

Discussion Paper No. 590

**ENDOGENOUS HUMAN CAPITAL  
ACCUMULATION, COMPARATIVE ADVANTAGE  
AND DIRECT vs. INDIRECT REDISTRIBUTION**

Hisahiro Naito

July 2003

The Institute of Social and Economic Research  
Osaka University  
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

# Endogenous Human Capital Accumulation, Comparative Advantage and Direct vs. Indirect Redistribution

Hisahiro Naito <sup>\*†</sup>  
Institute of Social and Economic Research  
Osaka University  
and  
Department of Economics  
University of California Irvine

Previous version August 30, 2002  
Current version June 7, 2003

Forthcoming from *Journal of Public Economics*

## Abstract

Recently, several papers have re-examined the so-called production efficiency theorem and the Atkinson and Stiglitz theorem on commodity taxes in the optimal taxation literature. Naito (1999) showed that indirect redistribution through production distortion or consumption distortion can Pareto-improve welfare and that the two theorems do not necessarily hold when different factors are imperfect substitutes and factor prices are endogenous. On the other hand, Saez (2001) argued that in the long run where human capital accumulation is endogenous, the two theorems are still valid. This paper develops reasonable alternative models where individuals accumulate human capital based on their comparative advantage. The present paper shows that the production efficiency theorem is not necessarily valid and that indirect redistribution from the able to the less able such as tariffs and production subsidies can increase efficiency even when skill accumulation is endogenous.

Keywords: Human capital accumulation, non-linear income taxation, and comparative advantage  
JEL Number: H21, H23

---

\*I appreciate the participants of the Public Economics Research Group around Osaka and Daiji Kawaguchi for his helpful information on the relationship between earnings and ability. I am very grateful for the comments from two anonymous referees and professor Robin Boadway, the editor of the Journal. Their comments have been very helpful for improving this paper. Of course, the author is responsible for all remaining errors.

†Address: Institute of Social and Economic Research, Osaka University, Mihogaoka 6-1, Ibaraki City, Osaka, Japan, postal code 567-0047  
phone:81-6-6879-8581; fax: 81-6-6878-2766.  
e-mail address: naito@iser.osaka-u.ac.jp

# 1 Introduction

Whether efficient income redistribution should be done through income taxation alone or should be complemented with other measures such as production distortion or consumption distortion is one of the key issues whenever optimal public policies are discussed. With this regard, the production efficiency theorem (Diamond and Mirrlees, 1971), which states that production distortion is not optimal and the Atkinson and Stiglitz theorem on optimal commodity taxation (Atkinson and Stiglitz, 1976, 1980), which shows that commodity taxation is not necessary in the presence of an optimal income tax system, are the most important results in public finance literature.

However, in public finance literature researchers started examining those results. For example, Cremer, Pestieau and Rochet (2001) showed that the Atkinson and Stiglitz theorem does not hold when individuals are different in ability and endowment. Saez (2003) showed that the Atkinson and Stiglitz theorem does not hold when tastes are heterogenous. Naito (1999) showed that in a model similar to the model of Stiglitz (1982), if multiple goods are produced and factor prices are endogenous, the Atkinson and Stiglitz theorem does not necessarily hold and the production efficiency result does not either.

On the other hand, many of the previous studies on optimal income taxation have received criticism that they did not focus on long term decisions such as human capital accumulation but focused on the short term choices such as labor supply. As a result, it is sometimes argued that the result obtained in the short run model might not hold in the long run.

In particular, Saez (2003) made skill accumulation endogenous in the model of optimal taxation and analyzed several issues of public policy. He showed that Naito's results are not valid and that the production efficiency theorem and the Atkinson and Stiglitz theorem on commodity taxation are valid when human capital accumulation is endogenous. Since accumulation of human capital has a strong effect on the economy in the long run and since the implications of the production efficiency theorem and the Atkinson and the Stiglitz theorem are important, the

contribution of Saez's paper is substantial.

Despite such contributions, however, we believe that a further investigation would be needed. In many previous analyses involving asymmetric information not only in public finance literature but also in other literature, conclusions were not often robust in the sense that they critically depended on the structure of information and the timing of information revelation. Thus, it is worthwhile investigating the robustness of the result of Saez (2003) with another reasonable set of assumptions. In particular, in this paper we will show that if higher ability persons have *comparative advantage* in the sense that the relative return from accumulating skilled human capital to unskilled human capital is higher than that of lower ability persons, the production efficiency theorem does not hold.<sup>1</sup>

To explain the intuition of the present paper, it would be useful to look at the differences between the assumptions in Saez (2003) and those in Naito (1999). In Naito (1999), there are two factors of production that are imperfect substitutes. In addition, from the beginning, each individual is attached to a particular labor market (the skilled labor market or the unskilled labor market) but the government cannot observe whether each individual is attached to the skilled or unskilled labor markets. The main idea in Naito (1999) is that when the government cannot observe an individual's type, the government can affect different individuals differently by using the response of the factor markets (Stolper and Samuelson theorem, Stolper and Samuelson (1941)). Since the income tax policy cannot discriminate the different types of agents attached to different labor markets but a commodity tax and a tariff can, using a commodity tax (in a case of a closed economy), or a tariff (in a case of an open economy) with the response of factor markets can increase the efficiency.

In Saez (2003), each job requires pre-determined skill levels. As a result, the income level

---

<sup>1</sup>In this paper, we only analyze the case of a small open economy due to the limitation of the space. As a result, we only prove that the production efficiency theorem does not hold. In a small open economy, a commodity tax cannot affect the producer prices and hence, factor prices. Therefore, the Atkinson and Stiglitz theorem holds in a small open economy. On the other hand, the result is changed in a closed economy. In the previous version of the paper, we proved that the Atkinson and Stiglitz theorem does not hold in a closed economy. See the section 5 in the present paper for more discussion. Also, for proof, please see our previous version of the present paper (Naito, 2002).

represents the amount of skill the individuals acquired. People have a heterogeneous ability to acquire skill. However, since such heterogeneity of ability is incorporated in the utility function as a difference of disutility to acquire skill, there is not any room for such heterogeneous ability to interact with the market reaction. Thus, the heterogeneity of abilities is intrinsically independent of the external environment of the economy. When the heterogeneity of abilities is independent of the response of the factor market, it is essentially equivalent to assuming that the dimension of factors used for production is one. In such a case, changes of factor prices due to government policy cannot increase the economic efficiency.

The key idea of the present paper is that in the presence of *comparative advantage* in accumulating different types of human capital, individuals with different abilities will be affected differently by the responses of factor markets even when skill accumulation is endogenous. In such a case, a policy that introduces inefficient production but affects the factor prices differently for different factors can indirectly redistribute from the less able to the able. Although such a policy cause a distortion, it has only the second order effect, but such an indirect redistribution has the first order effect on welfare. Thus, it can increase the social welfare.

For illustration, consider a situation where there are two types of human capital: skilled human capital and unskilled human capital and where those who have higher ability have *comparative advantage* in accumulating skilled human capital. *Comparative advantage* in accumulating skilled human capital for the able means that the relative benefit from accumulating skilled human capital to unskilled human capital for individuals with high ability is higher than for the less able. We could think that training, knowledge and experience in white collar jobs are skilled human capital and those in blue collar jobs are unskilled human capital. In such a situation, a decrease of the return from skilled human capital and an increase of the return from unskilled human capital will hurt the able relatively more and give relatively more benefit to the less able. The intuition of this paper is that when individual ability is not observable to the social planner but the social planner is interested in redistribution from high ability individuals to low ability

individuals, then a policy that will change the returns from skilled and unskilled human capital differently might be useful for an efficiency reason.

The crucial assumption in the present paper is the presence of comparative advantage in human capital accumulation. Whether such an assumption is reasonable or not is an interesting empirical question. Earlier literature of the human capital theory assumed that earning could be explained completely once it is conditioned by human capital level. Earlier empirical evidences showed that there is a strong correlation between earnings and the level of human capital and indicated that ability does not matter for explaining earnings once they are conditioned by the human capital levels. On the other hand, recent literature of labor economics and self-selection emphasizes that ability can also increase earning and play a systematic role for explaining earnings even after it is conditioned by the human capital level. This literature points out that even in an extreme case when human capital does not increase the productivity at all, if ability can increase the productivity and if higher ability agents tend to acquire more skills, there will be a correlation between human capital level and earnings. In the standard signaling literature, it is commonly assumed that a higher ability person would get more benefit from acquiring skill. In addition, recently, Dinardo and Tobias (2001) and Tobias (2003) examined whether the returns from schooling are higher for high ability individuals than for low ability individuals by using a non-parametric method. They found that the returns are higher for high ability individuals than for low ability individuals. This suggests that assuming the presence of comparative advantage is not unrealistic as an approximation of the reality.

At this point, one might wonder about the difference between Naito (1999) and the present paper. In the case of Naito (1999), each type of worker is attached to a different labor market. As a result, skilled workers can supply only skilled labor and unskilled workers can supply only unskilled labor. However, in the present paper, both high ability persons and low ability persons have options to accumulate both types of human capital or either type of human capital. Thus, the reasoning that an increase of the return from unskilled human capital and a decrease of the

return from skilled human capital always increases efficiency is not obvious.

The organization of this paper is as follows. In section 2, we present the model in a small open economy and analyze the production efficiency theorem by Diamond and Mirrlees (1971) when two factors are imperfect substitutes. In section 3, we analyze the same issue when two factors are perfect substitutes. In section 4, we shall give the implications and in section 5 we will give a brief conclusion.

## 2 The model

The economy is small and open and there are two output goods: good 1 and good 2. Good 1 is skilled human capital intensive good and good 2 is unskilled human capital intensive good. We assume that there are two types of human capital in this economy: skilled human capital and unskilled human capital. In this economy, there is a continuum of agents and all agents have identical, additive separable utility functions with respect to consumption, skilled human capital investment and unskilled human capital investment. We index all individuals' ability by  $i$  where  $i$  takes any value from one to two. We assume that the utility function of the type  $i$  agent has the following form:

$$u(c_{1i}, c_{2i}) - f_s(h_i^s) - f_u(h_i^u)$$

where  $u(c_{1i}, c_{2i})$  is strictly increasing with each argument and strictly concave and  $f_s(h_i^s)$  and  $f_u(h_i^u)$  are strictly increasing and strictly convex.  $c_{1i}$  and  $c_{2i}$  are the consumption of good 1 and good 2 by agent  $i$ . We assume that the labor supply is fixed and it is normalized to one.  $h_i^s$  and  $h_i^u$  are the levels of skilled and unskilled human capital of individual  $i$ .  $h_i^s$  and  $h_i^u$  can be interpreted as the knowledge levels, years of education, experience and training for each type of skill. In addition, to illustrate our point, we assume that  $f_s(h_i^s)$  and  $f_u(h_i^u)$  have the following functional forms:<sup>2</sup>

$$f_s(h_i^s) = (h_i^s)^{\gamma_s} \quad \text{and} \quad f_u(h_i^u) = (h_i^u)^{\gamma_u}$$

---

<sup>2</sup>Our main results can hold in more general functional forms.

where  $\gamma_s$  and  $\gamma_u$  measure the curvature of the disutility functions of skilled and unskilled human capital accumulation respectively and they are strictly greater than one. Given the amount of skilled human capital and unskilled human capital of individual  $i$ , we assume that the earning of individual  $i$  is determined as follows:

$$\text{earning}_i = g_s(i) \times w^s \times h_i^s + g_u(i) \times w^u \times h_i^u \quad (1)$$

where  $w^s$  and  $w^u$  are the returns from one efficient unit of skilled and unskilled human capital, respectively. (1) means that when individual  $i$  accumulates  $h_i^s$  units of skilled human capital and  $h_i^u$  units of unskilled human capital, the efficient unit of skilled human capital and unskilled human capital are  $g_s(i) \times h_i^s$  and  $g_u(i) \times h_i^u$  and the total return from skilled human capital and unskilled human capital are  $g_s(i) \times w^s \times h_i^s$  and  $g_u(i) \times w^u \times h_i^u$ , respectively. Let  $g_s(i) \times w^s$  and  $g_u(i) \times w^u$  be  $w_i^s$  and  $w_i^u$ .  $g'_s(i)/g_s(i)$  and  $g'_u(i)/g_u(i)$  measure the absolute advantage of an agent with ability  $i + \epsilon$  over agent  $i$  in accumulating skilled human capital and unskilled human capital, respectively. We assume that agents who have higher ability have absolute advantage in accumulating both skilled human capital and unskilled human capital:  $g'_s(i)/g_s(i) > 0$  and  $g'_u(i)/g_u(i)$ .<sup>3</sup> Also, as we discussed in the introduction, we assume that agents who have higher ability have comparative advantage in accumulating skilled human capital than unskilled human capital. Thus, we assume that

$$\frac{g'_s(i)}{g_s(i)} > \frac{g'_u(i)}{g_u(i)} \frac{\gamma_u}{\gamma_s} \quad (2)$$

The assumption (2) has a clear economic meaning. Consider a situation where the disutility functions of accumulating skilled human capital and unskilled human capital have the same degree of curvature ( $\gamma_s = \gamma_u$ ). In this case, (2) means that an agent whose ability is higher will have a larger rate of increase of  $w_i^s$ , the return from accumulating skilled human capital, than that of  $w_i^u$ , the return from accumulating unskilled human capital. When the curvature of the disutility functions are different, (2) says that the condition of the comparative advantage

---

<sup>3</sup>The assumption of the absolute advantage is not necessary. The assumption of the absolute advantage is a sufficient condition that guarantees that agents who have higher  $i$  will receive higher utility. As long as agents with higher ability can receive higher utility the assumption of the absolute advantage is not necessary.



must be adjusted by the ratio of the curvatures of marginal disutility of skilled human capital accumulation and unskilled human capital accumulation.<sup>4</sup>

At this point, note that (1) is different from the assumptions in Saez in several ways. In Saez, he assumed that heterogeneity of individuals is incorporated in the utility function, not the earning equation. Thus, once earning is conditioned by the human capital level, individual level heterogeneity of ability does not play any systematic role for explaining earnings. On the other hand, in (1) even after conditioned by the level of human capital, heterogeneity of ability plays a systematic role for explaining earnings and it induces higher earning for agents with higher ability. In addition, the relative return from one efficient unit of skilled human capital to unskilled human capital is higher for the agent who has higher ability than for the agent who has lower ability. As we discussed in the introduction, this interaction term between heterogeneity of ability and the return of human capital plays a crucial role in the present paper.

As for the objective of the government, we assume that the social planner will maximize the following utilitarian social welfare function:

$$\int_1^2 \{u(c_{1i}, c_{2i}) - f_s(h_i^s) - f_u(h_i^u)\} n_i di . \quad (3)$$

As for prices, we normalize the producer price and the consumer price of good 1 to one. Let  $p_2$ ,  $q_2$  and  $p_2^*$  be the consumer price and the producer price and the international price of good 2, respectively. As the purpose of this section is to examine whether introducing production distortion can increase the social welfare or not, we consider imposing a tariff on good 2. Although a tariff introduces not only a production distortion but also a consumption distortion, the first order effect of consumption distortion on welfare can be ignored as we will demonstrate. Let  $\sigma$

---

<sup>4</sup>The reason that the terms of  $g'_s/g_s$  and  $g'_u/g_u$  need to be adjusted by the curvature of the marginal disutility is as follows. For illustration, consider a condition that  $g'_s/g_s$  and  $g'_u/g_u$  must satisfy when agents with ability  $i + \epsilon$  and agents with ability  $i$  have the same degree of comparative advantage under the assumption of  $\gamma_s > \gamma_u$ . The assumption of  $\gamma_s > \gamma_u$  implies that the marginal disutility of skilled human capital changes faster than the marginal disutility of unskilled human capital when the amount of skilled and unskilled human capital respectively changes at the same rate. Note that the marginal disutility per return of skilled and unskilled human capital must be equal at the margin. This implies that in order that agents with ability  $i + \epsilon$  and agents with ability  $i$  have the same degree of comparative advantage,  $g'_s/g_s$  must be smaller than  $g'_u/g_u$ .

be a size of a tariff on good 2. Then, we will have

$$p_2 = q_2 = p_2^* + \sigma. \quad (4)$$

As for the equations determining the returns from skilled and unskilled human capital, we assume the standard two sector Heckscher-Ohlin model. In this economy, there are two sectors. Sector 1 is the skilled human capital intensive sector and it produces good 1. Sector 2 is the unskilled human capital intensive sector and it produces good 2. Each sector uses both skilled and unskilled human capital. Consumers (workers) are perfectly mobile between two sectors. When an agent who has  $h_i^s$  units of skilled human capital and  $h_i^u$  units of unskilled human capital works in sector  $k$ , it means that sector  $k$  uses  $g_s(i) \times h_i^s$  units of skilled human capital and  $g_u(i) \times h_i^u$  units of unskilled human capital. Each sector behaves as a price taker and maximizes its profit. Let  $F^k(H_k^s, H_k^u)$  be the production function in sector  $k = 1, 2$  where  $H_k^s$  and  $H_k^u$  are the total amount of skilled human capital and unskilled human capital used in sector  $k$ . We assume that  $F^k(H_k^s, H_k^u)$  exhibits constant returns to scale and it is concave with respect to both arguments. Let  $c^k(w^s, w^u)$  be the cost function in sector  $k$  to produce one unit of output in sector  $k$  when the returns of one efficient unit of skilled human capital and unskilled human capital are  $w^s$  and  $w^u$ , respectively. When both good 1 and good 2 are produced at the equilibrium,  $w^s$  and  $w^u$  are determined

$$1 = c_1(w^s, w^u) \text{ and } q_2 = c_2(w^s, w^u), \quad (5)$$

From the Stolper -Samuelson theorem,  $\partial w^s / \partial q < 0$  and  $\partial w^u / \partial q > 0$ .

The output of both goods are determined from the following factor market equilibrium conditions:

$$\frac{\partial c_1}{\partial w^s} y^1 + \frac{\partial c_2}{\partial w^s} y^2 = \int_1^2 g_s(i) \times h_i^s \times n_i di, \text{ and } \frac{\partial c_1}{\partial w^u} y^1 + \frac{\partial c_2}{\partial w^u} y^2 = \int_1^2 g_u(i) \times h_i^u \times n_i di \quad (6)$$

Although the output of both goods can be calculated from equation (6), it is more useful to work on the production possibility frontier for analytical reasons. Let  $H^s$  and  $H^u$  be the total amount of skilled human capital and unskilled human capital in this economy and define a

production possibility frontier as  $\Gamma(H^s, H^u)$ . Since the production functions are concave and the factor intensity of the two sectors are different, the production possibility set is convex. Let the producer price of good 1 and good 2 be 1 and  $q_2$ . Then, the output of good 1 and good 2 are determined as the solution of the following constrained maximization problem:

$$\max y_1 + q_2 y_2 \quad \text{s.t.} \quad (y_1, y_2) \in \Gamma(H^s, H^u) = 0$$

Thus, the output of good 1 and good 2 can be thought as a function of  $q_2$ ,  $H^s$  and  $H^u$ . Let  $Y(q, H_u, H_u)$  be the output function of good 2. At the optimum, the slope of production possibility set is equal to the relative producer price of good 2. Thus, we obtain  $Y_q \equiv \partial Y / \partial q_2 > 0$ . The Rybczynski theorem shows that  $Y_{H^u} \equiv \partial Y / \partial H^u > 0$  and  $Y_{H^s} \equiv \partial Y / \partial H^s < 0$ .

The purpose of the social planner is to maximize the utilitarian social welfare function. Given the additive separable utility function and the utilitarian social welfare function, the social planner wants to redistribute income from those who have higher ability to those who have lower ability. On the other hand, since the social planner cannot observe individual ability but rather individual earning, the social planner needs to design a non-linear income tax system  $T(R)$  to redistribute income where  $T(R)$  is a tax liability function and  $R$  is pre-tax income.

Before designing an income tax system, it is useful to consider the problem of designing the non-linear income tax system in two steps. The first step is to know how an individual  $i$  will choose skilled human capital and unskilled human capital to generate pre-tax income,  $R$ . The second step is to know, given an after-tax-income schedule of  $X = R - T(R)$ , how each individual chooses pre-tax income.

The first stage of the problem can be solved considering the following programming problem:

$$\begin{aligned} \min & f_s(h_i^s) + f_u(h_i^u) & (7) \\ \text{s.t.} & R = w_i^s \times h_i^s + w_i^u \times h_i^u \\ & \text{where } w_i^s = g_s(i) \times w^s \text{ and } w_i^u = g_u(i) \times w^u \end{aligned}$$

Let the minimized value of the above problem be  $Z(w_i^s, w_i^u, R)$ .  $Z(w_i^s, w_i^u, R)$  is the minimized

disutility to generate the pre-tax income  $R$  for an agent whose net returns from skilled human capital and unskilled capital are  $w_i^s$  and  $w_i^u$ , respectively. We denote the solution of the above problem as  $h_i^s(w_i^s, w_i^u, R)$  and  $h_i^u(w_i^s, w_i^u, R)$ . For the analysis later, it is useful to calculate compensated human capital supply. Consider the following dual problem of (7):

$$\begin{aligned} E(w_i^s, w_i^u, V) &\equiv \max w_i^s h_i^s + w_i^u h_i^u \\ \text{st. } f_s(h_i^s) + f_u(h_i^u) &\leq V \end{aligned}$$

Let the solution of the above problem be  $\tilde{h}_i^j(w_i^s, w_i^u, V)$  where  $j = s, u$ . Then, from the dual relationship, we will have

$$h_i^j(w_i^s, w_i^u, E(w_i^s, w_i^u, V)) \equiv \tilde{h}_i^j(w_i^s, w_i^u, V) \quad j=s,u$$

By taking derivative on both sides, we will have the Slutsky equation for  $h_i^s$  and  $h_i^u$ :

$$\frac{\partial h_i^j}{\partial w_i^s} + \frac{\partial h_i^j}{\partial R} h_i^s = \frac{\partial \tilde{h}_i^j}{\partial w_i^s} \quad \text{and} \quad \frac{\partial h_i^j}{\partial w_i^u} + \frac{\partial h_i^j}{\partial R} h_i^u = \frac{\partial \tilde{h}_i^j}{\partial w_i^u} \quad j=s,u$$

Note that the indifference curve of  $f_s(h_i^s) + f_u(h_i^u)$  is strictly concave. Therefore,  $\partial \tilde{h}_i^s / \partial w_i^s > 0$ ,  $\partial \tilde{h}_i^u / \partial w_i^u < 0$ ,  $\partial \tilde{h}_i^u / \partial w_i^s > 0$  and  $\partial \tilde{h}_i^s / \partial w_i^u < 0$ . This relationship means that if an individual maximizes his earnings holding the total disutility constant, an increase of the net return from skilled human capital will increase the supply of skilled human capital and an increase of the return of unskilled human capital will decrease the supply for skilled human capital. As for the properties of  $Z$ , let the Lagrangian multiplier of the disutility minimization problem be  $\alpha_i$ . Then,  $f'_s(h_i^s) = \alpha_i w_i^s$ ,  $f'_u(h_i^u) = \alpha_i w_i^u$ ,  $Z_{w_i^s} \equiv \partial Z / \partial w_i^s = -\alpha_i h_i^s$ ,  $Z_{w_i^u} \equiv \partial Z / \partial w_i^u = -\alpha_i h_i^u$  and  $Z_R \equiv \partial Z / \partial R = \alpha_i$ .

Let  $X(R)$  be the after-tax income schedule that the government designed. Then, at the second stage of the problem, given  $Z(w_i^s, w_i^u, R)$  and  $X(R)$ , each individual  $i$  will maximize his utility:

$$\max_{\{R\}} U(p_2, X(R)) - Z(w_i^s, w_i^u, R)$$

where  $U(p_2, x)$  is the indirect utility function from the consumption of two goods when the consumer price of good 2 is  $p_2$  and the after tax income is  $x$ . The objective of the social planner

is to design a schedule of  $X(R)$  to maximize the social welfare. On the other hand, the Revelation Principle shows that without loss of generality we can focus on the incentive compatible revelation mechanism. Thus, let  $(R_j, X_j)$  be the pre-tax income and after tax income when an agent announces that his type is  $j$ . Then define  $v(i)$  and  $\widehat{v}(j; i)$  as follows:

$$v(i) = \max_{\{j\}} U(p_2, X_j) - Z(w_i^s, w_i^u, R_j)$$

$$\widehat{v}(j; i) = U(p_2, X_j) - Z(w_i^s, w_i^u, R_j)$$

$v(i)$  is the maximized utility given the schedule of  $(R_j, X_j)$  and  $\widehat{v}(j; i)$  is the indirect utility when agent  $i$  announces that he is type  $j$ . The incentive compatibility condition implies that the type  $i$  agent has an incentive to announce that he is type  $i$ :

$$i = \arg \max_{\{j\}} \widehat{v}(j; i)$$

Assuming the differentiability of  $(X_j, R_j)$ , the first order condition of the incentive compatibility condition is

$$\left. \frac{\partial \widehat{v}(j, i)}{\partial j} \right|_{j=i} = \frac{\partial U}{\partial x} \frac{\partial x}{\partial j} - \frac{\partial Z}{\partial R} \frac{\partial R}{\partial j} = 0$$

On the other hand, by using the above first order condition, we have  $dv/di = -Z_{w_i^s} \times (dw_i^s/di)$ . Since  $\alpha_i$  is the Lagrangian multiplier of the required income constraint in the disutility minimization problem (7), from the FOC of the minimization problem we obtain

$$\frac{dv}{di} = \alpha_i R_i \left\{ \frac{g'_s(i)}{g_s(i)} \theta_{si} + \frac{g'_u(i)}{g_u(i)} \theta_{ui} \right\} \quad \text{where } \theta_{ji} = \frac{w_i^j h_i^j}{R_i} \quad (8)$$

Because of the assumption from the absolute advantage,  $dv/di > 0$ . (8) has a clear economic meaning. It means that the slope of the value function  $v(i)$  is proportional to the weighted average of the absolute advantage of skilled human capital accumulation and unskilled human capital accumulation. For analytical reasons, it is useful to eliminate  $\alpha_i$  in the above equation.

Using the first order condition for  $h_i^s$  and  $h_i^u$ , we can rewrite (8) as follows:

$$\frac{dv}{di} = \frac{g'_s(i)}{g_s(i)} f'_s(h_i^s) h_i^s + \frac{g'_u(i)}{g_u(i)} f'_u(h_i^u) h_i^u. \quad (9)$$

Given (9), as Mirrlees (1971) pointed out, it is more useful to assume that the social planner controls  $v_i$  and  $R_i$ .<sup>5</sup> Then,  $x_i$  is defined by the following relationship:

$$v(i) = U(p_2, X_i) - Z(w_i^s, w_i^u, R_i). \quad (10)$$

Let  $x(R, v, p_2, w_i^s, w_i^u)$  be the solution that solves (10) about  $X$ . Obviously,  $\partial x/\partial v = (U_x)^{-1}$ ,  $\partial x/\partial R_i = Z_R/U_x$  and  $\partial x/\partial p = -(U_p)/(U_x)$ ,  $\partial x/\partial w_i^s = Z_{w_i^s}/(U_x)$  and  $\partial x/\partial w_i^u = Z_{w_i^u}/(U_x)$ .

Finally, the government budget constraint implies that

$$\int_1^2 n_i \{R_i - x_i\} di + \sigma \left\{ \int_1^2 c_{2i} n_i di - Y(p_2^* + \sigma, H^s, H^u) \right\} = 0$$

The problem of the social planner is to solve the following constrained optimization program:

$$\begin{aligned} W(\sigma) &= \max_{\{R_i, v_i\}} \int_1^2 v(i) n_i di . \\ \text{st. } \frac{dv}{di} &= \frac{g'_s(i)}{g_s(i)} f'_s(h_i^s) h_i^s + \frac{g'_u(i)}{g_u(i)} f'_u(h_i^u) h_i^u \\ \int_1^2 n_i \{R_i - x_i(V_i)\} di + \sigma \left\{ \int_1^2 c_{2i} n_i di - Y(p_2^* + \sigma, H^s, H^u) \right\} &= 0 \\ &\sigma \text{ is given.} \end{aligned}$$

In the above programming problem,  $W(\sigma)$  is the maximized social welfare for given  $\sigma$ . Also note that  $h_i^s$  and  $h_i^u$  are functions of  $(R_i, w_i^s, w_i^u)$  and that  $w_i^s$  and  $w_i^u$  are the functions of  $\sigma$ .

Our interest is to know whether a change of  $\sigma$  from 0 will increase the social welfare or not. Analytically, by calculating  $dW/d\sigma$ , and evaluating at  $\sigma = 0$ , we can check whether introducing a distortion in production side (and consumption side too) can increase the social welfare. Let  $\mu_i$  and  $\lambda$  be the Lagrangian multiplier of the incentive compatibility constraint and the government

---

<sup>5</sup>One might think that the local incentive compatibility constraints are not sufficient for the global incentive compatibility constraints. On the other hand, the literature of the mechanism design shows that a single crossing property (SCP) and the monotonicity constraints are sufficient conditions for local incentive compatibility constraints to satisfy the global incentive compatibility constraints. (Fudenberg and Tirole, 1992). In this paper, we assume that the monotonicity constraints are always satisfied. This assumption is equivalent to assuming that there is no bunching. Many of the previous papers assumed that there is no bunching at the optimum. (Konishi 1995, Naito 1998). As for SCP, we can check it by examining  $\frac{\partial^2 Z}{\partial R \partial i} > 0$ . This is true as long as  $\frac{\partial h_i^s}{\partial R} > 0$ .

budget constraint. By using the envelope theorem, we obtain

$$\begin{aligned} \frac{dW}{d\sigma} \Big|_{\sigma=0} &= - \int_1^2 \mu_i \frac{d[f'(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \left\{ \frac{dh_i^s}{dw_i^s} \frac{dw_i^s}{d\sigma} + \frac{dh_i^s}{dw^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &\quad - \int_1^2 \mu_i \frac{d[f'(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \left\{ \frac{dh_i^u}{dw_i^u} \frac{dw_i^u}{d\sigma} + \frac{dh_i^u}{dw^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &\quad + \lambda \left\{ \int_1^2 c_{2i} n_i di - Y(p_2^* + \sigma, H^s, H^u) \right\} + \lambda \int_1^2 \left( -\frac{\partial x_i}{\partial p} - \frac{\partial x_i}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma} - \frac{\partial x_i}{\partial w^u} \frac{\partial w^u}{\partial \sigma} \right) n_i di \end{aligned}$$

After several calculations, we can obtain the following equation (See Appendix):

$$\frac{dW}{d\sigma} \Big|_{\sigma=0} = - \int_1^2 \mu_i \left\{ \gamma_s(g'_s/g_s) - \gamma_u(g'_u/g_u) \right\} f'_s(h_i^s) \left[ \frac{\partial \tilde{h}_i^s}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma} + \frac{\partial \tilde{h}_i^s}{\partial w^u} \frac{\partial w_i^u}{\partial \sigma} \right] di > 0 \quad (11)$$

Because of the property of the compensated supply function of  $h_i^s$ ,  $\partial \tilde{h}_i^s / \partial w_i^s > 0$  and  $\partial \tilde{h}_i^s / \partial w^u < 0$ . From the Stolper-Samuelson theorem,  $\partial w_i^s / \partial \sigma < 0$  and  $\partial w_i^u / \partial \sigma > 0$ . From the assumption on comparative advantage,  $\gamma_s(g'_s/g_s) - \gamma_u(g'_u/g_u) > 0$ . As for the sign of the Lagrangian multiplier of the incentive compatibility constraint, the standard argument shows that  $\mu_i \geq 0$  for all  $i$ . (See Appendix). Thus, we obtain  $dW/d\sigma > 0$ .

**Proposition 1** *Suppose that at the zero distortion on production and the consumption in a small open economy the social planner sets the income tax structure to maximize the social welfare function in an endogenous skill accumulation model. Then an introduction of a tariff (export subsidy) on an unskilled-labor-intensive good will increase the social welfare.*

The above equation (11) has several implications. For an illustration, consider a situation where the disutility functions of skilled and unskilled human capital accumulation have the same degree of curvature, i.e.  $\gamma_s = \gamma_u \equiv \gamma$ . Then, (11) shows that if  $(g'_s/g_s) = (g'_u/g_u)$ ,  $dW/d\sigma = 0$ . In other words, if there is no comparative advantage and if higher ability individuals are as good at accumulating skilled and unskilled human capital as lower ability individuals, then there is no welfare gain from changing the returns of skilled and unskilled human capital. Second, note that  $(\partial \tilde{h}_i^s / \partial w_i^s)(\partial w_i^s / \partial \sigma)$  and  $(\partial \tilde{h}_i^s / \partial w^u)(\partial w_i^u / \partial \sigma)$  measure how changes of returns from each type of human capital change the compensated supply of skilled human capital. Also note that

$\gamma \times f'_s(h_i^s) = f''_s(h_i^s)h_i^s + f'_s(h_i^s)$  and that  $f''_s(h_i^s)h_i^s + f'_s(h_i^s)$  is related with a change of  $\dot{v}$ . In addition, note that  $\mu_i$  measures how the social welfare increases when the incentive compatibility is relaxed. This implies that the term after the integration measures how a compensated change of the returns from skilled and unskilled human capital changes the slope of  $\dot{v}$  and increases the social welfare. Also the calculation needed for obtaining the equation shows that the effect of consumption distortion on welfare is zero, because as long as  $\sigma$  is small, such a distortion is of the second order.<sup>6</sup>

The intuition of the above proposition is as follows. In a situation where higher ability individuals have comparative advantage in accumulating skilled human capital and lower ability individuals have comparative advantage in accumulating unskilled human capital, a decrease of the return from skilled human capital and an increase of the return from unskilled human capital will hurt higher ability individuals and benefit lower ability individuals. If the social planner is interested in redistributing income from high ability individuals to low ability individuals, such changes of the returns from skilled and unskilled capital can indirectly redistribute income. On the other hand, starting from zero distortion, the deadweight loss of the production distortion is of the second-order but the welfare gain of relaxing the incentive problem has the first-order effect. As a result, introducing the production distortion increases the social welfare.

### 3 Extension: A case of Perfect substitute

In the previous section, we have assumed that two types of human capital are imperfect substitutes in order to assume differentiability of the human capital accumulation functions. As a result, people always accumulate both types of human capital. In reality, however, people sometimes accumulate only one type of human capital and, as a result, the choice of human capital becomes discrete. The purpose of this section is to analyze the welfare effect of direct versus indirect redistribution when human capital accumulation is endogenous and different types of human

---

<sup>6</sup>This can be easily checked from  $\lambda \int_1^2 c_{2i} n_i di = \lambda \int_1^2 (-\frac{\partial x_i}{\partial p}) n_i di$  in the Appendix.



capital are perfect substitutes.<sup>7</sup>

In this section, because the assumption that two types of human capital are perfect substitutes, we assume the following utility function for agent  $i$ :

$$u(c_{1i}, c_{2i}) = a^s h_i^s + a^u h_i^u$$

where  $u(c_{1i}, c_{2i})$  is strictly increasing with each argument and strictly concave.

As in the previous section, we assume that following comparative advantage condition holds:

$$\frac{g'_s(i)}{g_s(i)} > \frac{g'_u(i)}{g_u(i)} \quad (12)$$

The economic meaning of the above equation is the same as before. When two types of skill accumulation are perfect substitutes in the disutility function, the agent always solves the following constrained disutility minimization problem:

$$Z(w_i^s, w_i^u, R) \equiv \min a_s h_i^s + a_u h_i^u$$

$$\text{st } R = w_i^s h_i^s + w_i^u h_i^u$$

$$\text{where } w_i^s = g_s(i) \times w^s \text{ and } w_i^u = g_u(i) \times w^u$$

In the above problem, for an agent with ability  $i$ , if  $a_s/a_u < w_i^s/w_i^u$  he will accumulate only skilled human capital and if  $a_s/a_u > w_i^s/w_i^u$ , he will accumulate only unskilled human capital. Note that because of the assumption of comparative advantage (12),  $w_i^s/w_i^u$  is an increasing function of  $i$ . Let  $i^*$  be  $i$  that satisfies  $(w^s \times g_s(i))/(w^u \times g_u(i)) = a_s/a_u$ . Then, agents whose ability is greater than  $i^*$  accumulate only skilled human capital and agents whose ability  $i$  is less than  $i^*$  accumulate only unskilled human capital. We assume that such  $i^*$  is located within 1 and 2.<sup>8</sup>

---

<sup>7</sup>Besides the reason mentioned in the previous section, conducting a welfare analysis when individual behavior includes a discrete choice is useful from a theoretical standpoint as well. In many important economic situations such as the choice of location to live, the choice of technology by firms and labor market participation, decisions made by consumers or firms include discrete choices. Until very recently, a welfare analysis that includes discrete choices was rare. As far as the author knows, only Boadway and Cuff (2001) started to investigate this issue very recently. They analyzed an optimal taxation problem when some individuals are bunched at the bottom. Another purpose of this section is to contribute to such a literature as well.

<sup>8</sup>This assumption is not so restrictive as the following reason. For example, if  $i^*$  is greater than 2, all agents

Given such  $i^*$ ,  $Z(w_i^s, w_i^u, R)$  is

$$\begin{aligned} Z(w_i^s, w_i^u, R) &= a^s \left( \frac{R}{w_i^s} \right) \text{ for } i^* \leq i \leq 2 \\ Z(w_i^s, w_i^u, R) &= a^u \left( \frac{R}{w_i^u} \right) \text{ for } 1 \leq i < i^*. \end{aligned}$$

Let  $X(R)$  be an after-tax income schedule that the government designs. Then, each agent chooses his best  $R$  to maximize  $U(p_2, X(R)) - Z(w_i^s, w_i^u, R)$ . Once  $R$  is chosen, an agent chooses his optimal skill type and accumulates human capital to generate pre-tax income  $R$ . Let  $\tilde{v}(i)$  be the maximized value given the schedule  $X(R)$ :

$$\tilde{v}(i) \equiv \max_R U(p_2, X(R)) - Z(w_i^s, w_i^u, R).$$

For the analysis of the optimal schedule of  $X(R)$ , we assume that the schedule of  $X(R)$  is a continuous function. Although it is possible that the optimal schedule of  $X(R)$  is not continuous, the tax schedules of almost of all developed countries are continuous. When  $X(R)$  is a continuous function, it is straightforward to show that  $\tilde{v}(i)$  is continuous with respect to  $i$  from the theory of the maximum (Berg 1963). In addition, there is an interesting property on  $\tilde{v}(i)$  in the neighborhood of  $i^*$  that turns out to be crucial for our result. The following lemma shows that property of  $\tilde{v}(i)$ .

**Lemma 1** *When  $i$  increases, the graph of  $\tilde{v}(i)$  has a counter-clockwise kink at  $i^*$ .*

Proof. Let  $\tilde{v}_s(i)$  be the maximized utility of an agent with ability  $i$  given the tax schedule when he can accumulate only skilled human capital. Also, let  $\tilde{v}_u(i)$  be the maximized utility of an agent with ability  $i$  when he can accumulate only unskilled human capital. By the definition, the graph of  $\tilde{v}(i)$  is the upper envelope of  $\tilde{v}_s(i)$  and  $\tilde{v}_u(i)$  and  $i^*$  is at the intersection between  $\tilde{v}_s(i)$  and  $\tilde{v}_u(i)$ . This implies that there is a counter-clockwise kink at  $i^*$  (See also Figure 1). ■

---

will accumulate only unskilled human capital. However, the production needs both skilled and unskilled human capital. As a result, the return from skilled human capital will start to increase and the return from unskilled human capital will start to decrease. This implies that  $i^*$  will start to decrease. This process will continue until some agents start to accumulate skilled human capital.

Now consider the problem of designing a nonlinear income tax system. As in the previous section, we define  $v(i)$  as follows:

$$v(i) = \max_j U(p_2, X_j) - Z(w_i^s, w_i^u, R_j)$$

By using the same technique as in the previous section, we can calculate  $dv(i)/di$  for  $i$  in  $(1, i^*)$  and  $(i^*, 2)$ .

$$\frac{dv}{di} = a^s \frac{g'_s}{g_s} \frac{R_i}{g_s(i)w^s} \text{ for } i \in (i^*, 2) \quad (13)$$

$$\frac{dv}{di} = a^u \frac{g'_u}{g_u} \left( \frac{R_i}{g_u w^u} \right) \text{ for } i \in (1, i^*) \quad (14)$$

Next we will check a single crossing property of the utility function  $U(p_2, X) - Z(R, w^s, w^u, R)$ .

The marginal rate of substitution between  $X$  and  $R$  is

$$\begin{aligned} \text{MRS}(R, X) &= \frac{1}{U_x} \frac{a^s}{g_s w^s} \text{ for } i \in (i^*, 2) \\ &= \frac{1}{U_x} \frac{a^u}{g_u w^u} \text{ for } i \in (1, i^*) \end{aligned}$$

Thus the  $\text{MRS}(R, X)$  is a decreasing function of  $i$  and a single crossing property is satisfied. This means that the local incentive compatibility and the monotone condition of  $R$  are sufficient conditions for the global incentive compatibility (Fudenberg and Tirole, 1991). We assume that the monotonicity constraint is not binding.

As in the previous section, it is useful to think that the government controls  $v(i)$  and  $R_i$  and that  $X_i$  is defined from the following relationship:

$$v(i) = U(p_2, X) - Z(w_i^s, w_i^u, R_j)$$

Finally for analytical convenience, rewrite the first order condition of (13) and (14) :

$$\dot{v}^s = \frac{g'_s}{g_s} a^s h_i^s \quad \text{and} \quad \dot{v}^u = \frac{g'_u}{g_u} a^u h_i^u.$$

Based on the setup, the purpose of the government is to solve the following programming problem:

$$W(\sigma) = \max \int_1^{i^*} v^u(i) n_i di + \int_{i^*}^2 v^s(i) n_i di$$

$$\text{st. } \dot{v}^s = \frac{g'_s}{g_s} a^s h_i^s \quad \text{for } i^* < i \leq 2 \quad (\text{IC1})$$

$$\dot{v}^u = \frac{g'_u}{g_u} a^u h_i^u \quad \text{for } 1 < i < i^* \quad (\text{IC2})$$

$$v^s(i^*) = v^u(i^*) \quad (\text{BD1})$$

$$R_{i^*}^s = R_{i^*}^u \quad (\text{BD2})$$

$$\int_{i^*}^2 \{R_i^s - x(R_i^s, v_i^s, w_i^s, w_i^u)\} n_i di$$

$$+ \int_1^{i^*} \{R_i^u - x(R_i^u, v_i^u, w_i^s, w_i^u)\} n_i di$$

$$+ \sigma \left\{ \int_1^2 n_i c_{2i} di - Y(p_2^* + \sigma, H^s, H^u) \right\} \geq 0 \quad (\text{RC})$$

where

$$H^s = \int_{i^*}^2 h_i^s g_s(i) n_i di \quad \text{and} \quad H^u = \int_1^{i^*} h_i^u g_u(i) n_i di$$

The above programming problem deserves several comments. First, (IC1) and (IC2) are the local incentive compatibility constraints. Second, (BD1) comes from the assumption that the tax schedule that the government designs is continuous and, as a result, the utility level of the agents must be continuous. (BD2) comes from the assumption that individual  $i^*$  chooses only one  $R$ . Now let  $\mu_i^s, \mu_i^u$  and  $\lambda$  be the Lagrangian multipliers of (IC1), (IC2) and (RC). Let  $\beta_1$  and  $\beta_2$  be the Lagrangian multipliers of (BD1) and (BD2). The first order conditions can be calculated and we will write them in the Appendix to save the space. Then, what we need to know is the effect of increasing  $\sigma$  from zero on the social welfare, which is equivalent to  $dW/d\sigma$ . By using the envelope theorem, we have (See Appendix)

$$\left. \frac{dW}{d\sigma} \right|_{\sigma=0} = \frac{\partial i^*}{\partial \sigma} \left\{ \mu_{i^*}^s a^s h_{i^*}^s \frac{g'_s}{g_s} - \mu_{i^*}^u a^u h_{i^*}^u \frac{g'_u}{g_u} \right\} \quad (15)$$

From the FOC of  $v_{i^*}^s$  and  $v_{i^*}^u$ , we have  $\mu_{i^*}^s = \mu_{i^*}^u$ . In addition, as we show in the Appendix  $\mu_i^s$  and  $\mu_i^u$  are always non-negative. Furthermore  $a^s h_{i^*}^s (g'_s/g_s)$  and  $a^u h_{i^*}^u (g'_u/g_u)$  are the right hand

slope of  $v_i^s$  and the left hand slope  $v_i^u$  at  $i^*$ . From Lemma 1, the slope of  $v_i^s$  is steeper than the slope of  $v_i^u$  at  $i^*$ . Since  $\frac{\partial i^*}{\partial \sigma} > 0$ , we have  $dW/d\sigma > 0$ .

**Proposition 2** *Consider a small open economy where individuals accumulate human capital endogenously and different types of human capital are perfect substitutes. Suppose that the social planner designs a nonlinear income tax system to maximize the utilitarian social welfare function without any production distortion and that there is no-bunching at the switching point  $i^*$ . Then, introducing a tariff on an unskilled human capital intensive good will increase the social welfare.*

At this point, it would be useful to consider the economic meaning of (15). Figure 1 shows the graph of  $\tilde{v}(i)$ ,  $\tilde{v}_s(i)$  and  $\tilde{v}_u(i)$ . When the government increases the tariff  $\sigma$  from zero, the graph of  $\tilde{v}_s(i)$  will shift downward and the graph of  $\tilde{v}_u(i)$  shifts upward. As a result,  $i^*$  will increase. Also, notice that from (IC1) and (IC2), the slope of  $\tilde{v}_s(i)$  increases and the slope of  $\tilde{v}_u(i)$  decreases.

In the mechanism designs problem,  $\dot{v}$ , the slope of the value function, is related with how the compensation schedule must be sensitive with unobserved ability. When  $\dot{v}$  is higher, it means that the social planner needs to give higher utility to those with higher ability. With redistributive social welfare function, the social planner wants to give higher utility to agents with lower ability. Thus, when  $\dot{v}$  is high, the level of utility that the social planner can give to the agents with lower ability is limited since the amount of the resource is limited. In such a situation, if the government can make  $\dot{v}$  smaller exogenously, it is possible to increase the social welfare and changing  $\sigma$  can be a good policy tool for changing  $\dot{v}$ .

When  $\sigma$  increases, the change of  $\dot{v}$  is not the same for all individuals however. As Figure 1 shows, all individuals whose ability is lower than  $i^*$  will experience a decrease of  $\dot{v}$  and all individuals whose ability is greater than  $i^*$  will experience an increase of  $\dot{v}$  except the neighborhood of  $i^*$ . But, as the analysis in the Appendix shows, the effect of a change of  $\dot{v}$  for those agents is of the second order and can be replicated by the adjustment of the nonlinear income system. On

the other hand, there are some individuals who experience the first order change of  $\dot{v}$ . Individuals whose ability is in  $(i^*, i^* + \partial i^*/\partial\sigma)$  will switch from accumulating skilled human capital to unskilled human capital. Since the graph of  $v(i)$  has a counter-clockwise kink at  $i^*$ , individuals in  $(i^*, i^* + \partial i^*/\partial\sigma)$  will experience the first order decrease of  $\dot{v}$ . This implies that the government needs less ability-sensitive compensation schedules for those agents. Because this change of  $\dot{v}$  has the first order effect, it will increase the social welfare.

(15) can be interpreted in terms of the marginal tax schedule as well. Note that  $(\partial Z/\partial R^m)/U_x$  is equal to  $1 - T_i^m$  where  $T_i^m$  is the marginal tax rate of income of those who accumulated  $m = s, u$  type of skill and his ability is equal to  $i$ . From the FOC of  $R_i^s$  and  $R_i^u$ ,

$$\lambda n T_i^s = \mu_i^s \times a^s \frac{\partial h_i^s}{\partial R_i^s} \frac{g'_s}{g_s} \text{ and } \lambda n T_i^u = \mu_i^u \times a^u \frac{\partial h_i^u}{\partial R_i^u} \frac{g'_u}{g_u}$$

Thus, Since  $(\partial h/\partial R) \times R = h$ , we have

$$\left. \frac{dW}{d\sigma} \right|_{\sigma=0} = \frac{\partial i^*}{\partial \sigma} \lambda n (R_{i^*}^s T_{i^*}^s - R_{i^*}^u T_{i^*}^u).$$

$T_{i^*}^s$  and  $T_{i^*}^u$  are the marginal tax rates of individuals just above  $i^*$  and just below  $i^*$ , respectively. When  $\sigma$  increases, the individual just above  $i^*$  who initially accumulated skilled human capital will switch from accumulating skilled human capital to unskilled human capital. Since the marginal tax rate of those who accumulated skilled human capital is higher than the marginal tax rate for those who accumulated unskilled human capital around  $i^*$ , the marginal tax rate will decrease.<sup>9</sup> Thus,  $R_{i^*}^s T_{i^*}^s - R_{i^*}^u T_{i^*}^u$  is the earning that is affected by a change of the marginal tax rates. Since this change of the marginal tax rate is of the first order, it can increase the social welfare.

## 4 Discussion

In the above two sections, we have shown that indirect redistribution through an increase of the return from unskilled human capital and a decrease of the return from skilled human capital

---

<sup>9</sup>Readers still might wonder why the marginal tax rate for those who accumulated skilled human capital is higher than those who accumulated unskilled human capital around  $i^*$ . The reason is around the right hand side of  $i^*$ , the marginal return from ability is higher at the right hand side of  $i^*$  than at the left hand side of  $i^*$  because  $i^*$  is a switching point.

would increase the social welfare. One natural question at this point would be why such changes of returns do not cause the adverse effect on human capital accumulation and, if they cause it, why we can ignore it. The answer to such a question is that it will cause the adverse effect on human capital accumulation but a redistributive income tax system also causes such an incentive problem. In a circumstance where each individual's comparative advantage is not observable and human capital accumulation is endogenous, the redistributive income taxation necessarily introduces adverse incentive effects on human capital accumulation. In such a situation, the question is not whether redistributive change of returns from different types of skill causes the adverse incentive effect but whether it can mitigate the existing incentive problem caused by income taxation. Given that income taxation is subject to asymmetric information due to the unobservability of comparative advantage at individual levels, redistributive changes of return from different types of skill mitigates the asymmetric information problem since the government can affect agents with different types of comparative advantage differently. As a result, the redistributive changes of returns from different types of skill will increase the economic efficiency.

At this point we should emphasize that the assumption of *comparative advantage* plays a crucial role in our analysis. This implies that empirical studies that examined the returns from human capital accumulation for individual with different abilities such as Dinardo and Tobias (2001) and Tobias (2003) are important. In addition, the results in the empirical studies and the result in this paper can have important implications for public policy. For example, it might be possible that encouraging skilled human capital accumulation through government funding does not necessarily increase the social welfare if comparative advantage in human capital accumulation exists and individuals who get the benefit most from the government funding are individuals with higher ability.

In this paper, we examined the production efficiency theorem in a small open economy setting. It is worth mentioning that in a small open economy setting, the Atkinson and Stiglitz theorem holds since a commodity tax cannot affect the factor prices due to the factor equalization theorem

(Samuelson, 1949). On the other, it is possible to extend the intuition of the present paper to a closed economy setting by using the two sector-two factor general equilibrium model (Harberger, 1962). In this case, we can prove that the Atkinson and Stiglitz theorem does not hold since a commodity tax can affect the producer prices and factor prices in a closed economy. More specifically, we can prove that imposing a commodity tax on skilled human capital intensive good will increase the social welfare.<sup>10</sup>

## 5 Conclusion

In this paper, we have examined whether indirect redistribution such as tariffs and production subsidies can complement income taxation in the long run where human capital accumulation is endogenous. For that purpose, I developed two models where individuals can choose the amount of both skilled and unskilled human capital based on their *comparative advantage*. In the first model, we assumed that skilled human capital and unskilled human capital are imperfect substitutes and that individuals accumulate both skilled and unskilled human capital. In the second model, we assumed that skilled human capital and unskilled human capital are perfect substitutes and that individuals accumulate only one type of human capital. Assuming that individuals with higher ability have comparative advantage in accumulating skilled human capital, we have shown that indirect redistribution such imposing a tariff on an unskilled human capital intensive good can increase the efficiency and complement an income tax system. This suggests that the validity of the production efficiency theorem depends on how the process of human capital accumulation is modelled. The result of this paper also suggests that empirical studies such as Dinardo and Tobias (2001) and Tobias (2003) that showed the returns from human capital were different among individuals with different abilities have important implications for public policy.

---

<sup>10</sup>For a formal proof, see our previous version of the paper (Naito, 2002).



## Appendix

### Derivation of equation (11)

Let  $\mu_i$  and  $\lambda$  be the Lagrangian multiplier of the incentive constraint and the resource constraint.

Then, the Lagrangian function is

$$W(\sigma) = \int_1^2 v(i)n_i di + \int_1^2 \mu_i \left[ \frac{dv}{di} - f'_s(h_i^s)h_i^s(g'_s/g_s) - f'_u(h_i^u)h_i^u(g'_u/g_u) \right] di + \lambda \int_1^2 n_i \{R_i - x_i(v_i)\} di + \sigma \left\{ \int_1^2 n_i c_i^2 di - y_2(p_2^* + \sigma, H^s, H^u) \right\}$$

By using the integration by parts, we can obtain

$$\begin{aligned} W(\sigma) &= \int_1^2 v_i n_i di + \int_1^2 \mu_i \frac{dv}{di} di - \int_1^2 \mu_i f'_s(h_i^s)h_i^s(g'_s/g_s) di - \int_1^2 \mu_i f'_u(h_i^u)h_i^u(g'_u/g_u) di \\ &+ \lambda \int_1^2 n_i \{R_i - x_i\} di + \sigma \int_1^2 n_i c_{2i} di - \sigma Y(p_2^* + \sigma, H^s, H^u) \\ &= \int_1^2 v_i n_i di + \mu_2 v_2 - \mu_1 v_1 - \int_1^2 \dot{\mu}_i v_i di - \int_1^2 \mu_i f'_s(h_i^s)h_i^s(g'_s/g_s) di - \int_1^2 \mu_i f'_u(h_i^u)h_i^u(g'_u/g_u) di \\ &+ \lambda \int_1^2 n_i \{R_i - x_i\} di + \lambda \sigma \left\{ \int_1^2 n_i c_{2i} di - Y(p_2^* + \sigma, H^s, H^u) \right\} \end{aligned}$$

Therefore, the first-order-conditions are

$$\begin{aligned} n_i - \dot{\mu}_i - \lambda n_i \frac{\partial x_i}{\partial v_i} + \lambda \sigma \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial v_i} &= 0 \\ -\mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \frac{\partial h_i^u}{\partial R_i} - \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \frac{\partial h_i^u}{\partial R_i} + \lambda n_i - \lambda n_i \frac{\partial x_i}{\partial R_i} + \lambda \sigma n_i \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial R_i} &= 0 \\ \mu_1 = 0 \text{ and } \mu_2 = 0 & \end{aligned}$$

By using the envelope theorem, we obtain

$$\begin{aligned} \frac{dW}{d\sigma} \Big|_{\sigma=0} &= - \int_1^2 \mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \left\{ \frac{dh_i^s}{dw_i^s} \frac{dw_i^s}{d\sigma} + \frac{dh_i^s}{dw_i^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &- \int_1^2 \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \left\{ \frac{dh_i^u}{dw_i^s} \frac{dw_i^s}{d\sigma} + \frac{dh_i^u}{dw_i^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &+ \lambda \left\{ \int_1^2 c_{2i} n_i di - Y(p_2^* + \sigma, H^s, H^u) \right\} + \lambda \int_1^2 \left( -\frac{\partial x_i}{\partial p} - \frac{\partial x_i}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma} - \frac{\partial x_i}{\partial w_i^u} \frac{\partial w_i^u}{\partial \sigma} \right) n_i di \end{aligned}$$

Note that  $-\partial x_i / \partial p_2 = (U_{p_2}) / (U_x)$ . From the Roy's identity,  $(U_{p_2}) / (U_x) = -c_{2i}$ . Therefore,

$\lambda \int_1^2 c_{2i} n_i di = \lambda \int_1^2 \left( -\frac{\partial x_i}{\partial p} \right) n_i di$ . In addition,  $\frac{\partial x_i}{\partial w_i^s} = z_{w_i^s} / U_x$  and  $\frac{\partial x_i}{\partial w_i^u} = z_{w_i^u} / U_x$  and  $\frac{\partial x_i}{\partial R_i} = Z_{R_i} / U_x$ .

Using the definition of  $Z_{w_i^s}$  and  $Z_{w_i^u}$ ,  $\frac{\partial x_i}{\partial w_i^s} = -\alpha_i h_i^s / U_x$ ,  $\frac{\partial x_i}{\partial w_i^u} = -\alpha_i h_i^u / U_x$  and  $Z_R / U_x = \alpha_i / U_x$ .

On the other hand, the FOC of  $R_i$  at  $\sigma = 0$  is that

$$-\mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \frac{\partial h_i^s}{\partial R_i} - \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \frac{\partial h_i^u}{\partial R_i} + \lambda n_i = \lambda n_i \alpha_i / U_x$$

Thus,  $dW/d\sigma$  becomes

$$\begin{aligned} \frac{dW}{d\sigma} \Big|_{\sigma=0} &= - \int_1^2 \mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \left\{ \frac{dh_i^s}{dw_i^s} \frac{dw_i^s}{d\sigma} + \frac{dh_i^s}{dw_i^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &\quad - \int_1^2 \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \left\{ \frac{dh_i^u}{dw_i^s} \frac{dw_i^s}{d\sigma} + \frac{dh_i^u}{dw_i^u} \frac{dw_i^u}{d\sigma} \right\} di \\ &\quad - \lambda y_2 + \int_1^2 [-\mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \frac{\partial h_i^s}{\partial R_i} - \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \frac{\partial h_i^u}{\partial R_i} + \lambda n_i] h_i^s \frac{\partial w_i^s}{\partial \sigma} di \\ &\quad + \int_1^2 [-\mu_i \frac{d[f'_s(h_i^s)h_i^s(g'_s/g_s)]}{dh_i^s} \frac{\partial h_i^s}{\partial R_i} - \mu_i \frac{d[f'_u(h_i^u)h_i^u(g'_u/g_u)]}{dh_i^u} \frac{\partial h_i^u}{\partial R_i} + \lambda n_i] h_i^u \frac{\partial w_i^u}{\partial \sigma} di \\ &\quad - \lambda y_2 + \int_1^2 \lambda n_i h_i^s \frac{\partial w_i^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w_i^u}{\partial \sigma} di \end{aligned}$$

Note that  $\int_1^2 \lambda n_i h_i^s \frac{\partial w_i^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w_i^u}{\partial \sigma} di = \int_1^2 \lambda n_i h_i^s i \frac{\partial w^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w^u}{\partial \sigma} di$ .  $\int_1^2 \lambda n_i h_i^s i \frac{\partial w^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w^u}{\partial \sigma} di$  is a change of total earning due to a tariff when levels of human capital of all individuals are fixed. On the other hand, from perfect competition, for given level of human capital of all individuals, the total revenue of the firm should be equal to the total payment to factor owners. Thus,  $y_1 + (p_2^* + \sigma)y_2 = w^s \int_1^2 \lambda n_i h_i^s i di + w^u \int_1^2 \lambda n_i h_i^u di$  always holds. Let  $Q(\sigma)$  be the total revenue of firms when all human capital level of all individuals are fixed. Then,  $dQ/d\sigma = \int_1^2 \lambda n_i h_i^s i \frac{\partial w^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w^u}{\partial \sigma} di$ . By definition of  $Q(\sigma)$

$$Q(\sigma) = \max y_1 + (p_2^* + \sigma)y_2 \quad \text{s.t.} \quad (y_1, y_2) \in \Gamma(H^s, H^u) = 0$$

$H^s$  and  $H^u$  are fixed.

From the envelope theorem,  $\frac{dQ}{d\sigma} = y_2$ . Therefore,  $-\lambda y_2 + \int_1^2 \lambda n_i h_i^s i \frac{\partial w^s}{\partial \sigma} di + \int_1^2 \lambda n_i h_i^u \frac{\partial w^u}{\partial \sigma} di = 0$ .

Note that from the definition of  $\tilde{h}_i^s$  and  $\tilde{h}_i^u$ , we have

$$f'_s(\tilde{h}_i^s) \frac{\partial \tilde{h}_i^s}{\partial w_i^s} + f'_u(\tilde{h}_i^u) \frac{\partial \tilde{h}_i^u}{\partial w_i^s} = 0 \quad \text{and} \quad f'_u(\tilde{h}_i^s) \frac{\partial \tilde{h}_i^s}{\partial w_i^u} + f'_u(\tilde{h}_i^u) \frac{\partial \tilde{h}_i^u}{\partial w_i^u} = 0$$

By using the Slutsky equation for  $h_i^s$  and  $h_i^u$  and the above equation, we have

$$\begin{aligned} \frac{dW}{d\sigma} \Big|_{\sigma=0} &= - \int_1^2 \mu_i \left\{ \left[ \frac{f_s''(h_i^s)h_i^s}{f_s'(h_i^s)} + 1 \right] (g_s'/g_s) - \left[ \frac{f_u''(h_i^u)h_i^u}{f_u'(h_i^u)} + 1 \right] (g_u'/g_u) \right\} f_s'(h_i^s) \frac{\partial \tilde{h}_i^s}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma} di \\ &\quad - \int_1^2 \mu_i \left\{ \left[ \frac{f_s''(h_i^s)h_i^s}{f_s'(h_i^s)} + 1 \right] (g_s'/g_s) - \left[ \frac{f_u''(h_i^u)h_i^u}{f_u'(h_i^u)} + 1 \right] (g_u'/g_u) \right\} f_s'(h_i^s) \frac{\partial \tilde{h}_i^s}{\partial w_i^u} \frac{\partial w_i^u}{\partial \sigma} di \\ &= - \int_1^2 \mu_i \left\{ \left[ \frac{f_s''(h_i^s)h_i^s}{f_s'(h_i^s)} + 1 \right] (g_s'/g_s) - \left[ \frac{f_u''(h_i^u)h_i^u}{f_u'(h_i^u)} + 1 \right] (g_u'/g_u) \right\} f_s'(h_i^s) \left[ \frac{\partial \tilde{h}_i^s}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma} + \frac{\partial \tilde{h}_i^s}{\partial w_i^u} \frac{\partial w_i^u}{\partial \sigma} \right] di \end{aligned}$$

From the condition of the comparative advantage, the inside of the large bracket is positive.

Also, both  $\frac{\partial \tilde{h}_i^s}{\partial w_i^s} \frac{\partial w_i^s}{\partial \sigma}$  and  $\frac{\partial \tilde{h}_i^s}{\partial w_i^u} \frac{\partial w_i^u}{\partial \sigma}$  are positive. Thus, we have  $\frac{dW}{d\sigma} \Big|_{\sigma=0} > 0$ .

### Proof of $\mu_i \geq 0$

From the FOC of  $v_i$ , we will have  $n_i - \dot{u}_i - \lambda n_i \frac{\partial x_i}{\partial v_i} + \lambda \sigma \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial v_i} = 0$ . Thus, we have

$$n_i - \lambda n_i \frac{\partial x_i}{\partial v_i} = \mu_i$$

at  $\sigma = 0$ . By integrating both sides and using the definition of  $\frac{\partial x_i}{\partial v_i}$  and  $\mu_1 = 0$ , we will have

$$\int_1^i n_i \left\{ 1 - \frac{\lambda}{U_x} \right\} = \mu_i$$

From the first order condition of the revelation problem,  $U_x(p_2, X)X'(i) = Z_R R'(i)$ . This means that the sign of  $X'(i)$  and  $R'(i)$  are the same. Since  $v(i)$  is strictly increasing,  $X'(i)$  and  $R'(i)$  must be increasing. When  $X'(i)$  is increasing,  $\frac{\lambda}{U_x}$  is increasing. This implies that if at some  $i^{**}$ ,  $1 - \lambda/U_x = 0$ , then for any  $i > i^{**}$ ,  $1 - \lambda/U_x < 0$ . However,  $\mu_2 = 0$  from the FOC of  $v_2$ . This implies that  $\mu_1$  is initially strictly positive until  $i^{**}$  and then it begins to decrease and reaches to zero at  $i = 2$ . Therefore,  $\mu_i \geq 0$  for all  $1 \leq i \leq 2$ .

### Proof of Proposition 2

The Lagrangian is:

$$\begin{aligned}
L = & \int_1^{i^*} v^u(i)n_i di + \int_{i^*}^2 v^s(i)n_i di + \int_1^{i^*} \mu_i^u \{v^u - a^u h_i^u (g'_u/g_u)\} di + \int_{i^*}^2 \mu_i^s \{v^s - a^s h_i^s (g'_s/g_s)\} di \\
& + \beta_1 \{v_{i^*}^s - v_{i^*}^u\} + \beta_2 \{R_{i^*}^s - R_{i^*}^u\} \\
& + \lambda \int_1^{i^*} \{R_i^u - x(R_i^u, v_i^u, w_i^s, w_i^u)\} n_i di + \lambda \int_{i^*}^2 \{R_i^s - x(R_i^s, v_i^s, w_i^s, w_i^u)\} n_i di \\
& + \lambda \sigma \left\{ \int_1^2 n_i c_{2i} di - Y(p_2^* + \sigma, H^s, H^u) \right\}
\end{aligned}$$

By using the integration by parts, we obtain

$$\begin{aligned}
L = & \int_1^{i^*} v^u(i)n_i di + \int_{i^*}^2 v^s(i)n_i di + \mu_{i^*}^u v_{i^*}^u - \mu_1^u v_1^u - \int_1^{i^*} \dot{\mu}_i^u v_i^u di - \int_1^{i^*} \mu_i^u a^u h_i^u (g'_u/g_u) di \\
& \mu_2^s v_2^s - \mu_{i^*}^s v_{i^*}^s - \int_{i^*}^2 \dot{\mu}_i^s v_i^s di - \int_{i^*}^2 \mu_i^s a^s h_i^s (g'_s/g_s) di + \beta_1 \{v_{i^*}^s - v_{i^*}^u\} + \beta_2 \{R_{i^*}^s - R_{i^*}^u\} \\
& + \lambda \int_1^{i^*} \{R_i^u - x(R_i^u, v_i^u, q, w_i^s, w_i^u)\} n_i di + \lambda \int_{i^*}^2 \{R_i^s - x(R_i^s, v_i^s, w_i^s, w_i^u)\} n_i di \\
& + \lambda \sigma \left\{ \int_1^2 n_i c_{2i} di - Y(p_2^* + \sigma, H^s, H^u) \right\}
\end{aligned}$$

The first order condition for  $v_i^s$ ,  $v_2^s$ ,  $v_{i^*}^s$ ,  $R_i^s$ ,  $R_{i^*}^s$ ,  $v_i^u$ ,  $v_{i^*}^u$ ,  $v_1^u$ ,  $R_i^u$  and  $R_{i^*}^u$  are

$$v_i^s : n_i - \dot{\mu}_i^s - \lambda n_i \frac{\partial x}{\partial v_i^s} + \lambda \sigma \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial v_i^s} = 0$$

$$v_2^s : \mu_2^s = 0$$

$$v_{i^*}^s : -\mu_{i^*}^s + \beta_1 = 0$$

$$R_i^s : -\mu_i^s \times a^s \frac{\partial h_i^s}{\partial R_i^s} \frac{g'_s}{g_s} + \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^s} + \lambda \sigma n_i \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial R_i^s} = 0$$

$$R_{i^*}^s : \beta_2 = 0$$

$$v_i^u : n_i - \dot{\mu}_i^u - \lambda n_i \frac{\partial x}{\partial v_i^u} + \lambda \sigma \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial v_i^u} = 0$$

$$v_{i^*}^u : \mu_{i^*}^u - \beta_1 = 0$$

$$v_1^u : \mu_1^u = 0$$

$$R_i^u : -\mu_i^u \times a^u \frac{\partial h_i^u}{\partial R_i^u} \frac{g'_u}{g_u} + \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^u} + \lambda \sigma n_i \frac{\partial c_{2i}}{\partial x_i} \frac{\partial x_i}{\partial R_i^u} = 0$$

$$R_{i^*}^u : \beta_2 = 0$$

Now we characterize those first order conditions. First, note that at  $\sigma = 0$ ,

$$\mu_i^s = \mu_{i^*}^s + \int_{i^*}^i n_j (1 - \lambda \frac{\partial x}{\partial v_j}) dj \text{ for } i \in (i^*, 2) \text{ and } \mu_{i^*}^u = \mu_1^u + \int_1^{i^*} n_j (1 - \lambda \frac{\partial x}{\partial v_j}) dj \quad (16)$$

Thus, since  $\mu_1^u = 0$ ,  $\mu_i^s = \int_1^i n_j (1 - \lambda \frac{\partial x}{\partial v_j}) dj$  for  $i \in (i^*, 2)$ . Note that  $\frac{\partial x}{\partial v_j} = 1/(U_x)$ . A single crossing property and the monotonicity of  $R_i^s$  and  $R_i^u$  guarantee that  $x_i$  is increasing. This implies that  $\frac{\partial x}{\partial v_j}$  is increasing and the inside of the integral is a decreasing function of  $i$ . Since  $\mu_2^s = 0$ ,  $\mu_i^u$  and  $\mu_i^s$  are non-negative.

Now we examine  $dW/d\sigma$  and evaluate at  $\sigma = 0$ . From the envelope theorem,

$$\begin{aligned} \frac{dW}{d\sigma} = & \frac{\partial i^*}{\partial \sigma} \left\{ v^u(i^*)n_{i^*} - v^s(i^*)n_{i^*} + \mu_{i^*}^u v^u(i^*) + \mu_{i^*}^s v_{i^*}^u - \mu_{i^*}^u v^u(i^*) - \mu_{i^*}^u a^u h_i^u \frac{g'_u}{g_u} \right. \\ & - \mu_{i^*}^s v(i^*) - \mu_{i^*}^s \dot{v}_{i^*}^s + \dot{\mu}_{i^*}^s v(i^*) + \mu_{i^*}^s a^s h_i^s \frac{g'_s}{g_s} + \beta_1 \{v_{i^*}^s - \dot{v}_{i^*}^u\} + \beta_5 \dot{R}_{i^*}^u - \beta_5 \dot{R}_{i^*}^s \\ & \left. - \lambda \{R_{i^*}^s - x(R_{i^*}^s, v_{i^*}^s, q, w_{i^*}^s, w_{i^*}^u)\} n_{i^*} + \lambda \{R_{i^*}^u - x(R_{i^*}^u, v_{i^*}^u, q, w_{i^*}^s, w_{i^*}^u)\} n_{i^*} \right\} \\ + \frac{\partial w^s}{\partial \sigma} & \left\{ - \int_{i^*}^2 \mu_{i^*}^s a^s \frac{\partial \bar{h}_i^s}{\partial w_{i^*}^s} g'_s di - \lambda \int_{i^*}^2 \frac{\partial x}{\partial w_{i^*}^s} n_i g_s(i) di \right\} + \frac{\partial w^u}{\partial \sigma} \left\{ - \int_{i^*}^2 \mu_{i^*}^u a^u \frac{\partial \bar{h}_i^u}{\partial w_{i^*}^u} g'_u di - \lambda \int_1^{i^*} \frac{\partial x}{\partial w^u} g_u(i) n_i di \right\} \\ & + \lambda \int_1^2 n_i c_{2i} di - \lambda Y(p_2^* + \sigma, H^s, H^u) + \int_{i^*}^2 - \frac{\partial x}{\partial q} \frac{\partial q}{\partial \sigma} n_i di - \frac{\partial x}{\partial q} \frac{\partial q}{\partial \sigma} \int_1^{i^*} n_i di \end{aligned}$$

From the Roy's identity,  $c_{2i} = -\frac{\partial x}{\partial q} \frac{\partial q}{\partial \sigma}$ . Thus,

$$\begin{aligned} \frac{dW}{d\sigma} = & \frac{\partial i^*}{\partial \sigma} \left\{ \mu_{i^*}^s a^s h_i^s \frac{g'_s}{g_s} - \mu_{i^*}^u a^u h_i^u \frac{g'_u}{g_u} \right\} + \frac{\partial w^s}{\partial \sigma} \left\{ - \int_{i^*}^2 \mu_{i^*}^s a^s \frac{\partial \bar{h}_i^s}{\partial w_{i^*}^s} g'_s di - \lambda \int_{i^*}^2 \frac{\partial x}{\partial w_{i^*}^s} n_i g_s(i) di \right\} \\ & + \frac{\partial w^u}{\partial \sigma} \left\{ - \int_{i^*}^2 \mu_{i^*}^u a^u \frac{\partial \bar{h}_i^u}{\partial w_{i^*}^u} g'_u di - \lambda \int_1^{i^*} \frac{\partial x}{\partial w^u} n_i g_u di \right\} - \lambda Y(p_2^* + \sigma, H^s, H^u) \end{aligned}$$

Now we need to calculate the inside of the integral. Note that from the definition of  $h_i^s$  and  $h_i^u$ , we have

$$\frac{\partial h_i^s}{\partial w_{i^*}^s} = -h_i^s \frac{\partial h_i^s}{\partial R_{i^*}^s} \text{ and } \frac{\partial h_i^u}{\partial w_{i^*}^u} = -h_i^u \frac{\partial h_i^u}{\partial R_{i^*}^u}$$

This implies that

$$-\mu_{i^*}^s a^s \frac{\partial h_i^s}{\partial w_{i^*}^s} g'_s = \mu_{i^*}^s a^s h_i^s \frac{\partial h_i^s}{\partial R_{i^*}^s} g'_s \text{ and } -\mu_{i^*}^u a^u \frac{\partial h_i^u}{\partial w_{i^*}^u} g'_u = \mu_{i^*}^u a^u h_i^u \frac{\partial h_i^u}{\partial R_{i^*}^u} g'_u$$

By using the FOC of  $R_i^s$  and  $R_i^u$ ,

$$\begin{aligned}\mu_i^s a^s h_i^s \frac{\partial h_i^s}{\partial R_i^s} g'_s &= h_i^s g_s \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^s} \right\} \\ \mu_i^u a^u h_i^u \frac{\partial h_i^u}{\partial R_i^u} g'_u &= h_i^u g_u \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^u} \right\}\end{aligned}$$

Thus,  $\frac{dW}{d\sigma}$  is

$$\begin{aligned}\frac{dW}{d\sigma} &= \frac{\partial i^*}{\partial \sigma} \left\{ \mu_{i^*}^s a^s h_{i^*}^s \frac{g'_s}{g_s} - \mu_{i^*}^u a^u h_{i^*}^u \frac{g'_u}{g_u} \right\} \\ &+ \frac{\partial w^s}{\partial \sigma} \left\{ \int_{i^*}^2 h_i^s g_s \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^s} \right\} di - \lambda \int_{i^*}^2 \frac{\partial x}{\partial w_i^s} g_s n_i di \right\} \\ &+ \frac{\partial w^u}{\partial \sigma} \left\{ \int_{i^*}^2 h_i^u g_u \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R_i^u} \right\} di - \lambda \int_1^{i^*} \frac{\partial x}{\partial w_i^u} g_u n_i di \right\} \\ &- \lambda Y(p_2^* + \sigma, H^s, H^u)\end{aligned}$$

Next, we need to calculate  $\lambda \frac{\partial w^s}{\partial \sigma} \int_{i^*}^2 h_i^s g_s(i) n_i + \lambda \frac{\partial w^u}{\partial \sigma} \int_1^{i^*} h_i^u g_u(i) n_i di$ . From the argument of the previous section, we have

$$\lambda y_2 = \lambda \frac{\partial w^s}{\partial \sigma} \int_{i^*}^2 h_i^s g_s(i) n_i + \lambda \frac{\partial w^u}{\partial \sigma} \int_1^{i^*} h_i^u g_u(i) n_i di.$$

Third, we will show that  $h_i^s \frac{\partial x}{\partial R_i^s} = -\frac{\partial x}{\partial w_i^s}$  and  $h_i^u \frac{\partial x}{\partial R_i^u} = -\frac{\partial x}{\partial w_i^u}$ . From the definition of  $Z$ , we have

$$\begin{aligned}\frac{\partial Z}{\partial R_i^s} &= a^s / w_i^s \text{ and } \frac{\partial Z}{\partial w_i^s} = -a^s h_i^s (1/w_i^s) \text{ for } i \in (i^*, 2) \\ \frac{\partial Z}{\partial R_i^u} &= a^u / w_i^u \text{ and } \frac{\partial Z}{\partial w_i^u} = -a^u h_i^u (1/w_i^u) \text{ for } i \in (1, i^*)\end{aligned}$$

Thus, by using the definition of  $\frac{\partial x}{\partial R_i^s}$ ,  $\frac{\partial x}{\partial w_i^s}$ ,  $\frac{\partial x}{\partial R_i^u}$ ,  $\frac{\partial x}{\partial w_i^u}$ , we can check that  $h_i^s \frac{\partial x}{\partial R_i^s} = -\frac{\partial x}{\partial w_i^s}$  and  $h_i^u \frac{\partial x}{\partial R_i^u} = -\frac{\partial x}{\partial w_i^u}$ .

Therefore,  $dW/d\sigma$  is

$$\frac{dW}{d\sigma} \Big|_{\sigma=0} = \frac{\partial i^*}{\partial \sigma} \left\{ \mu_{i^*}^s a^s h_{i^*}^s \frac{g'_s}{g_s} - \mu_{i^*}^u a^u h_{i^*}^u \frac{g'_u}{g_u} \right\}$$

From the FOC of  $v_{i^*}^s$  and  $v_{i^*}^u$ , we have  $\mu_{i^*}^s = \mu_{i^*}^u$ . In addition,  $a^s h_{i^*}^s \frac{g'_s}{g_s}$  and  $a^u h_{i^*}^u \frac{g'_u}{g_u}$  are the right side slope of  $v_{i^*}^s$  and the left side slope  $v_{i^*}^u$  at  $i^*$ . From Lemma 1, the slope of  $v_{i^*}^s$  is steeper than the slope of  $v_{i^*}^u$  at  $i^*$ . Since  $\frac{\partial i^*}{\partial \sigma} > 0$ ,  $\frac{dW}{d\sigma} > 0$ .

## References

- [1] Atkinson, A. and Joseph E. Stiglitz *Lectures on Public Economics* , McGraw Hill, 1980.
- [2] Atkinson, A. and Stiglitz, Joseph E., “The Design of Tax Structure: Direct versus Indirect Taxation,” *Journal of Public Economics* 6, July-Aug. 1976, pp 55-75.
- [3] Berg, Claude *Topological Spaces*, New York, Macmillan, 1963.
- [4] Boadway, Robin and Katherine Cuff, “A Minimum Wage Can Be Welfare-Improving and Employment-Enhancing,” *European Economic Review* 45(3), March 2001, pages 553-76.
- [5] Cremer, Helmuth, Pierre Pestieau and Jean-Charles Rochet, 2001, “Direct versus Indirect Taxation: The Design of the Tax Structure Revisted,” *International Economic Review*, vol 42(3),pp781-99
- [6] Diamond, Peter and James Mirrlees, 1971, “Optimal Taxation and Public Production,” *American Economic Review*, 61, pp8-27 and pp 261-278.
- [7] DiNardo, John and Justin Tobias, 2001, “Nonparametric Density and Regression Estimation,” *Journal of Economic Perspectives* 15(4), pp 11-28.
- [8] Fudenberg, Drew and Jean Tirole, *Game Theory*, MIT Press, 1992.
- [9] Hargerger, Arnold C., 1962, “The Incidence of the Corporation Income Tax,” *Journal of Political Economy*, pp215-240
- [10] Mirrlees, James A.,1971, “An Exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, 38, pp 175-208.
- [11] Naito, Hisahiro, 1996, “Tariff As A Device to Relax the Incentive Problem of a Progressive Income Tax System,” *Research Seminar of International Economics Working Papers*, No 391, The Department of Economics and School of Public Policy,<http://www.spp.umich.edu/rsi>, The University of Michigan.

- [12] Naito, Hisahiro, 1999, "Re-examination of Uniform Commodity Taxes under A Non-linear Income Tax System and Its Implication for Production Efficiency," February 1(2),. *Journal of Public Economics*, pp65-88
- [13] Naito, Hisahiro, 2002, "Endogenous Human Capital Accumulation and Direct vs. Indirect Redistribution," working paper, Institute of Social Economic Research Osaka University
- [14] Stiglitz, Joseph, 1982, "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17, pp 213-240.
- [15] Samuelson, Paul A., 1949, "International Factor-Price Equalisation Once Again", *The Economic Journal*, pp 181-197.
- [16] Saez, Emmanuel, 2003, "Direct or Indirect Instruments for Redistribution: Short-run versus Long-Run", forthcoming from *Journal of Public Economics*
- [17] Saez, Emmanuel, 2002, "The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes," , *Journal of Public Economics*, vol83(2), pp217-30.
- [18] Samuelson, Paul A., 1949, "International Factor-Price Equalisation Once Again", *The Economic Journal*, pp 181-197.
- [19] Stolper, Wolfgang and Paul Samuelson, 1941, "Protection and Real Wages," *Review of Economic Studies*, 9, pp 58-73.
- [20] Tobias, L. Justin, 2003, "Are Returns to Schooling Concentrated Among the Most Able? A Semiparametric Analysis of the Ability Earnings Relationships", *Oxford Bulletin of Economics and Statistics*, vol65, pp 1-29.



Figure 1: The Graph of  $v(i)$

