Appendices

This appendix will not be included in the main text. It will be provided through the webpage of the Journal.

1 Proof of Proposition 1

The Lagrangian is:

\[ L = \int_1^\infty v^u(i)n_i di + \int_1^2 v^v(i)n_v di + \int_1^\infty \mu^\mu i \{ v^u - a^\mu h^\mu_i (g^\mu_{ui}/g_{ui}) \} di + \int_1^2 \mu^v i \{ v^v - a^v h^v_i (g^v_{vi}/g_{vi}) \} di + \beta_1 \{ v^v_p - v^v_i \} + \beta_2 \{ R^v_p - R^v_i \} + \lambda \int_1^\infty \{ R^u_p - x(R^u_i, v^u_i, w^u_i, w^u_i, p_1 + t) \} n_i di + \lambda \int_1^2 \{ R^v_p - x(R^v_i, v^v_i, w^v_i, w^v_i, p_1 + t) \} n_v di + \lambda t \int_1^\infty n_c t_1 di + \beta_3 \int_1^\infty g_{ui} h^u_i di - H^u \} + \beta_4 \int_1^\infty g_{vi} h^v_i di - H^v \}.

By using the integration by parts, we can obtain:

\[ L = \int_1^\infty v^u(i)n_i di + \int_1^2 v^v(i)n_v di + \mu^\mu_i v^u_i - \int_1^\infty \mu^\mu i v^u_i di + \mu^v i v^v_i - \int_1^2 \mu^v i v^v_i di + \int_1^\infty \mu^\mu_i a^\mu h^\mu_i (g^\mu_{ui}/g_{ui}) di + \beta_1 \{ v^v_p - v^v_i \} + \beta_2 \{ R^v_p - R^v_i \} + \lambda \int_1^\infty \{ R^u_p - x(R^u_i, v^u_i, w^u_i, w^u_i, p_1 + t) \} n_i di + \lambda \int_1^2 \{ R^v_p - x(R^v_i, v^v_i, w^v_i, w^v_i, p_1 + t) \} n_v di + \lambda t \int_1^\infty n_c t_1 di + \beta_3 \int_1^\infty g_{ui} h^u_i di - H^u \} + \beta_4 \int_1^\infty g_{vi} h^v_i di - H^v \}.

Denote \( x(R^u_i, v^u_i, w^u_i, w^u_i, p_1 + t) \) as \( x_i \). The first order condition for \( v^u_i, v^v_i, R^u_i, R^v_i, v^u_p, v^v_p, R^u_p, R^v_p, H^u \) and \( H^v \) are:

\[ v^u_i : n_i - \mu^u_i - \lambda_n i \frac{\partial x_i}{\partial v^u_i} + \lambda m_i \frac{\partial c_{ui}}{\partial x_i} \frac{\partial x_i}{\partial v^u_i} = 0 \]

\[ v^v_i : \mu^v_i = 0 \]

\[ v^v_p : -\mu^v_p + \beta_1 = 0 \]

\[ R^u_i : [ -\mu^u_i \times a^u i \frac{g^u_{ui}}{g_{ui}} + \beta_3 i g_{ui} ] \frac{\partial h^u_i}{\partial R^u_i} + \lambda_n i - \lambda_n i \frac{\partial x_i}{\partial R^u_i} + \lambda m_i \frac{\partial c_{ui}}{\partial x_i} \frac{\partial x_i}{\partial R^u_i} = 0 \]

\[ R^v_i : \beta_2 = 0 \]

\[ v^u_p : n_i - \mu^u_p - \lambda_n i \frac{\partial x_i}{\partial v^u_p} + \lambda m_i \frac{\partial c_{ui}}{\partial x_i} \frac{\partial x_i}{\partial v^u_p} = 0 \]

\[ v^v_p : \mu^v_p - \beta_1 = 0 \]
\[ R_{\mu}^i : [\mathbf{-}\mu^u_t \times d^{\mu}_a g_{\mu} + \beta_a g_{\mu}] \frac{\partial h^u}{\partial R_{\mu}^i} + \lambda n_i - \lambda n_i \frac{\partial x_i}{\partial R_{\mu}^i} + \lambda \alpha_1 n_i \frac{\partial x_i}{\partial R_{\mu}^i} = 0 \]

\[ R_{\mu}^i : \beta_2 = 0 \]

\[ H^i : \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial H^i} = \beta_s \]

\[ H^u : \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial H^u} = \beta_u \]

Now we characterize those first order conditions. First, note that:

\[ \mu^i_t = \mu^s_t + \int_1^{i'} n_j(1 - \frac{\lambda}{\partial v_j})d j \quad \text{and} \quad \mu^u_t = \mu^1_t + \int_1^{i'} n_j(1 - \lambda \frac{\partial x_j}{\partial v_j})d j \quad (1) \]

Since \( \mu^s_t = \mu^\ast_t, \mu^i_t = \int_1^{i'} n_j(1 - \lambda \frac{\partial x_j}{\partial v_j})d j \) for \( i \in (i^*, 2) \) and \( \mu^u_t = \int_1^{i'} n_j(1 - \lambda \frac{\partial x_j}{\partial v_j})d j \) for \( i \in (1, i^*) \). Note that \( \frac{\partial x_j}{\partial v_j} = 1/(u_x) \). A single crossing property and the monotonicity of \( R_{\mu}^i \) and \( R_{\mu}^u \) guarantee that \( x_i \) is increasing. This implies that \( \frac{\partial x_j}{\partial v_j} \) is increasing and the inside of the integral is a decreasing function of \( i \). Since \( \mu^s_t = 0 \) and \( \mu^u_t = 0 \), the only way that these conditions are satisfied is that initially \( n_i(1 - \lambda \frac{\partial x_j}{\partial v_j}) \) is positive and after some \( i^* \), it becomes negative. In this case, for all \( i \in [i^*, 2) \) and \( i_u \in (1, i^*) \), \( \mu^s_t \) and \( \mu^u_t \) are strictly positive.

Now we examine \( dW/\partial t \) and evaluate at \( t = 0 \). From the envelope theorem,

\[ \frac{dW}{\partial t} = \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial t} + \lambda \int_1^{i'} n_i c_1 d i - \lambda \int_1^{i'} \frac{\partial x_i}{\partial (p_1 + t)} n_i d i - \lambda \int_1^{i'} \frac{\partial x_i}{\partial (p_1 + t)} n_i d i \]

where

\[ \frac{\partial L}{\partial p_1} = d_{\mu} \left\{ v^s_t n_t - v^s_t n_r + \mu^s_t v^s_t + \mu^u_t v^u_t - \mu^u_r v^u_t - \mu^u_t a^u h^u_t g^u_{\mu} \right\} \]

\[-\mu^s_t v^s_t - \mu^u_t v^s_t + \mu^u_t v^u_t i^s_t + \mu^u_t a^u h^u_t g^u_{\mu} + \beta_1 \{ v^s_t - v^u_t \} + \beta_2 R_{u^s} - \beta_2 R_{u^u} - \lambda \{ R_{\mu}^s - x_r \} n_r + \lambda \{ R_{\mu}^s - x_r \} n_r \]

\[-\beta_g g_{\mu} h^u_t + \beta_0 g_{\mu} h^u_t \}

\[ + \left\{ \int_1^{i'} [\mu^s_t a^t (g_{\mu}/g_{\mu}) + \beta_0 g_{\mu}] \frac{\partial h^u_t}{\partial p_1} \frac{\partial w^s_t}{\partial p_1} d i - \lambda \int_1^{i'} \frac{\partial x_i}{\partial w^s_t} \frac{\partial x_i}{\partial p_1} n_i d i \right\} \]

\[ + \left\{ \int_1^{i'} [\mu^u_t a^u (g_{\mu}/g_{\mu}) + \beta_0 g_{\mu}] \frac{\partial h^u_t}{\partial p_1} \frac{\partial w^u_t}{\partial p_1} d i - \lambda \int_1^{i'} \frac{\partial x_i}{\partial w^u_t} \frac{\partial x_i}{\partial p_1} n_i d i \right\} \]

\[ - \lambda \int_1^{i'} \frac{\partial x_i}{\partial (p_1 + t)} n_i d i - \lambda \int_1^{i'} \frac{\partial x_i}{\partial (p_1 + t)} n_i d i \]
By using the above first order conditions, we have:

\[
\frac{\partial L}{\partial p_1} = \frac{di^*}{dp_1} \left\{ \mu^1_t a^w h^t_1 \frac{g_1^t}{g_u} - \mu^u_t a^w h^u_t \frac{g_u^t}{g_u} - \beta_s g_{su} h^s_r + \beta_s g_{su} h^s_r \right\} \\
+ \left\{ \int_{r^*}^{r^*} \left[ -\mu^1_t a^s (g_{si}^t / g_u) + \beta_s g_{su} \right] \frac{\partial h^t_r}{\partial p_1} \frac{\partial w^s_i}{\partial p_1} di - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial w^s_i} \frac{\partial w^s_i}{\partial p_1} n_i di \right\} \\
+ \left\{ \int_{r^*}^{r^*} \left[ -\mu^u_t a^s (g_{ui}^t / g_u) + \beta_s g_{su} \right] \frac{\partial h^u_r}{\partial p_1} \frac{\partial w^u_i}{\partial p_1} - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial w^u_i} \frac{\partial w^u_i}{\partial p_1} n_i di \right\} \\
- \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial (p_1 + t)} n_i di - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial (p_1 + t)} n_i di
\]

Now we need to calculate the inside of the integral. Note that from the definition of \( h^1 \) and \( h^u \), we have:

\[
\frac{\partial h^1_r}{\partial w^s_i} = -h^1_r \frac{\partial h^1_r}{\partial R^1_r} \quad \text{and} \quad \frac{\partial h^u_r}{\partial w^u_i} = -h^u_r \frac{\partial h^u_r}{\partial R^u_r}
\]

This implies that:

\[
[-\mu^1_t a^s (g_{si}^t / g_u) + \beta_s g_{su}] \frac{\partial h^1_r}{\partial w^s_i} = \left[ \mu^1_t a^s (g_{si}^t / g_u) - \beta_s g_{su} \right] h^1_r \frac{\partial h^1_r}{\partial R^1_r}
\]

and

\[
[-\mu^u_t a^s (g_{ui}^t / g_u) + \beta_s g_{su}] \frac{\partial h^u_r}{\partial w^u_i} = \left[ \mu^u_t a^s (g_{ui}^t / g_u) - \beta_s g_{su} \right] h^u_r \frac{\partial h^u_r}{\partial R^u_r}
\]

By using the FOC of \( R^1_r \) and \( R^u_r \),

\[
\left[ \mu^1_t a^s (g_{si}^t / g_u) - \beta_s g_{su} \right] h^1_r \frac{\partial h^1_r}{\partial R^1_r} = h^1_r \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R^1_r} \right\}
\]

\[
\left[ \mu^u_t a^s (g_{ui}^t / g_u) - \beta_s g_{su} \right] h^u_r \frac{\partial h^u_r}{\partial R^u_r} = h^u_r \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R^u_r} \right\}
\]

Thus, \( \frac{\partial L}{\partial p_1} \) is:

\[
\frac{\partial L}{\partial p_1} = \frac{di^*}{dt} \left\{ \mu^1_t h^t_1 \frac{g_1^t}{g_u} - \mu^u_t h^u_t \frac{g_u^t}{g_u} - \beta_s g_{su} h^s_r + \beta_s g_{su} h^s_r \right\} \\
+ \left\{ \int_{r^*}^{r^*} h^t_r \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R^1_r} \right\} \frac{\partial w^s_i}{\partial p_1} di - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial w^s_i} \frac{\partial w^s_i}{\partial p_1} n_i di \right\} \\
+ \left\{ \int_{r^*}^{r^*} h^u_r \left\{ \lambda n_i - \lambda n_i \frac{\partial x}{\partial R^u_r} \right\} \frac{\partial w^u_i}{\partial p_1} di - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial w^u_i} \frac{\partial w^u_i}{\partial p_1} n_i di \right\} \\
- \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial (p_1 + t)} n_i di - \lambda \int_{r^*}^{r^*} \frac{\partial x_i}{\partial (p_1 + t)} n_i di
\]

Next, we need to calculate \( \lambda \int_{r^*}^{r^*} h^t_r n_i \frac{\partial w^s_i}{\partial p_1} di + \lambda \int_{r^*}^{r^*} h^u_r n_i \frac{\partial w^u_i}{\partial p_1} di \). Note that \( \lambda \int_{r^*}^{r^*} h^t_r n_i \frac{\partial w^s_i}{\partial p_1} + \lambda \int_{r^*}^{r^*} h^u_r n_i \frac{\partial w^u_i}{\partial p_1} di = \lambda \int_{r^*}^{r^*} h^t_r g_{su} \frac{\partial w^s_i}{\partial p_1} + \lambda \int_{r^*}^{r^*} h^u_r g_{su} \frac{\partial w^u_i}{\partial p_1} di \).
is a change in total earnings due to a change in the price of good 1 when levels of human capital of all individuals are fixed. On the other hand, in perfect competition, for given levels of human capital of all individuals, the total revenue of a firm should be equal to the total payments to factor owners. Thus, \( p_1y_1 + y_2 = w_1^e \int_{s_1}^{2} n_i h_i g_{su} di + w_1^e \int_{s_1}^{1} n_i h_i g_{su} di \)

always holds true. Let \( Q(p_1) \) be the total revenue of firms when all human capital levels of all individuals are fixed. Then, \( \frac{dQ}{dp_1} = \lambda \int_{s_1}^{2} h_i g_{su} n_i \frac{\partial w^e}{\partial p_1} + \lambda \int_{s_1}^{1} h_i g_{su} n_i \frac{\partial w^e}{\partial p_1} di \).

By definition of \( Q(p_1) \)

\[
Q(p_1) \equiv \max_{p_1 y_1 + y_2} \quad \text{s.t.} \quad (y_1, y_2) \in \Gamma(H^e, H^u) \\
H^e \text{ and } H^u \text{ are fixed.}
\]

From the envelope theorem, \( \frac{dQ}{dp_1} = y_1 \). Therefore, we have:

\[
\lambda y_1 = \lambda \frac{\partial w^e}{\partial \sigma} \int_{s_1}^{2} h_i g_{su} n_i + \lambda \frac{\partial w^e}{\partial \sigma} \int_{s_1}^{1} h_i g_{su} n_i di.
\]

Third, we will show that \( h_i \int_{s_1}^{2} \frac{\partial s_i}{\partial H^e} = -\frac{\partial s_i}{\partial \sigma} \) and \( h_i \int_{s_1}^{1} \frac{\partial s_i}{\partial H^u} = -\frac{\partial s_i}{\partial \sigma} \). From the definition of \( Z \), we have:

\[
\frac{\partial Z}{\partial R^e_i} = a_i^t / w_i^e \text{ and } \frac{\partial Z}{\partial w_i^e} = -h_i^1 (1/w_i^e) \text{ for } i \in (i^*, 2) \\
\frac{\partial Z}{\partial R^u_i} = a_i^t / w_i^u \text{ and } \frac{\partial Z}{\partial w_i^u} = -h_i^2 (1/w_i^u) \text{ for } i \in (1, i^*)
\]

Thus, by using the definition of \( \frac{\partial s_i}{\partial H^e}, \frac{\partial s_i}{\partial H^u}, \frac{\partial s_i}{\partial \sigma}, \frac{\partial s_i}{\partial \sigma} \), we can confirm that \( h_i \int_{s_1}^{2} \frac{\partial s_i}{\partial H^e} = -\frac{\partial s_i}{\partial \sigma} \)

and \( h_i \int_{s_1}^{1} \frac{\partial s_i}{\partial H^u} = -\frac{\partial s_i}{\partial \sigma} \).

Therefore, \( \partial L / \partial p_1 \) is:

\[
\left. \frac{\partial L}{\partial p_1} \right|_{\sigma=0} = \frac{d_1^t}{dp_1} \left\{ \mu_i^* h_i^e g_i^e \frac{g_i^e}{g_s} - \mu_i^* h_i^u g_i^u \frac{g_i^u}{g_s} - \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial H^e} g_{su} h_i^e \right\} + \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial H^e} g_{su} h_i^u
\]

From the FOC of \( H^e \) and \( H^u \), we have:

\[
\frac{\partial L}{\partial p_1} \left\{ 1 + \frac{d_1^t}{dp_1} \frac{\partial p_1}{\partial H^e} g_{su} h_i^e \right\} - \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial H^u} g_{su} h_i^u \right\} = \frac{d_1^t}{dp_1} \left\{ \mu_i^* a_i^t h_i^e g_i^e \frac{g_i^e}{g_s} - \mu_i^* a_i^t h_i^u g_i^u \frac{g_i^u}{g_s} \right\}
\]

Therefore, this implies that

\[
\left. \frac{\partial L}{\partial p_1} \right|_{\Delta^t} = \frac{d_1^t}{dp_1} \left\{ \mu_i^* a_i^t h_i^e g_i^e \frac{g_i^e}{g_s} - \mu_i^* a_i^t h_i^u g_i^u \frac{g_i^u}{g_s} \right\}
\]

where \( \Delta^t = 1 + \frac{d_1^t}{dp_1} \left( \frac{\partial p_1}{\partial H^e} g_{su} h_i^e - \frac{\partial p_1}{\partial H^u} g_{su} h_i^u \right) > 0 \). By using the definition of \( \partial p_1 / \partial t, \partial p_1 / \partial H^e \) and \( \partial p_1 / \partial H^u \), we have:

\[
\left. \frac{dW}{dt} \right|_{t=0} = \Psi_1 \frac{d_1^t}{dp_1} \left\{ \mu_i^* a_i^t h_i^e g_i^e \frac{g_i^e}{g_s} - \mu_i^* a_i^t h_i^u g_i^u \frac{g_i^u}{g_s} \right\}
\]
\[ \Psi_1 = \frac{-RD_p}{RD_p - RS_p + RS_p \frac{\partial^2}{\partial p^2} g_{si} h_{si}^c + RS_u \frac{\partial^2}{\partial p^2} g_{ui} h_{ui}^c}. \]

From the FOC of \( v_s^i \) and \( v_u^i \), we have \( \mu_s^i = \mu_u^i \). In addition, \( \alpha^s h_s^i \frac{\partial}{\partial s^i} \) and \( \alpha^u h_u^i \frac{\partial}{\partial u^i} \) are the right side slope of \( v_s^i \) and the left side slope \( v_u^i \) at \( i^* \). From Lemma 1, the slope of \( v_s^i \) is steeper than the slope of \( v_u^i \) at \( i^* \). Since \( \frac{dW}{dt} > 0 \), \( \frac{\partial^2}{\partial p^2} g_{si} h_{si}^c < 0 \).

### 2 A Case of Imperfect Substitutes

In the main text, I considered a case where two types of human capital accumulation are perfect substitutes in the disutility function. As a result, each person accumulates only one type of human capital. In reality, however, individuals might accumulate both types of human capital. It is important to check the robustness of the proposition when two types of human capital are imperfect substitutes.

In order to make two types of human capital accumulation process imperfect substitutes, we assume that the utility function of the type \( i \) agent has the following form:

\[ u(c_{1i}, c_{2i}) = f_s(h_s^i) - f_s(h_u^i). \]

Regarding \( u(c_{1i}, c_{2i}) \), I make the same assumption as in the previous section. As for \( f_j(h_j^i) \) (\( j = s, u \)), \( f_j(h_j^i) \) is strictly increasing and strictly convex. The labor supply is fixed. In addition, to simplify the analysis, it is assumed that \( f_s(h_s^i) \) and \( f_u(h_u^i) \) have the following functional forms:

\[ f_s(h_s^i) = (h_s^i)^{\gamma_s} \quad \text{and} \quad f_u(h_u^i) = (h_u^i)^{\gamma_u} \]

where \( \gamma_s \) and \( \gamma_u \) measure the curvature of the disutility functions of skilled and unskilled human capital accumulation respectively and are strictly greater than one. Given the amount of skilled human capital and unskilled human capital of individual \( i \), I assume that the earnings of individual \( i \) is determined in the same way as in the previous subsection. Agents who have greater ability have more comparative advantage in accumulating skilled human capital than unskilled human capital. More specifically, I assume that:

\[ \frac{g_{si}^c}{g_{si}} > \frac{g_{ui}^c}{g_{ui}} \frac{\gamma_u}{\gamma_s}. \]

(2) implies that when the disutility of accumulating human capital is not constant, the comparative advantage condition needs to be adjusted by the curvature of the disutility function.

\[ \frac{g_{si}^c}{g_{si}} > \frac{g_{ui}^c}{g_{ui}} \frac{\gamma_u}{\gamma_s}. \]

1 Our main results can hold in more general functional forms.

2 The economic interpretation of (2) is as follows. Consider a condition that type \( i \) and type \( i + \epsilon \) agents have the same degree of comparative advantage. Equation (2) says that if the marginal disutility of accumulating skilled human capital grows faster than the marginal disutility of accumulating unskilled human capital (\( \gamma_s > \gamma_u \)), then the increase of the return from skilled human capital, \( g_{si}^c / g_{si} \), can be lower than the increase of the return from unskilled human capital, \( g_{ui}^c / g_{ui} \), in order to have the same comparative advantage. This is because accumulating skilled human capital accompanies larger disutility from the first place.
The objective of the social planner is to design a schedule of social welfare. By using the same technique as in the previous section, we can calculate note that the indifference curve of his utility:

\[
Z = \{ h \mid f(h) = \text{constant} \}
\]

Let the minimized value of the above problem be \( R(h) \). Then, from the dual relationship, we will have:

\[
h^j_i(w^j_i, w^u_i; R, h^s_i) \equiv h^j_i(w^j_i, w^u_i; V) ; j=s,u.
\]

By taking derivatives from both sides, we will have the Slutsky equation for \( h^s_i \) and \( h^u_i \):

\[
\frac{\partial h^s_i}{\partial w^j_i} + \frac{\partial h^s_i}{\partial R} h^s_i = \frac{\partial h^s_i}{\partial w^u_i} \quad \text{and} \quad \frac{\partial h^s_i}{\partial w^u_i} + \frac{\partial h^s_i}{\partial R} h^s_i = \frac{\partial h^s_i}{\partial w^u_i} ; j=s,u.
\]

Note that the indifference curve of \( f_s(h^s_i) + f_u(h^u_i) \) is strictly concave. Therefore, \( \frac{\partial h^s_i}{\partial w^s_i} > 0, \frac{\partial h^s_i}{\partial w^u_i} < 0, \frac{\partial h^u_i}{\partial w^s_i} > 0 \) and \( \frac{\partial h^u_i}{\partial w^u_i} < 0 \). This relationship means that if an individual maximizes his earnings holding the total disutility constant, an increase in the net return from skilled human capital, will increase the supply of skilled human capital and an increase in the return of unskilled human capital will decrease the supply of skilled human capital.

Let \( X(R) \) be the government-designed after-tax income schedule. Then, at the second stage of the problem, given \( Z(w^s_i, w^u_i; R) \) and \( X(R) \), each individual \( i \) will maximize his utility:

\[
\max_{\{R\}} U(p + t, X(R)) - Z(w^s_i, w^u_i; R).
\]

The objective of the social planner is to design a schedule of \( X(R) \) to maximize the social welfare. By using the same technique as in the previous section, we can calculate \( dv/dv_i: dv/di = -\sum_{j=s,u} Z_{w^j_i} \times (dw^j_i/di) \). Let \( \alpha_i \) be the Lagrangian multiplier of the
required income constraint in the disutility minimization problem (3). From the FOC of the minimization problem for \( Z(w^s_i, w^u_i, R) \), we obtain:

\[
\frac{dv}{di} = \alpha_i R_i \left\{ \frac{g'_s}{g_s} \theta_{si} + \frac{g'_u}{g_u} \theta_{ui} \right\} \text{ where } \theta_{ji} = \frac{w^j_i h^j_i}{R_i} ; j = s, u.
\] (4)

Because of the assumption of an absolute advantage, \( \frac{dv}{di} > 0 \). (4) has a clear economic significance. 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 \( h^s_i \) and \( h^u_i \), we can rewrite (4) as follows:

\[
\frac{dv}{di} = \frac{g'_s}{g_s} f'_s(h^s_i) h^s_i + \frac{g'_u}{g_u} f'_u(h^u_i) h^u_i.
\] (5)

Given (5), as in the previous section, it is more useful to assume that the social planner controls \( v_i \) and \( R_i \) and \( x_i \) are defined by the following relationship:

\[
v(i) = U(p_1 + t, X_i) - Z(w^s_i, w^u_i, R_i).
\]

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

\[
W(t) = \max_{\{R_i, v_i\}} \int_1^2 v(i)n_i di
st. \frac{dv}{di} = \frac{g'_s}{g_s} f'_s(h^s_i) h^s_i + \frac{g'_u}{g_u} f'_u(h^u_i) h^u_i
\]

\[
\int_1^2 n_i [R_i - x_i] di + t \int_1^2 c_1 n_i di = 0
\]

\[
H^s = \int g_s h^s_i di, \quad H^u = \int g_u h^u_i di,
\]

where \( p_1 = p_1(t, H^s, H^u) \)

and \( t \) is given.

In the above programming problem, \( W(t) \) is the maximized social welfare for given \( t \). Also note that \( h^s_i \) and \( h^u_i \) are functions of \( (R_i, w^s_i, w^u_i) \) and that \( w^s_i \) and \( w^u_i \) are functions of \( p_1 \).

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

\[
\frac{dW}{dt} \bigg|_{t=0} = -\Psi_2 \left\{ \int_1^2 \mu_i [\gamma'(g^s_i/g_s) - \gamma'(g^u_i/g_u)] f'_s \frac{\partial h^s_i}{\partial p_1} d\mu_i + f'_s \frac{\partial h^s_i}{\partial p_1} d\mu_i \right\} (6)
\]

As for SCP, we can check it by examining \( \frac{\partial^2 Z}{\partial R^2} > 0 \). This is true as long as \( \frac{\partial h^j_i}{\partial R} > 0 \) for \( j = s, u \).
for a given level of the commodity tax, the equilibrium price can be determined from
\[ \frac{\partial P}{\partial t} = \Psi_1(RD_p) \]
and \( \mu_i \) is the Lagrangian multiplier of the incentive compatibility constraint. Because of the properties of the compensated supply function of \( h_i \) and \( \mu_i \), the Stolper-Samuelson theorem, \( \partial w_i^s/\partial p_1 > 0 \) and \( \partial w_i^u/\partial p_1 < 0 \). From the Rybczynski theorem, \( RS_{Hs} > 0 \) and \( RS_{Hu} < 0 \). From the assumption on comparative advantage, \( \gamma(g_{su}/g_{us}) > 0 \). As for the sign of the Lagrangian multiplier of the incentive compatibility constraint, the standard argument shows that \( \mu_i > 0 \) for all \( i \in (1, 2) \) (See Section 3 of this appendices). Thus, we obtain \( dW/\partial t > 0 \).

**Proposition 2**
Suppose that the social planner sets the income tax structure to maximize the social welfare function in an endogenous skill accumulation model at a zero commodity tax. Then, an introduction of a commodity tax on skilled-labor-intensive goods will increase the social welfare.

Equation (6) 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, (6) shows that if \( (g_{su}/g_{us}) = (g_{us}/g_{su}) \), \( dW/\partial t = 0 \). In other words, if there is no comparative advantage and if greater ability individuals are as good at accumulating skilled and unskilled human capital as lower ability individuals, then there is no gain in social welfare from changing the returns of skilled and unskilled human capital. Second, \( (\partial h_i^s/\partial w_i^s)(\partial w_i^s/\partial p_1) \) and \( (\partial h_i^u/\partial w_i^u)(\partial w_i^u/\partial p_1) \) measure how changes of returns from each type of human capital change the compensated supply of skilled human capital. Third, \( \Psi_2 \) measures how a change in the commodity tax \( t \) will change the relative price of good 1, taking the effect of changes in the supply of human capital into consideration. Also note that \( \gamma \times f''_s(h_i^s) = f''_u(h_i^u)h_i^u + f''_u(h_i^u) \) and that \( f''_s(h_i^s)h_i^u + f''_s(h_i^u) \) is related to a change in \( \bar{v} \).

In addition, note that \( \mu_i \) measures how social welfare increases when the incentive compatibility is relaxed. This implies that the term after integration measures how a compensated change in returns from skilled and unskilled human capital changes the slope of \( \bar{v} \) and increases the social welfare.

The intuition of the above proposition is as follows: In a situation where individuals with greater ability have comparative advantage in accumulating skilled human capital and individuals with lesser ability have comparative advantage in accumulating unskilled human capital, a decrease in the return from skilled human capital and an increase in the return from unskilled human capital will hurt individuals with greater

\[ RS(t + p_1) = RS(p_1, H^s(p_1), H^u(p_1)) \]
Second, \( f_1^s g_1^s(\partial h_i^s/\partial p_1) + \frac{\partial w_i^s}{\partial p_1} | \Psi_1 \) and \( f_1^u g_1^u(\partial h_i^u/\partial p_1) + \frac{\partial w_i^u}{\partial p_1} | \Psi_1 \) are the compensated change of supply of skilled and unskilled human capital when the price of good 1 increases. Thus, we obtain \( dp_1/\partial t = \Psi_2 \).
ability and benefit individuals with lesser ability. If the social planner is interested in redistributing income from high ability individuals to low ability individuals, such changes in the returns from skilled and unskilled capital can indirectly redistribute income. On the other hand, starting from zero distortion, the deadweight loss of the commodity tax is of the second-order but the social welfare gain of relaxing the incentive problem has a first-order effect. As a result, introducing production distortion increases social welfare.

3 Proof of the Proposition 2

Let $\mu_i$ and $\lambda$ be the Lagrangian multiplier of the incentive constraint and the resource constraint. Denote $x(R_i, v_i, w_i, p_i + t)$ as $x'$. Then, the Lagrangian function is:

$$W(t) = \int_1^2 v_i n_i di + \int_1^2 \mu_i \frac{dv_i}{dt} - f'_s(h'_i)(g'_s/g_s) - f'_u(h'_u)h'_u(g'_u/g_u)di + \lambda \int_1^2 n_i \{R_i - x_i\} di + \int_1^2 n_i c_1 di + \beta_i \{\int_1^2 g_u h'_u di - H'\} + \beta_h \{\int_1^2 g_u h''_u di - H''\}$$

By using the integration by parts, we can obtain:

$$W(t) = \int_1^2 v_i n_i di + \int_1^2 \mu_i \frac{dv_i}{dt} - \int_1^2 \mu_i f'_s(h'_i)(g'_s/g_s)di - \int_1^2 \mu_i f'_u(h'_u)h'_u(g'_u/g_u)di + \lambda \int_1^2 n_i \{R_i - x_i\} di + \int_1^2 n_i c_1 di + \beta_i \{\int_1^2 g_u h'_u di - H'\} + \beta_h \{\int_1^2 g_u h''_u di - H''\}$$

the first-order-conditions are:

$$v_i : n_i - \dot{u}_i - \lambda n_i \frac{\partial x_i}{\partial v_i} + \lambda n_i \frac{\partial c_1}{\partial x} \frac{\partial x_i}{\partial v_i} = 0$$

$$R_i : -\mu_i \frac{d[f'_u(h'_u)(g'_u/g_u)]}{dh'_u} \frac{\partial h'_u}{\partial R_i} - \mu_i \frac{d[f'_u(h'_u)(g'_u/g_u)]}{dh'_u} \frac{\partial h'_u}{\partial R_i} + \lambda n_i$$

$$\frac{\partial h'_u}{\partial R_i} + \beta_u g_u \frac{\partial h'_u}{\partial R_i} - \lambda n_i \frac{\partial x_i}{\partial x_i} + \lambda n_i \frac{\partial c_1}{\partial x} \frac{\partial x_i}{\partial v_i} = 0$$

$$\mu_1 = 0 and \mu_2 = 0$$

From the FOC of $v_i$, we will have $n_i - \dot{u}_i - \lambda n_i \frac{\partial x_i}{\partial v_i} + \lambda n_i \frac{\partial c_1}{\partial x} \frac{\partial x_i}{\partial v_i} = 0$

$$n_i - \lambda n_i \frac{\partial x_i}{\partial v_i} = \mu_i$$
at $t = 0$. By integrating both sides and using the definitions of $\frac{\partial x}{\partial w}$ and $\mu_1 = 0$, we will have

$$\int_t^i n_i \{1 - \frac{\lambda}{U_x}\} = \mu_i$$

From the first order condition of the revelation problem, $U_x(p_1, X) X'(i) = Z_R R'(i)$. 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, $\frac{\partial x}{\partial w}$ is increasing. This implies that if at some $i^*$, $1 - \frac{\lambda}{U_x} = 0$, then for any $i > i^*$, $1 - \frac{\lambda}{U_x} < 0$. 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 zero at $i = 2$. Therefore, $\mu_i > 0$ for all $1 < i < 2$.

Now, we calculate the effect of increasing the commodity tax from $t = 0$. By using the envelope theorem, we have:

$$\left. \frac{dW}{dt} \right|_{t=0} = \frac{\partial L}{\partial p_1} \frac{\partial p_1}{\partial t} + \lambda \int_1^2 c_1 n_i d\lambda + \int_1^2 \left[ - \frac{\partial x_i}{\partial (p_1 + t)} \right] d\lambda$$

where $\frac{\partial L}{\partial p_1}$ is:

$$\frac{\partial L}{\partial p_1} = \int_1^2 \left\{ - \mu_i \frac{d[f(h_i^*)h_i^*(g_{ai}/g_{ui})]}{dh_i^*} + \beta_{sgui} \right\} \left\{ \frac{dh_i^*}{dw_i^*} \frac{dw_i^*}{dp_1} + \frac{dh_i^*}{dw_i^*} \frac{dw_i^*}{dp_1} \right\} d\lambda + \lambda \int_1^2 \left[ - \frac{\partial x_i}{\partial (p_1 + t)} \right] d\lambda$$

Note that $\frac{\partial x_i}{\partial (p_1 + t)} = -(U_{p_1})/(U_x)$. From the Roy’s identity, $-(U_{p_1})/(U_x) = c_{1i}$. Therefore, $\lambda \int_1^2 c_1 n_i d\lambda = \lambda \int_1^2 \left\{ \frac{\partial x_i}{\partial p_1} \right\} n_i d\lambda$. In addition, $\frac{\partial x_i}{\partial w_i^*} = z_{wii}/U_{x_i}$ and $\frac{\partial x_i}{\partial w_i^*} = z_{wii}/U_{x_i}$. Using the definition of $Z_{ai}$ and $Z_{aui}$, $\frac{\partial x_i}{\partial w_i^*} = -\lambda h_i^*/U_x$, $\frac{\partial x_i}{\partial w_i^*} = -\lambda h_i^*/U_x$. Thus, $\int_1^2 c_1 n_i d\lambda = \int_1^2 \left[ - \frac{\partial x_i}{\partial (p_1 + t)} \right] n_i d\lambda$. Therefore, on the other hand, the FOC of $R_i$ at $t = 0$ is:

$$\{ - \mu_i \frac{d[f(h_i^*)h_i^*(g_{ai}/g_{ui})]}{dh_i^*} + \beta_{sgui} \} \frac{\partial h_i^*}{\partial R_i} + \{ - \mu_i \frac{d[f(h_i^*)h_i^*(g_{ai}/g_{ui})]}{dh_i^*} + \beta_{sgui} \} \frac{\partial h_i^*}{\partial R_i}$$

$$+ \lambda n_i = \lambda n_i \alpha_i/U_x$$

Now, we will calculate $\lambda \int_1^2 \left[ - \frac{\partial x_i}{\partial w_i^*} - \frac{\partial x_i}{\partial w_i^*} \frac{\partial w_i^*}{\partial p_1} \right] n_i d\lambda$. Note that $\lambda \frac{\partial x_i}{\partial w_i^*} n_i = \lambda \alpha_i h_i^*/U_x$ and $-\lambda \frac{\partial x_i}{\partial w_i^*} n_i = \lambda \alpha_i h_i^*/U_x$. From the FOC of $R_i$, $\lambda \int_1^2 \left[ - \frac{\partial x_i}{\partial w_i^*} - \frac{\partial x_i}{\partial w_i^*} \frac{\partial w_i^*}{\partial p_1} \right] n_i d\lambda$ is
equal to:

$$
\int_1^2 \left\{ \left( - \mu_i \frac{d[f_i'(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} + \beta_u g_i \right) \frac{\partial h_i^u}{\partial R_i} + \left( - \mu_i \frac{d[f'_i(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} + \beta_u g_i \right) \frac{\partial h_i^u}{\partial R_i} + \lambda_i \right\} \frac{\partial w_i^u}{\partial p_i^1} \, di 
+ \int_1^2 \left\{ \left( - \mu_i \frac{d[f_i'(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} + \beta_u g_i \right) \frac{\partial h_i^u}{\partial R_i} + \lambda_i \right\} \frac{\partial h_i^u}{\partial p_i^1} \, di 
$$

Therefore, \( \frac{\partial L}{\partial p_1} \) becomes

$$
\left. \frac{\partial L}{\partial p_1} \right|_{t=0} = \int_1^2 \left\{ \left( - \mu_i \frac{d[f_i'(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} + \beta_u g_i \right) \frac{\partial h_i^u}{\partial w_i^u} \, d\beta_i^1 + \int_1^2 \lambda_n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di + \int_1^2 \lambda_n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di \right\}
$$

Note that \( \int_1^2 \lambda_n h_i^u \frac{\partial w_i^u}{\partial \sigma} \, di \) from the argument in the previous sub-section. Therefore, \(+ \int_1^2 [ - \lambda \frac{\partial x_i}{\partial (p_1 + t)} + \int_1^2 \lambda n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di ] + \int_1^2 \lambda n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di = \lambda y^3 \) from the argument in the previous sub-section. Therefore, \(+ \int_1^2 [ - \lambda \frac{\partial x_i}{\partial (p_1 + t)} + \int_1^2 \lambda n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di ] + \int_1^2 \lambda n h_i^u \frac{\partial w_i^u}{\partial p_i} \, di = 0 \). We have:

$$
\frac{\partial L}{\partial p_1} = \int_1^2 \left\{ \left( - \mu_i \frac{d[f_i'(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} + \beta_u g_i \right) \frac{\partial h_i^u}{\partial w_i^u} \frac{dw_i^u}{dp_1} + \frac{\partial h_i^u}{\partial w_i^u} \frac{dw_i^u}{dp_1} \, di \right\}
$$

From the FOC of \( H^u \) and \( H^u \), \( \beta_u = \frac{\partial w}{\partial p_1} \frac{\partial p_1}{\partial p_1} \) and \( \beta_u = \frac{\partial w}{\partial p_1} \frac{\partial p_1}{\partial p_1} \). Thus,

$$
\frac{\partial L}{\partial p_1} = \int_1^2 \left\{ \left( - \mu_i \frac{d[f_i'(h_i^g) h_i^u(g'_i, g_i)]}{dh_i} \right) \frac{\partial h_i^u}{\partial w_i^u} \frac{dw_i^u}{dp_1} + \frac{\partial h_i^u}{\partial w_i^u} \frac{dw_i^u}{dp_1} \, di \right\}
$$
Solving $\frac{\partial L}{\partial p_1}$, we have:

$$\frac{\partial L}{\partial p_1} = \frac{1}{\Delta} \left\{ \int_1^2 \left( -\mu \frac{d[f'_u(h')h'_u(g'_u/g_u)]}{dh'_u} \right) \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] \right\}$$

$$\frac{1}{\Delta} \left\{ \int_1^2 \left( -\mu \frac{d[f'_u(h')h'_u(g'_u/g_u)]}{dh'_u} \right) \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] \right\}$$

where $\Delta = 1 - \left( \int_1^2 \frac{\partial L}{\partial p_1} g'_s \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] dt \right) > 0$.

Therefore, we have:

$$f'_u(\tilde{h}_1') \frac{\partial \tilde{h}'_1}{\partial w'_1} + f'_u(\tilde{h}'_u) \frac{\partial \tilde{h}'_u}{\partial w'_u} = 0$$

and $f'_u(\tilde{h}_1') \frac{\partial \tilde{h}'_1}{\partial w'_1} + f'_u(\tilde{h}'_u) \frac{\partial \tilde{h}'_u}{\partial w'_u} = 0$.

Using the definition of $\partial p_1 / \partial H^s$ and $\partial p_1 / \partial H^u$, this implies that $dW / dt$ is equal to

$$\frac{dW}{dt} = -\Psi_2 \left\{ \int_1^2 \mu [\gamma'(g'_s/g_s) - \gamma'(g'_u/g_u)] f'_u \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] dt \right\}$$

where

$$\Psi_2 = \frac{-RD_p}{RD_{p_1} - RS_{p_1} - RS_{H^s} \left( \frac{1}{\gamma'_s} \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] dt \right) - RS_{H^u} \left( \frac{1}{\gamma'_u} \left[ \frac{\partial \tilde{h}'_1}{\partial w'_1} dw'_1 + \frac{\partial \tilde{h}'_u}{\partial w'_1} dw'_u \right] dt \right)}$$

From the condition of the comparative advantage, $\gamma'(g'_s/g_s) - \gamma'(g'_u/g_u) > 0$. In addition, $\Psi_2 < 0$. Thus, $\frac{dW}{dt} > 0$. 

12