asset-tested disability program. Insurance for all skill shocks with the exception of disability is accomplished through a direct mechanism that stands in for the income and wealth tax system. Moreover, this paper also proved that when agents are heterogeneous and market is incomplete, the linear taxation in Ramsey model is not the Pareto tax. In the above series of papers, they assumed that the agents’ ability was heterogeneous and unobservable, while they did not consider the knowledge spillovers between different kinds of people. Obviously, the people with high ability have positive external effect on the people with low ability. Meanwhile, they did not consider the human capital accumulation, or just considered the time was the only input for human capital accumulation, without taking into account the function of physical capital in human capital accumulation.

In this article, I consider the optimal taxation with the dynamic Mirrlees models in the neoclassical economics. The basic assumptions are as follows: there exist heterogeneous agents with private information about their personal abilities. But differ from the above dynamic model with regarding the human capital accumulation. In the paper Golosov, Kocherlakota and Tsyvinski (2003), they assume the agents’ skill levels which could be regarded as human capital are heterogeneous and independent to each other. While human capital introduces increasing
returns in the function that relates an agent’s supply of effective labor to her time input, agents could use their capital and time to improve their human capital. Besides, the agents’ skill levels are interaction other than independent. As we know, production can be given a spatial dimension by postulation a production externality that makes any individual more productive if other productive people are nearby (Glomm(1992)). Eaton and Eckstein (1997) point out that an externality affects the technology for accumulating human capital rather than the technology for producing goods. Thus, I use the two-sector economy in Rebelo (1991) and Jones, Maneulli and Rossi (1993, 1997) with regarding knowledge spillovers.

The rest of the paper is structured as follows: In the section 2, I state the assumptions in the paper. Then, I use the three-period model to work out the conditions for first best allocation when the information is complete. In the section 4, when the information is incomplete, I get the conditions for second best allocation. In the section 5, I have the conclusions about optimal taxation in the discrete economy.

2. Setup

The model economy is similar to those used in representative agent Ramsey models, such as Golosov, Tsyvinski and Werning (2006) and Bohacek and Kapicka(2008). The main difference is that there exists knowledge spillover effect between different kinds of agents, and in order to improve the personal ability, they also need to input physical capital besides the time. The reason of these assumptions is to emphasis the externality of human capital accumulation and the importance of physical capital, whatever in the process of producing the final goods or increase their skill levels, in the spirit of Mirrlees’s private information framework.

In the economy, there exists two kinds of people and the only difference between them is their personal ability denoted by \( \theta_i, \theta_H > \theta_h \), \( \theta_i \in R_{++} \), here the \( \theta_i \) is exogenous, unobservable and unchangeable (i = H, h). Furthermore, we could regard the difference of ability is decided by nature such as the intelligence.

Just as stated in Rebelo (1991) and Jones, Maneulli and Rossi (1993, 1997), throughout the paper I consider two-sector economy: one sector produces the final production and the other sector only produces the human capital. Although they have different production function, both
of them need input physical capital and labor. Moreover, the consumption market is perfect competitive.

Everyone has one unit endowment of time, which uses for leisure $L_t$, producing physical goods $l_t$ and human capital accumulation $N_t$, obviously, there exists $L_t = 1 - N_t - l_t$ and $l_t, N_t \in (0, 1)$.

### 2.1 Preferences and Technology

Each agent lives for three periods, which just like life cycle, in the first period, both types of agents have the similar human capital level; in the second period, as the learning ability is different across agents, they will have the different human capital level, and the externality will happen during this period; in the last period, they will allocate any resource in physical and human capital accumulation since they would like to use all their income for consumption. Meanwhile, they have identical preferences, presented by the discounted sum of utility

$$\sum_{t=0}^{3} \beta^t U(c_t, L_t)$$

where $\beta < 1, U : R_+ \times [0,1] \rightarrow R$, $c_t \in R_+$ is the agent’s consumption in period $t$ and $L_t \in [0,1]$ is the agent’s leisure in period $t$. I assume that in $R_+ \times (0,1]$, the utility function $U$ is continuous differential, strictly increasing and strictly concave, which mean $U_c, U_L > 0, U_{cc}, U_{LL} < 0$, and $\lim_{c \to 0} U_c = \infty, \lim_{L \to 0} U_L = \infty$. According to Atkinson and Stiglitz (1972, 1980), it is normal to assume the utility function is weakly separable between consumption and leisure, which means $U_{cl} = 0$, and we could rewrite the utility function as $U(c_t, L_t) = u(c_t) + \nu(1 - N_t - l_t)$

Before we describe the productive function, we need to know the concept of efficient labor. Just as in Golosov, Kocherlakota and Tsyvinski (2003), Kapicka (2005) and Kocherlakota (2005, 2007), the efficient labor is $y = HN$, where $H$ denotes the human capital and $N$ is the agent’s labor input. Efficient labor $y$ is observable, but human capital $H$ and actual labor $N$ are not.

In period $t$, the human capital’s productive function, $G : R_+ \times R_+ \times [0,1] \rightarrow R_+$, depends on the agent’s personal ability, time spending and physical capital. The productive function of high ability agents is

$$H_{H,t+1} = G(\theta_H, x_{Ht}, y_{Ht}) + (1 - \delta_H)H_{Ht}$$
where effective labor \( y^H_{ht} = H_{ht}N_{ht} \), \( N_{ht} \in (0,1) \) is the agent’s labor input in producing human capital; \( x_{ht} \) is the agent’s physical capital input; the productive function satisfies the assumptions about the new-classical productive function which mean it is continuous differential, strictly increasing and strictly concave in \( R^{++} \times (0,1) \) and \( \lim_{y \to 0} G_y = \infty \); \( \lim_{x \to 0} G_x = \infty \); \( \delta_{ht} \in (0,1) \) is the human capital’s discount rate.

With considering the knowledge spillovers, just as in Eaton and Eckstein (1997), Lucas (2004), in the proceeding of human capital accumulation, as a result of the difference among agents, the levels of human capital would be diversified. Moreover, the agents with high ability will have positive external effect on people with low ability, the productive function of low ability agents is

\[
H_{ht,t+1} = \left( \frac{H_{ht}}{H_{ht}} \right)^\delta G(\theta_{ht}, x_{ht}, y^H_{ht}) + (1 - \delta_{ht})H_{ht}
\]

Here, \( \delta > 0 \) is exogenous.

In period \( t \), the agent’s productive function of final goods is

\[ Y_i(t) = F(K_i, y^i_{it}) \]

where effective labor \( y^i_{it} = H_{it}I_{it} \), the \( I_{it} \in (0,1) \) is the agent’s labor input in producing final goods; \( K_i \) is the agent’s physical capital and the accumulation function is \( K_{it,t+1} = I_{it} + (1 - \delta_k)K_{it} \), here \( I_{it} \) is the agent’s capital investment in period \( t \) and \( \delta_k \) is the capital’s discount rate; \( F : R^+ \times [0,1] \to R^+ \) satisfies the assumptions about the new-classical productive which mean it is continuous differential, strictly increasing and strictly concave in \( R^{++} \times (0,1) \) and \( \lim_{y \to 0} F_y = \infty \); \( \lim_{k \to 0} F_k = \infty \).

Especially, just as in Barro, Sala-i-Martin (1995), there exists different density in using physical capital and labor between the productive functions of final goods and human capital. Obviously, the productive function of final goods is relative density in using physical capital while the function of human capital is relative density in using labor.

**2.2 Texas Policy**

As in the Mirrlees private information formwork, the taxation policy is the function of observed
variables. Individuals face the sequence of budget constraints with the initial wealth
\[ c_t + x_t + I_t + T_t \leq F(K_t, y_t^i) \]
Here \( T_t = T(y_t^H, y_t^l, x_t, I_t) \) is the taxes policy that depends on the individual’s efficient labor input and physical capital level.

In every period, the government uses the tax income to keep balance, which is implied by the individuals’ budget constraint and the resource constraints.

3. First Best Allocation

The economy lasts for 3 periods. Initially, the agents with different ability have the same human capital and physical capital, according to their difference in ability, their human capital will diversify in the next period.

Firstly, we discuss the optimal allocation when information is perfect and government knows their personal ability. The purpose of this section is to provide a benchmark against which the results in other sections may be compared.

The social planner characterizes optimal allocation \((c_u, N_u, l_u, x_u, I_u)_{i=H,h}\) by solving the program
\[
\max_{(c_u, N_u, l_u, x_u, I_u)} \sum_{i=H,h} \left( \sum_{t=1}^3 \beta^{t-1} U(c_{it}, 1 - N_{it} - l_{it}) + \beta^2 U(c_{i3}, 1 - l_{i3}) \right)
\]
Here, \( \beta \) is the time preference factor which satisfies \( 0 < \beta < 1 \).

In the first best allocation, the constraints for social planner are just the feasible conditions in every period. In period one, the feasible condition for government is
\[
\sum_{i=H,h} (c_{i1} + x_{i1} + I_{i1}) + g_1 = \sum_{i=H,h} F(K_{i1}, H_{i1} l_{i1})
\]
In period two,
\[
\sum_{i=H,h} (c_{i2} + x_{i2} + I_{i2}) + g_2 = \sum_{i=H,h} F(K_{i2}, H_{i2} l_{i2})
\]
Here, \( H_{i2} = G(\theta_i, x_{i1}, H_{i1} N_{i1}) + (1 - \delta_i) H_{i1}, K_{i2} = (1 - \delta_h) K_{i1} + I_{i1}, i = H, h \)

In period three,
\[
\sum_{i = i}^{n} c_{i3} + g_{3} = \sum_{i = i}^{n} F(K_{i3}, H_{i3}, l_{i3})
\]

(3-4)

Here, the physical capital \( K_{i3} = (1 - \delta_{t})K_{t2} + I_{t2} \); the human capital of agent with high ability is \( H_{h3} = G(\theta_{h}, x_{h2}, H_{h2}N_{h2}) + (1 - \delta_{h})H_{h2} \).

With considering the knowledge spillovers, the human capital of agent with low ability is \( H_{h3} = \left( \frac{H_{h2}}{H_{s2}} \right)^{y} G(\theta_{h}, x_{h2}, H_{h2}N_{h2}) + (1 - \delta_{h})H_{h2} \).

It is easy to solve the problem (3-1) under the constraints of (3-2) (3-3) (3-4), and get the following first order conditions:

As for the consumptions in each period,

\[
\frac{\partial U(c_{it}, L_{it})}{\partial L_{it}} = \frac{\partial U(c_{it}, L_{it})}{\partial c_{it}} \frac{\partial F(K_{it}, y_{it})}{\partial y_{it}} H_{it}
\]

(3-5)

Here, \( t = 1, 2, 3 \) while \( U_{it} = 0 \), the equation (3-5) means different agents should have the consumptions when the information is perfect.

As for the agent’s labor \( l_{i} \) input in producing final goods, there exists,

\[
\frac{\partial U(c_{it}, L_{it})}{\partial L_{it}} = \frac{\partial U(c_{it}, L_{it})}{\partial c_{it}} \frac{\partial F(K_{it}, y_{it})}{\partial y_{it}} H_{it}
\]

(3-6)

Here, \( i = H, h; t = 1, 2, 3 \) this equation reflects the marginal substitution between the labor input and consumption in the first best condition.

In period one, the labor \( N_{h1} \) and capital \( x_{h1} \) inputs in producing human capital for high ability should satisfy,

\[
U_{L_{H1}} = \beta U_{c_{H2}} F(y_{H2} \frac{\partial H_{h2}}{\partial N_{h1}} + \beta U_{c_{H3}} F(y_{H3} \frac{\partial H_{h3}}{\partial N_{h1}} + \beta U_{c_{h3}} F(y_{h3} \frac{\partial H_{h3}}{\partial N_{h1}}
\]

(3-7)

\[
U_{c_{H1}} = \beta U_{c_{H2}} F(y_{H2} \frac{\partial H_{h2}}{\partial x_{h1}} + \beta U_{c_{H3}} F(y_{H3} \frac{\partial H_{h3}}{\partial x_{h1}} + \beta U_{c_{h3}} F(y_{h3} \frac{\partial H_{h3}}{\partial x_{h1}}
\]

(3-8)

There are \( N_{h1} \) and \( x_{h1} \) ‘s Euler equations. The last parts of (3-7) (3-8) show the external effect when high ability agent uses his resources in producing human capital.
Theorem 1: When consider the knowledge spillovers, the agent with high ability should invest more resources in producing human capital. 

If we do not consider the external effect, the last parts of (3-7) (3-8) would not exist in the optimal condition. While \( \beta^2 U_{c_{h3}} F_{y_{h3}} l_{h3} \frac{\partial H_{h3}}{\partial H_{h2}} \frac{\partial H_{h2}}{\partial N_{h1}} > 0 \) and 

\( \beta^2 U_{c_{h3}} F_{y_{h3}} l_{h3} \frac{\partial H_{h3}}{\partial H_{h2}} \frac{\partial H_{h2}}{\partial x_{h1}} > 0 \), then, \( U_{c_{h1}} \) and \( U_{c_{u1}} \) would be higher when consider the knowledge spillovers. As \( U_{L_1} < 0 \) and \( U_{c_L} < 0 \), the agent has to spend less time in leisure and consume less goods. In other words, the agent should use more labor and capital in producing human capital.

In period one, the labor \( N_{h1} \) and capital \( x_{h1} \) inputs in producing human capital for low ability should satisfy,

\[
U_{L_{h1}} = \beta U_{c_{h2}} F_{y_{h2}} l_{h2} \frac{\partial H_{h2}}{\partial N_{h1}} + \beta^2 U_{c_{h3}} F_{y_{h3}} l_{h3} \frac{\partial H_{h3}}{\partial H_{h2}} \frac{\partial H_{h2}}{\partial N_{h1}} \tag{3-9}
\]

\[
U_{c_{h1}} = \beta U_{c_{h2}} F_{y_{h2}} l_{h2} \frac{\partial H_{h2}}{\partial x_{h1}} + \beta^2 U_{c_{h3}} F_{y_{h3}} l_{h3} \frac{\partial H_{h3}}{\partial H_{h2}} \frac{\partial H_{h2}}{\partial x_{h1}} \tag{3-10}
\]

There are \( N_{h1} \) and \( x_{h1} \)'s Euler equations which are the same as the conditions without considering the external effect.

In period two, the labor \( N_{i2} \) and capital \( x_{i2} \) inputs in producing human capital should satisfy,

\[
U_{L_{i2}} = \beta U_{c_{i3}} F_{y_{i3}} l_{i3} \frac{\partial H_{i3}}{\partial N_{i2}} \tag{3-11}
\]

\[
U_{c_{i2}} = \beta U_{c_{i3}} F_{y_{i3}} l_{i3} \frac{\partial H_{i3}}{\partial x_{i2}} \tag{3-12}
\]

Here \( i = H, h \). As in the three-period model, the agents will not invest their resources in producing human capital in the last period, then the inputs in producing human capital in period two would just have effects on the final goods’ output. Therefore, these Euler equations are the same as the conditions without considering the external effects.

In period one, the physical capital \( I_{i1} \) input in producing final goods, we get

\[
U_{c_{i1}} = \beta U_{c_{i2}} F_{k_{i2}} + \beta^2 U_{c_{i3}} F_{k_{i3}} (1 - \delta_k) \tag{3-13}
\]

\[
U_{c_{i2}} = \beta U_{c_{i3}} F_{k_{i3}} \tag{3-14}
\]
Here $i = H, h$.

From equation (3-5), we could find out that in every period, each type has the same consumption. However, from equation (3-6), the agent with high ability has to invest more resources in human capital accumulation. Under the condition with private information, the agent with high ability would like to hide his ability and pretend to be low ability; therefore, the allocation will not be Pareto optimal.

### 4. Second Best Allocation

In this part, I consider the economy with private information, and the social planner could just observe the set $(c, y^H, y^L, x, I) = (c^i_H, y^H_{i_H}, y^L_{i_H}, x^i, I^i_H)$, here $i = H, h; t = 1, 2, 3$. Besides the feasible conditions in every period just as the equation (3-2), (3-3), (3-4), the social planner uses the incentive-compatibility conditions to distinguish agents with different types, and the government recognizes a given agent’s ability at the beginning of period.

In third period, the incentive-compatibility condition for high ability agent,

$$U(c_{H3}, L_{H3}) = u(c_{H3}) + v(1 - \frac{y^H_{H3}}{H_{H3}}) \geq U(c_{h3}) + v(1 - \frac{y^H_{h3}}{H_{H3}}) \quad (4-1)$$

Here $H_{H3} = G(\theta_H, x_{H2}, y^H_{H2}) + (1 - \delta_H)H_{H2}$

In second period, the incentive-compatibility condition for high ability agent,

$$U(c_{H2}, 1 - \frac{y^H_{H2}}{H_{H2}} - \frac{y^L_{H2}}{H_{H2}} + \beta(U(c_{H3}, 1 - \frac{y^H_{H3}}{H_{H3}})) \geq U(c_{h2}, 1 - \frac{y^H_{h2}}{H_{h2}} - \frac{y^L_{h2}}{H_{h2}} + \beta(U(c_{h3}, 1 - \frac{y^H_{h3}}{H_{h3}})) \quad (4-2)$$

Here $H_{H2} = G(\theta_H, x_{H1}, y^H_{H1}) + (1 - \delta_H)H_{H1}, H_{H3} = G(\theta_H, x_{h2}, y^H_{h2}) + (1 - \delta_H)H_{H2}$

In period one, the incentive-compatibility condition for high ability agent,

$$\sum_{t=1}^{2} \beta^{t-1} U(c_{Ht}, 1 - \frac{y^H_{Ht}}{H_{Ht}} - \frac{y^L_{Ht}}{H_{Ht}}) + \beta^2 U(c_{Ht}, 1 - \frac{y^H_{H3}}{H_{H3}}) \geq U(c_{h1}, 1 - \frac{y^H_{h1}}{H_{h1}} - \frac{y^L_{h1}}{H_{h1}}) +$$

$$\beta U(c_{h2}, 1 - \frac{y^H_{h2}}{H_{h2}} - \frac{y^L_{h2}}{H_{h2}}) + \beta^2 U(c_{h3}, 1 - \frac{y^H_{h3}}{H_{h3}}) \quad (4-3)$$

Here $H_{H2} = G(\theta_H, x_{h1}, y^H_{h1}) + (1 - \delta_H)H_{H1}, H_{H3} = G(\theta_H, x_{h2}, y^H_{h2}) + (1 - \delta_H)H_{H2}$

At the beginning of period, the social planner considers the optimal problem under the constraints of feasible conditions, equation (3-2), (3-3), (3-4) and IC, equation (4-1), (4-2), (4-3).
In this case, at the optimal condition, the agent is indifferent between truthful announcement and adopting the strategy corresponding to the above binding IC constraints. However, once the social planner recognizes the agent’s real type and the series of allocation would be decided, then, among the three IC constraints, only the (4-3) IC constraint will be valid.

The first order conditions with respect to \((c, y^H, y^l, x, I) = (c_H, y^H_H, y^l_l, x_l, I_l)\) are as follow, and see the detail calculation process in Appendix:

The consumption of high ability agent \(c_H\),
\[
\beta^{-1}(1 + \mu) \frac{\partial U(c_H, L_H)}{\partial c_H} = \lambda
\]  
(4-4)

The consumption of low ability agent \(c_l\),
\[
\beta^{-1}(1 - \mu) \frac{\partial U(c_l, L_l)}{\partial c_l} = \lambda
\]  
(4-5)

Here \(\lambda, t = 1, 2, 3\) denote the Lagrange multipliers of the feasible conditions (3-2) (3-3) (3-4), and \(H\) denotes the Lagrange multiplier of the IC constraint (4-3), which mean the shadow price of the variables.

**Theorem 2:** When the agents’ ability is heterogeneous and private information, the optimal allocation for consumption should be \(c_H > c_l\), which means we need to pay the information rent to high ability agent.

From equation (4-4) (4-5), we have \((1 + \mu)U_{c_H} = (1 - \mu)U_{c_l}\), while the IC condition (4-3) is binding, the Lagrange multiplier \(\mu > 0\), then \(U_{c_H} < U_{c_l}\), as \(U_{c_l} < 0\), then we have \(c_H > c_l\).

The high ability agent’s effective labor \(y^H_H\) which is used for producing final goods,
\[
\frac{\partial U(c_H, L_H)}{\partial L_H} H_H = \frac{\partial U(c_H, L_H)}{\partial c_H} \frac{\partial F(K_H, y^H_H)}{\partial y^H_H}
\]  
(4-6)

Here \(t = 1, 2, 3\), it is the same as the conditions in the first best allocation.

The low ability agent’s effective labor \(y^l_l\) which is used for producing final goods,
\[
\frac{\partial U(c_l, L_l)}{\partial L_l} H_l = \frac{\partial U(c_l, L_l)}{\partial c_l} \frac{\partial F(K_l, y^l_l)}{\partial y^l_l}
\]  
(4-7)
Here $t = 2, 3$, except the initial period, the optimal condition of $y^t_{th}$ is different from the first best allocation.

The high ability agent’s effective labor $y^H_{th}$ which is used for producing human capital, in period one,

$$-rac{U_{t_{h1}}}{H_{h1}} + \beta U_{t_{h2}} \frac{y^H_{h2} + y^f_{h2}}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} + \beta^2 U_{t_{h3}} \frac{y^f_{h3} + \mu y^H_{h3} + \partial H_{h2}}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}}$$

$$+ \beta^2 U_{t_{h3}} \frac{y^f_{h3}}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y^H_{h1}} + \frac{\mu}{1 + \mu} \mu \frac{y^f_{h3} + \partial H_{h3}}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} = (4-8)$$

In period two,

$$\frac{U_{t_{h2}}}{H_{h2}} = \beta U_{t_{h3}} \frac{y^f_{h3}}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y^H_{h2}} (4-9)$$

Obviously, the equation (4-8) is different from the first best allocation while (4-9) is the same.

The low ability agent’s effective labor $y^H_{th}$ which is used for producing human capital, in period one,

$$-rac{U_{t_{h1}}}{H_{h1}} + \beta U_{t_{h2}} \frac{y^H_{h2} + y^f_{h2}}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} + \beta^2 U_{t_{h3}} \frac{y^f_{h3} + \partial H_{h2}}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} =$$

$$\frac{U^*_{t_{h1}}}{H_{h1}} + \beta U^*_{t_{h2}} \frac{y^H_{h2} + y^f_{h2}}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} + \beta^2 U^*_{t_{h3}} \frac{y^f_{h3} + \partial H_{h2}}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial y^H_{h1}} = (4-10)$$

In period two,

$$-\frac{U_{t_{h2}}}{H_{h2}} + \beta U_{t_{h3}} \frac{y^f_{h3}}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y^H_{h2}} = \frac{U^*_{t_{h2}}}{H_{h2}} + \beta U^*_{t_{h3}} \frac{y^f_{h3}}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y^H_{h2}} = (4-11)$$

Here $U^*$ denotes the utility when high ability agent pretends to be low ability agent.

The high ability agent’s physical capital $x_{th}$ which is used for human capital accumulation, in period one,

$$U_{t_{h1}} = \beta U_{t_{h2}} \frac{y^H_{h2} + y^f_{h2}}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial x_{h1}} + \beta^2 U_{t_{h3}} \frac{y^f_{h3} + \partial H_{h2}}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h1}}$$
\[
\frac{1}{1+\mu} (\beta^2 U_{l_h}) \left( \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h2}} \right)
\]  
(4-12)

In period two,
\[
U_{e_{h2}} = \beta U_{l_h} \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h2}}
\]  
(4-13)

Compared with equation (3-12) in the first best allocation, the right hand side of equation (4-12) has less \( \frac{\mu}{1+\mu} (\beta^2 U_{l_h}) \left( \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h2}} \right) \), which means the value of \( U_{e_{h1}} \) is less than that in the first best allocation and the high ability agent would consume more in period one when the agents’ ability is private information.

The low ability agent’s physical capital \( x_{h_t} \) which is used for human capital accumulation, in period one,
\[
U_{e_{h1}} > \beta U_{l_h} \frac{y_{h2}^H + y_{h2}^l \frac{\partial H_{h2}}{\partial x_{h1}} + \beta^2 U_{l_h} \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h2}}}{(H_{h2})^2}
\]  
(4-14)

In period two,
\[
U_{e_{h2}} > \beta U_{l_h} \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h2}}{\partial x_{h2}}
\]  
(4-15)

The first order condition of \( I_{it} \) is consistent with the first best allocation, here \( i = H, h \) and \( t = 1, 2 \).

5. Optimal Taxation

When agents are heterogeneous, the linear taxation is not the Pareto optimal taxation (Golosov(2006)). Here, I discuss the non-linear taxation which depends on the observed information \( (c, y^H, y^l, x, I) = (c_{it}, y_{it}^H, y_{it}^l, x_{it}, I_{it}) \)

In the discrete economy, the agent \( i \) faces the following problem:
\[
\max_{(c_{it}, y_{it}^H, y_{it}^l, x_{it}, I_{it})} \sum_{t=1}^{2} \beta^{t-1} U(c_{it}, 1 - \frac{y_{it}^H}{H_{it}} - \frac{y_{it}^l}{H_{it}}) + \beta^2 U(c_{i3}, 1 - \frac{y_{i3}^l}{H_{i3}})
\]

The feasible conditions in every period are the constraints, in the perfect competitive market, as the firms are owned by agents; we have the feasible condition in period one,
\[ c_{i1} + x_{i1} + I_{i1} + T(y_{i1}^{H}, y_{i1}^{f}, x_{i1}, I_{i1}) = F(K_{i1}, y_{i1}^{f}) \]  
(5-1)

In period two,
\[ c_{i2} + x_{i2} + I_{i2} + T(y_{i2}^{H}, y_{i2}^{f}, x_{i2}, I_{i2}) = F(K_{i2}, y_{i2}^{f}) \]  
(5-2)

In period three,
\[ c_{i3} + T(y_{i3}^{f}) = F(K_{i3}, y_{i3}^{f}) \]  
(5-3)

It is easy to have the first order conditions, the effective labor \( y_{it}^{f} \) used for producing final goods,
\[
\frac{U_{t_{it}}}{H^{it}} = U_{c_{it}} \left( F_{y_{it}^{f}} - \frac{\partial T(y_{it}^{H}, y_{it}^{f}, x_{it}, I_{it})}{\partial y_{it}^{f}} \right) 
\]  
(5-4)

The effective labor \( y_{i1}^{H} \) used for producing human capital in period one,
\[
- \frac{U_{t_{i1}}}{H^{i1}} + \beta U_{t_{i2}} \frac{y_{i1}^{H} + y_{i2}^{H}}{(H_{i2})^{2}} \frac{\partial H_{i2}}{\partial y_{i1}^{H}} + \beta^{2} U_{t_{i3}} \frac{y_{i2}^{f} + y_{i3}^{f}}{(H_{i3})^{2}} \frac{\partial H_{i3}}{\partial y_{i2}^{f}} = U_{c_{i1}} \left( F_{y_{i1}^{H}} + \frac{\partial T(y_{i1}^{H}, y_{i1}^{f}, x_{i1}, I_{i1})}{\partial y_{i1}^{H}} \right) 
\]  
(5-5)

In period two,
\[
- \frac{U_{t_{i2}}}{H^{i2}} + \beta U_{t_{i3}} \frac{y_{i2}^{f} + y_{i3}^{f}}{(H_{i3})^{2}} \frac{\partial H_{i3}}{\partial y_{i2}^{f}} = U_{c_{i2}} \left( F_{y_{i2}^{f}} + \frac{\partial T(y_{i2}^{H}, y_{i2}^{f}, x_{i2}, I_{i2})}{\partial y_{i2}^{f}} \right) 
\]  
(5-6)

The physical capital \( x_{i1} \) used for human capital accumulation in period one,
\[
U_{c_{i1}} \left( 1 + \frac{\partial T(y_{i1}^{H}, y_{i1}^{f}, x_{i1}, I_{i1})}{\partial x_{i1}} \right) = \beta U_{t_{i2}} \frac{y_{i2}^{H} + y_{i2}^{f}}{(H_{i2})^{2}} \frac{\partial H_{i2}}{\partial x_{i1}} + \beta^{2} U_{t_{i3}} \frac{y_{i3}^{f}}{(H_{i3})^{2}} \frac{\partial H_{i3}}{\partial x_{i1}} 
\]  
(5-7)

In period two,
\[
U_{c_{i2}} \left( 1 + \frac{\partial T(y_{i2}^{H}, y_{i2}^{f}, x_{i2}, I_{i2})}{\partial x_{i2}} \right) = \beta U_{t_{i3}} \frac{y_{i3}^{f}}{(H_{i3})^{2}} \frac{\partial H_{i3}}{\partial x_{i2}} 
\]  
(5-8)

The physical capital \( I_{i1} \) used for producing final goods, in period one,
\[
U_{c_{i1}} \left( 1 + \frac{\partial T(y_{i1}^{H}, y_{i1}^{f}, x_{i1}, I_{i1})}{\partial I_{i1}} \right) = \beta U_{c_{i2}} F_{k_{i2}} + \beta^{2} U_{c_{i3}} F_{k_{i3}} (1 - \delta_{k}) 
\]  
(5-9)
In period two,  
\[ U_{i_2}^* (1 + \frac{\partial T(y_{i_1}^H, y_{i_1}^l, x_{i_1}, I_{i_1})}{\partial I_{i_2}}) = \beta U_{i_3} F_{k_{i_3}} \]  
(5-10)

Here \( i = H, h; t = 1, 2, 3 \)

Compared the above conditions with the conditions in second best allocation, we have the following conclusions:

**Theorem 3:** When the information is incomplete, other than the initial period, the marginal labor income tax on high ability agent should be zero, while on low ability agent should be positive.

According to the equation of (4-6), we could find out in equation (5-4),  
\[ \frac{\partial T(y_{h_{i_1}}^H, y_{h_{i_1}}^l, x_{h_{i_1}}, I_{h_{i_1}})}{\partial y_{h_{i_1}}^l} = 0, \]  
while for low ability agent, we have the inequation of (4-7), it is easy to have  
\[ \frac{\partial T(y_{h_{i_1}}^H, y_{h_{i_1}}^l, x_{h_{i_1}}, I_{h_{i_1}})}{\partial y_{h_{i_1}}^l} > 0. \]  
Then, we have the above conclusion.

Since the positive marginal tax rate has a discouraging effect on the agent’s effort, and compared with the low ability agent, the individual with high ability has higher elasticity of labor supply, then if the revenue collected by tax is the same, the zero marginal tax on earning beyond some level could be optimal.

As Mankiw, Weinzierl and Yagan (2009) state the optimal marginal taxation depends on the distribution of ability, while in practice, as we could see from the figure 1, the top marginal income tax rate in US has declined from 70% in 1979 to 35% in 2010, which is partially consistent with the above results.
Theorem 4: When the information is incomplete and the high ability agents’ inputs of producing human capital have external effects, it is optimal for government to subsidize them.

At period one, the effective labor \( y_{H1}^H \) have the equation of (5-5), with considering the equation (4-7), we have,

\[
\frac{\partial T(y_{H1}^H, y_{H1}^I, x_{H1}, I_{H1})}{\partial y_{H1}^H} = -\left( \frac{1}{1+\mu} \beta^2 U_{t_h^3} \frac{y_{H1}^I}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{H1}^H} \frac{\partial H_{h2}}{\partial y_{H1}^H} \right)
\]  

(5-11)

Obviously, we have \( \frac{\partial T(y_{H1}^H, y_{H1}^I, x_{H1}, I_{H1})}{\partial y_{H1}^H} < 0 \); as for the physical capital \( x_{H1} \), we have equation (5-7), compared with equation (4-12), we get,

\[
\frac{\partial T(y_{H1}^H, y_{H1}^I, x_{H1}, I_{H1})}{\partial x_{H1}} = -\left( \frac{1}{1+\mu} \beta^2 U_{t_h^3} \frac{y_{H1}^I}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{H1}} \frac{\partial H_{h2}}{\partial x_{H1}} \right)
\]  

(5-12)

thus, \( \frac{\partial T(y_{H1}^H, y_{H1}^I, x_{H1}, I_{H1})}{\partial x_{H1}} < 0 \)

Rauch (1993) finds that a one-year increase in the average education in a metropolitan area is associated with a three percent increase in wages even in regression that control for observed characteristics. In modern economics, the emergence of high-skilled workers is often treated as resource of economic growth which is often argued as an important barrier to the development of countries. Since the high ability agent’s investment in human capital has the social scale effects,
the subsidy will improve the social welfare.

**Theorem 5:** When the information is incomplete, the government should subsidize the low ability agent’s physical capital which is used for human capital accumulation.

In the second best economy, the conditions for physical capital input are inequation \((4-14)\) \((4-15)\), compared with \((5-7)\) \((5-8)\), we have \(\frac{\partial T(y^{H}_{i}, y^{I}_{i}, x_{i}, I_{i})}{\partial x_{i}} < 0, t = 1, 2\)

Even though the low-skilled agents’ increasing of human capital accumulation does not have the external effect, considering the catching-up effect, it is still worth for government to subsidize the low-skilled agents, which could narrow the gap between the rich and poor so that the social welfare could be improved.

**Theorem 6:** When the information is incomplete, the marginal tax rate of physical capital which is used for producing final goods is zero.

With the equation \((5-9)\) \((5-10)\) and \((3-13)\) \((3-14)\) in the second best economy, we have \(\frac{\partial T(y^{H}_{i}, y^{I}_{i}, x_{i}, I_{i})}{\partial I_{i}} = 0, i = H, h\)

Just as Chamley (1986) and Judd (1985) show that the optimal capital income ought to be untaxed, at least in expectation, the intuition behind the zero capital taxes is that capital taxes could yield large distortions to intertemporal consumption plans and discourage saving, which would decrease the aggregate output in the future.

**6. Conclusion**

In this paper, we explored the optimal taxation issue in the neoclassical economy with considering the positive externality between different kinds of agents. Different marginal taxation emerges due to the agent heterogeneity. Under the assumption that both physical capital and time are necessary, not only to produce the final goods, but also to improve the skill levels, it gave the theoretical support for some public policies, such as subsidizing the education. There
are several assumptions for the analysis, and now I just speculate on their role in the mail result.

Firstly, in the dynamic environment, I just ignore any aggregate or idiosyncratic uncertainty, which allow us to reduce the optimal taxation issues to a simple static subproblem. In turn, we could provide the clear analysis about the direction of binding incentive constraints and the sign of marginal income taxation, which would be influenced by the uncertainty of the economics.

Secondly, the model just consider the issue about how to minimize the distortion when the government need to collect the given amount of revenue, but we ignore the important function caused by the government expenditure, such as the supplying the public goods and the redistribution effects. Barro (1990) supply the endogenous model with considering the contribution of government spending in utility and production, which we could use for analyzing the optimal taxations.

Reference

Appendix

In the economy with private information, the government as the social planner solves the optimal program:

$$\max_{(c_{it}, N_{it}, l_{it}, t_{it})} \sum_{t=1}^{2} \left( \sum_{i=1}^{2} (\beta^{t-1} U(c_{it}, 1 - N_{it} - l_{it}) + \beta^{t-1} U(c_{iti}, 1 - l_{iti})) \right),$$

s.t. equation (3-2) (3-3) (3-4) and (4-3).

Construct the Lagrange equation:

$$L = \sum_{t=1}^{2} \left( \sum_{i=1}^{2} (\beta^{t-1} U(c_{it}, 1 - \frac{y_{Ht}}{H_{it}} - \frac{y_{lt}}{H_{lt}}) + \beta^{t-1} U(c_{iti}, 1 - \frac{y_{Hiti}}{H_{iti}})) + \lambda_{1}(\sum_{i=H,h} F(K_{it}, y_{it}) - \sum_{i=H,h} (c_{i1} + x_{i1} + I_{i1}) - g_{1}) + \lambda_{2}(\sum_{i=H,h} F(K_{i2}, y_{i2}) - \sum_{i=H,h} (c_{i2} + x_{i2} + I_{i2}) - g_{2})

+ \lambda_{3}(\sum_{i=H,h} F(K_{i3}, y_{i3}) - \sum_{i=H,h} c_{iti} - g_{3}) + \mu(\sum_{i=1}^{2} (\beta^{t-1} U(c_{it}, 1 - \frac{y_{Ht}}{H_{it}} - \frac{y_{lt}}{H_{lt}}))

+ \beta^{2} U(c_{Ht}, 1 - \frac{y_{Ht}}{H_{Ht}}) - U^{*}(c_{Ht}, 1 - \frac{y_{Ht}}{H_{Ht}} + \frac{y_{Ht}}{H_{Ht}}) - \beta U^{*}(c_{Ht}, 1 - \frac{y_{Ht}}{H_{Ht}} + \frac{y_{Ht}}{H_{Ht}}) - \beta U^{*}(c_{Ht}, 1 - \frac{y_{Ht}}{H_{Ht}} + \frac{y_{Ht}}{H_{Ht}}))

Here $\lambda_{t}, t = 1, 2, 3$ denote the Lagrange multipliers of the feasible conditions (3-2) (3-3) (3-4), and

$\mu$ denotes the Lagrange multiplier of the IC constraint (4-3), $U^{*}$ denotes the utility function when high ability agent pretends to be low ability agent.

F.O.C:

$$c_{Ht} : \beta^{t-1}(1 + \mu) \frac{\partial U(c_{Ht}, L_{Ht})}{\partial c_{Ht}} = \lambda_{t}$$

$$c_{Ht} : \beta^{t-1}(1 - \mu) \frac{\partial U(c_{Ht}, L_{Ht})}{\partial c_{Ht}} = \lambda_{t}$$
\[ y_{ln}^l : \frac{\partial U(c_{ln}, L_{hn})}{\partial L_{hn}} 1 + \mu \frac{\partial F(K_{hn}, y_{ln}^l)}{\partial y_{ln}^l} = \lambda \frac{\partial F(K_{hn}, y_{ln}^l)}{\partial y_{ln}^l} \]

\[ y_{hn}^l : \frac{\partial U(c_{ln}, L_{hn})}{\partial L_{hn}} \frac{\beta^{-1}}{H_{hn}} = \lambda \frac{\partial F(K_{hn}, y_{ln}^l)}{\partial y_{ln}^l} + \mu \frac{\beta^{-1}}{H_{hn}} \frac{\partial U^*(c_{ln}, L_{hn})}{\partial L_{hn}} \]

\[ y_{h1}^H : (-\frac{U_{h1}^l}{H_{h1}} + \beta U_{h2}^l \frac{y_{h2}^l + \gamma_{h2}^l}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial y_{h2}^l} + \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{h3}^l} \frac{\partial H_{h2}}{\partial y_{h2}^l})(1 + \mu) \]

\[ = -\frac{1}{1 + \mu} \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{h3}^l} \frac{\partial H_{h2}}{\partial y_{h2}^l} \]

\[ y_{h2}^H : (-\frac{U_{h2}^l}{H_{h2}} + \beta U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{h3}^l})(1 + \mu) = 0 \]

\[ y_{h2}^H : -\frac{U_{h2}^l}{H_{h2}} + \beta U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{h3}^l} - \mu(-\frac{U_{h2}^l}{H_{h2}} + \beta U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial y_{h3}^l}) = 0 \]

\[ \lambda_1 = (\beta U_{h2}^l \frac{y_{h2}^l}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial x_{h2}} + \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h1}} \frac{\partial H_{h2}}{\partial x_{h1}})(1 + \mu) \]

\[ x_{h1} = \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h2}}(1 + \mu) \]

\[ x_{h2} : \lambda_2 = \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h2}}(1 + \mu) \]

\[ x_{h1}^l : \lambda_1 = \beta U_{h2}^l \frac{y_{h2}^l}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial x_{h1}} + \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h1}} \frac{\partial H_{h2}}{\partial x_{h1}} \]

\[ -\mu(\beta U_{h2}^l \frac{y_{h2}^l}{(H_{h2})^2} \frac{\partial H_{h2}}{\partial x_{h1}} + \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h1}} \frac{\partial H_{h2}}{\partial x_{h1}}) \]

\[ x_{h2} : \lambda_2 = \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h2}} - \mu \beta^2 U_{h3}^l \frac{y_{h3}^l}{(H_{h3})^2} \frac{\partial H_{h3}}{\partial x_{h2}} \]

\[ I_1 = \lambda_2 F_{h2}^l + \lambda_3 F_{h3}^l (1 - \delta_k) \]
\[ I_{12} = \lambda_2 F k_{13} \]

As for inequation (4-7), the optimal condition of \( y_{ht} \):

\[
\frac{\partial U(c_{ht}, L_{ht})}{\partial L_{ht}} \frac{1}{H_{ht}} = \lambda \frac{\partial F(K_{ht}, y_{ht}^l)}{\partial y_{ht}^l} + \mu \frac{\partial U^*(c_{ht}, L_{ht})}{\partial L_{ht}},
\]

Other than initial period, the human capital of high ability agent is higher than low ability agent, obviously, facing the same effective labor, high ability will work less and enjoy more leisure, then we get \[
\frac{\partial U(c_{ht}, L_{ht})}{\partial L_{ht}} \frac{(1 - \mu)}{H_{ht}} < \lambda \frac{\partial F(K_{ht}, y_{ht}^l)}{\partial y_{ht}^l}.
\]

According to the optimal condition of \( c_{ht} \), we have the (4-7).

With considering the optimal condition of \( x_{hi} \), while \( U_{h_2}^* < U_{h_2}^* \), \( U_{h_3}^* < U_{h_3}^* \), and as for human capital accumulation, with the same inputs, the low ability agent’s marginal production of physical capital and labor is higher than high ability agent, it is easy to get (4-11) (4-12).