Redistributive Tax and Growth in a Model with Discrete Occupation choice

Debasis Bandyopadhyay* Parantap Basu†

*University of Auckland, d.bandyopadhyay@auckland.ac.nz
†parantap.basu@durham.ac.uk
This paper is posted at ResearchSpace@Auckland.
http://researchspace.auckland.ac.nz/ecwp/192
Redistributive Tax and Growth
in a Model with Discrete Occupational Choice

Debasis Bandyopadhyay*
Department of Economics
The University of Auckland
Private Bag 92019
Auckland, New Zealand
E-mail: debasis@auckland.ac.nz.

Parantap Basu
Department of Economics
Fordham University
Bronx, NY 10458. USA
E-mail: basu@murray.fordham.edu

August 1999

Abstract

An optimal redistributive tax-subsidy formula is derived for a growth model where inequality is endogenously driven by an adult's choice of occupation between work and management. The human capital or knowledge is the only engine of growth and there is externality associated with human capital à la Lucas (1988) and Romer (1990). How much available knowledge would be exploited in the economy depends on the proportion of innovators who are called managers in our model. A redistributive tax reform directly impacts this proportion of managers by influencing the occupational choice. A growth maximizing redistributive scheme involves an educational subsidy to the innovators financed by taxing workers. Such an optimal educational subsidy is, however, path dependent. An economy with excessively high proportion of managers warrants lesser educational subsidy.

* This paper was written while the second author was visiting the Department of Economics, the University of Auckland in December 1998 and, between July and August 1999. Basu's first visit was funded by the Marsden Grant # 96-UOA-SOC-0018 of the Royal Society of New Zealand and the second visit was funded by the Economics Department of the University of Auckland. We gratefully acknowledge the support and secretarial assistance from the Economics Department and the Royal Society of New Zealand.
1. Introduction

Whether redistributive policies hurt or promote growth is an important issue. The conventional neo-classical wisdom suggests that redistributing income from rich to poor may hurt growth because the rich undertake all the investment and innovations and thus a redistributive taxation adversely affects their incentives. A number of papers including Chamley (1986) and Judd (1985) argue that the optimal long run capital income tax is zero. Judd (1999) finds zero capital income tax the optimal policy even when human capital is included in the model. Wright (1996) makes an argument that in most empirically relevant cases, faster growing economies will have lower taxes on capital income.

However, because of the assumption of homogenous agents and complete markets in all these models, the steady state is path independent in the sense that it does not depend on the initial distribution of wealth. As a result, the optimal tax policy is also independent of the initial distribution of wealth. A wave of recent papers including Galor and Zeira (1993), and Banerjee and Newman (1994) point out the path dependence property of the steady state. If the poor inherit less human capital than the rich and have no access to a credit market, they will continue to remain poor and would not invest in human capital unlike the rich. Although Galor and Zeira (1993) provide important insight into the relationship between credit market imperfection and income distribution, it does not address any fiscal policy issues and moreover their model does not allow steady state growth. ¹

Although the issue of income distribution is becoming increasingly relevant in macroeconomics, hardly any paper addresses the issue of optimal redistributive tax policy in an environment where growth and inequality are endogenous. Freeman (1996)

¹ A number of papers have, however, examined the issue whether zero capital income tax is socially optimal when credit market is not perfect in the sense that there is borrowing constraint. Hubbard and Judd (1986) show that positive capital income tax is socially optimal if agents face binding credit constraint. Basu (1987) also find that if workers have no access to the credit market and a benevolent government maximizes worker's welfare, optimal capital income tax may not be zero. Benabou (1996) addresses the issue of redistributive taxation in an environment with imperfect credit market and finds cases where redistribution may positively influence growth.
develops an overlapping generations model with discrete occupational choice and makes the point that income inequality may arise endogenously and perpetuate even when the agents have identical tastes and initial endowments. However, Freeman's model abstracts from any consideration of growth as well as redistributive taxation. In this paper, we address the issue of redistributive capital income taxation in an environment where there is growth with discrete occupational choice. The model is an extension of Bandyopadhyay (1993). The human capital is the only engine of growth in this model. There is externality associated with human capital à la Lucas (1988). However, unlike Lucas (1988) and Romer (1990), the total factor productivity in our model depends on the intensity of innovation activity represented by the proportion of innovators (who we call managers) in the economy. The fraction of people choosing managerial occupation is the driving force behind innovation in our model. 2

A redistributive taxation impacts the long-run growth in our model through two distinct channels. First, it directly impacts the steady state rate of investment by driving a wedge between before and after tax return on capital. This is the usual distortionary effect of capital income taxation characterized in several papers including Judd (1999). Second, it indirectly influences growth by impacting the most fundamental state variable in our model, which is the occupational distribution. The extant growth models as well as the optimal tax literature accompanying it do not incorporate this latter distributional effect of capital income tax.

Why is the distributional effect of capital income taxation relevant? It is well known that growth is significantly determined by the total factor productivity. Recent literature on endogenous growth models following Lucas (1988) and Romer (1990) stresses the importance of non-rival knowledge in determining the total factor productivity. In a recent paper, Prescott (1998) argues that it is the extent of non-rival knowledge that a

---

2 The proportion of manager proxies the proportion of educated people in the economy because in our model, in the steady state only managers undertake investment in education. The motivation for including the proportion of managers in the total factor productivity function stems from the study of Bandyopadhyay (1997) who finds that the proportion of educated people significantly explain the cross country disparity in growth rates.
country exploits rather than the available stock of knowledge itself that accounts for the cross country disparity in income. How much available knowledge would be exploited in the economy depends on the intensity of innovations, which in turn could be directly related to the proportion of innovators or managers in the economy. A change in capital income taxation by altering the steady state proportion of these managers may influence growth. This is how growth and distribution interact in our model when the fiscal authority undertakes redistributive tax reforms.

We show that any fiscal redistributive program that attempts to equalize income among workers and managers is socially harmful because it hurts long run growth. The optimal tax policy thus turns out to be negative capital income taxation, which is equivalent to a proportional educational subsidy in our model. In the steady state, it is optimal for the fiscal authority to subsidize managers by taxing workers. We find that the negative capital income taxation holds regardless of the initial state of distribution of skills. On the other hand, the extent of educational subsidy is, however, path dependent. An economy with higher proportion of managers beyond a threshold level warrants lesser educational subsidy.

In order to understand the rationale behind the negative capital income taxation in our model, one has to see how the proportion of managers influences the growth via its effect on the rate of return on investment. In our economy the rate of return on investment in education depends on the proportion of managers in two different ways. First, a higher proportion of managers enhances the total factor productivity through externality. Second, a higher proportion of managers also accentuates the relative scarcity of workers in the economy and thus lowers the rate of return on education by reducing the relative income of the managers. Thus there exists a socially optimal proportion of managers at which these two opposing effects just balance each other. If an economy starts off with a proportion of managers above this socially optimal level, one may bring a case for a positive redistributive capital income taxation because it may raise growth by eliminating

---

3 In a similar vein Galor and Tsiddon (1997) highlight the importance of high-ability individuals in determining economic growth.
some surplus managers. However, we prove that the distortionary effect of such a positive redistributive capital income taxation always outweighs the growth enhancing effect thus making a positive capital income taxation a sub-optimal tax policy.

In the context of fiscal redistribution, our model parallels the results of Rogerson and Ferenandez (1996) and Freeman (1996). Both these papers indicate that a transfer of resources from lower to higher-income individuals is an equilibrium outcome. In Rogerson and Fernandez (1996), this transfer of resource takes place by a redistributive tax rate, which is chosen by majority vote while in our setting it is determined by utility maximization. Freeman (1996), uses lottery as a device to finance public education. Both these papers abstract from long-run growth considerations while our emphasis here is to understand the interaction between a redistributive tax policy, occupational choice and long-run growth.

The rest of the paper is organized as follows. In the following section, we lay out the model and its comparative statics. Section 3 derives the steady state properties of the model. Section 4 deals with the issue of optimal redistributive tax policy. Section 5 ends with concluding comments.

2. The Model

Consider an environment with variable human capital, labor, and a single perishable consumption good. An agent lives two periods, one as a child being attached to an adult and one as an adult when she receives a child of her own. There is a continuum of dynasties with measure one and at each date $t$, a typical dynasty consists of an adult and a child. The adult has one unit of labor and $h$ units of human capital. She earns her income by choosing between the occupations of manager and worker and then divides her income between current consumption and investment in her child’s education. Investment in human capital is the only means of transferring consumption in our model. We assume incomplete markets for human capital because human capital cannot be used as a
collateral for loans and there is no separate tangible capital in the economy. This rules out a viable credit market in the model.

Preferences display intergenerational altruism, and so the adult maximizes the present discounted value of consumption of her dynasty. Dynasties differ only in terms of the adult's endowment of human capital at date 0. At date t, \( \Psi_t \) denotes the cumulative distribution of human capital among the date \( t \) adults. The history specifies the initial distribution \( \Psi_0 \). For simplicity consider a two-point distribution for \( \Psi_0 \), where \( h \) can take two possible values, 0 and \( h_0 \), which means

\[
(1) \quad \Psi_0(h) = 1 - m
\]

which means that initially the skill distribution is such that there are \( m \) proportion of managers and \( 1 - m \) proportion of workers.

Production is done by groups, each of which consisting of a manager and one or more workers. The output \( q \) of a group depends on the manager's human capital \( h \), the number \( n^d \) of workers she employs and the total factor productivity level \( A > 0 \) such that

\[
q = Ah^{1-a} (n^d)^a, \text{ where } 0 < a < 1 \text{ measures the output elasticity of worker. Assume }
\]

that the total factor productivity \( A \) depends, following Lucas (1988), on a knowledge spillover process that increases with the quality of labor measured by the economy's average human capital stock \( H \) and, following Romer (1990), on the stock \( A_0 \) of non-rival knowledge.

In contrast with Lucas (1988) and Romer (1990) in our model the total factor productivity \( A \) also depends on the intensity of the innovative activity proxied by the proportion of managers, \( m \). Two countries may have the same average human capital but experience different growth pattern because of different \( m \). The knowledge spillover increases with the intensity of innovative activities in the economy measured by the proportion \( m \) of adults who are managers and the institutional barriers to spillover of

---

\(^4\) The incomplete market for human capital is crucial in our model for preserving the dynastic
knowledge. In particular, assume that \( A = A_0 m^\theta H^b \), where \( b > 0 \) is a parameter measuring the degree of externality and \( \theta \geq 0 \) is a parameter measuring the degree of institutional barriers to the spillover of knowledge.

A few additional clarifications about the specification of total factor productivity are in order. The total factor productivity \( A \) depends on the aggregate level of human capital or knowledge, \( H \) and the extent of exploited knowledge which is partly determined endogenously by the proportion of innovators, \( m \) and partly exogenously by the institutional barriers to the dissemination of information. The latter institutional barrier is parameterized by \( \theta \). If \( \theta = 0 \), there is no such barrier.

To summarize, at each date \( t \geq 0 \) the output \( q \) of a manager is given by

\[
q(h, n^d_t; H_{mt}, m_t) = A_0 m^\theta H^b h^{1-a} (n^d_t)^a, \quad t = 0, 1, 2, \ldots
\]

Since our central interest is in understanding the interaction between tax policy and long-run growth, we assume that \( b = a \). This assumption makes the aggregate production function linear in the reproducible input \( H \) implying endogenous growth.\(^6\) Observe that there are \( m \) firms per capita in this economy, each being run by one manager. On average, there are \( (1-m)/m \) workers per firm. Note that \( \theta > 0 \) implies that the total factor productivity in the economy decreases as number of workers per manager increases. At any given point in time, manager’s human capital is given. An increase in the number of workers per managers gives rise to overcrowding and results in diseconomies of scale.

At each date \( t \geq 0 \) given the wage rate \( w_t \), and the two external factors \( H_{mt} \) and \( m_t \), a manager with \( h \) units of human capital employs \( n^d_t \) number of workers so as to

\(^5\) These institutional barriers may be of the form of patent laws which prevents instant diffusion of new technology or knowledge from one firm to another.

\(^6\) To see this note that in a steady state equilibrium the distribution of human capital coincides with the initial distribution laid out in (1). This means the workers stay as workers and managers stay as managers. In a steady state equilibrium described in section 4, workers do not invest in human capital meaning \( H_t = H_{mt} \), which means the production function becomes linear in \( H \) making the production function "Ak" type à la Rebelo (1991). The details of the steady state property of the model are discussed in section 3.
(3) \[ \text{Maximize } q(H_t, m_t, n_t^d, h) w_t n_t^d \]
\[ n_t^d > 0 \quad t=0, 1, 2, \ldots \]

The first order condition of (3) yields \( w_t = aA_0 H_t^\alpha m_t^\theta h^{1-a}(n_t^d)^{a-1} \), or, equivalently, the optimal number \( n_t^d(h) \) of workers employed by a manager with \( h \) units of human capital is

(4) \[ n_t^d(h) = \left( \frac{am_t^\theta H_t^\alpha}{w_t} \right)^{\frac{1}{1-a}} h, \quad t=0, 1, 2, \ldots \]

By (3) and (4) at each date \( t \) the indirect profit of a manager is proportional to her human capital stock \( h \) and is given by \( r_t h \), where,

(5) \[ r_t = (1-a)A_0 m_t^\theta H_t^\alpha \left( aA_0 m_t^\theta H_t^\alpha / w_t \right)^a h^{1-a}, \quad t=0, 1, 2, \ldots \]

**The Government**

There is a benevolent government in this economy undertaking a redistributive program. At each date, the government taxes managers’ incomes and redistributes all the tax proceeds to the workers in the form of a wage subsidy. Let \( T_t \) stand for the tax revenue at date \( t \) generated from taxing managers’ profit. Since the manager’s profit is nothing but returns to human capital, the tax revenue \( T_t \) can be written as:

(6) \[ T_t = \tau r_t H_{mt} \]

where \( \tau \) is a constant redistributive tax rate on the implicit rental income of the managers. Let \( Z_t \) stand for the total subsidy at date \( t \) given to workers and \( z_t \) represent per worker subsidy. Then \( Z_t = (1-m_t)z_t \). Since the government returns all tax revenue to workers as subsidy, the government budget constraint is:
(7) \[ Z_t = T_t \]

which means that the per worker subsidy, \( z_t \) is:

(8) \[ z_t = \frac{\tau \tau_t H_m t}{(1 - m_t)} \]

Note that we do not impose any restriction on the sign of the redistributive tax rate \( \tau \) meaning \( \tau \) may as well be negative in which case we have a transfer from workers to managers. The sign of \( \tau \) is determined by equilibrium consideration, which is the central theme of our paper.\(^7\)

The Breakeven Skill Level

At each date \( t \), \( x_t \) denotes the level of \textit{breakeven skill} such that an adult with \( x_t \) units of human capital earns an equal amount either as a manager or as a worker. By (5), \( x_t \) satisfies

(9) \[ \bar{w}_t = \bar{r}_t x_t, \quad t = 0, 1, 2, \ldots \]

where \( \bar{w}_t = w_t + z_t \), is the post subsidy wage and \( \bar{r}_t = (1 - \tau) r_t \), is the after tax rate of return on human capital. The adult’s occupation \( n_t(.) \) is an indicator function such that if she is a worker, \( n_t(.) = 1 \), otherwise, if she is a manager, \( n_t(.) = 0 \). At each date \( t \geq 0 \), her occupational choice \( n_t(.) \) and the resulting income \( y_t(.) \) as functions \( h \geq 0 \) are

(10) \[ n_t(h, \tau) = 1, \text{ if } h < x_t; \quad n_t(h, \tau) = 0 \text{ if } h > x_t; \]

\[ n_t(h, \tau) = 1 \text{ or } 0, \text{ if } h = x_t, \]

(11) \[ y_t(h, \tau) = n_t(h, \tau) \cdot w_t + (1 - n_t(h, \tau)) \cdot r_t h. \]

Figure 1 illustrates how the basic skill level divides the adults into two occupational groups, workers and managers, according to their individual stock of human capital.
An adult’s human capital $h_{t+1}$ at a date $t+1$ is positively related to her parent’s human capital $h_t$ and the investment $s_t$ in her schooling made by her parent at date $t$. In particular,

$$ h_{t+1} = (1 - \delta)h_t + s_t, \quad 0 < \delta < 1 $$

The above formulation presumes an externality $\delta < 1$ associated with family upbringing in the tradition of Benabou (1996). It also assumes that without investment in schooling the current generation can transfer only a fraction $\delta > 0$ of existing knowledge to the future generation. Consequently, knowledge is maintained or accumulated only if a generation acquires them through parental investment in schooling. This feature is similar to Mankiw, et al. (1992).

Following Barro (1974) we assume intergenerational altruism. At each date $t$, the utility $v_t$ of the adult is a function of her family’s consumption $c_t$ and her child’s utility $v_{t+1}$ as a grown-up adult. In other words,

$$ v_t = V(c_t, v_{t+1}) = u(c_t) + \beta v_{t+1}, \quad t=0, 1, 2, ... $$

We assume that $u(.)$ is strictly concave, bounded above, $u(0) = -\infty, u'(0) = \infty, 0 < \beta < 1$, such that $v_0 = \sum_{t=0}^{\infty} \beta^t u(c_t)$.

The adult with $h$ units of human capital chooses a suitable occupation $n(h)$ following (10) and divides her income $y(h)$, given by (11), between consumption $c_t$ and investment $s_t$ such that

\footnote{We show in section 4 that the optimal $\tau$ is negative in a steady state.}
(14) \[ c_t + s_t \leq y_t(h, \tau) \]
\[ t=0, 1, 2, \ldots \]

At \( t = 0 \) the optimization problem of the adult with \( h \geq 0 \) units of human capital is to choose a sequence \( \{ c_t(h, \tau) \geq 0, s_t(h, \tau) \geq 0, n_t(h, \tau) \in [0,1]\}_{t=0,1,2,\ldots} \), given, so as to

(15) \[ \text{Maximize } \sum_{t=1} \beta^t u(c_t) \text{ subject to } (10), (11), (12) \text{ and } (14). \]
\[ t=0, 1, 2, \ldots \]

\textbf{Characteristics of Equilibrium}

The set of sequences \( \{(c_t(h, \tau), s_t(h, \tau), n_t(h, \tau), n_t^d(h, \tau) : h \geq 0; x_t, r_t, m_t, H_{mt}, w_t, \tau)\}_{t=0,1,2,\ldots} \) and the initial distribution \( \Psi_0 \) describe the model’s equilibrium such that at each \( t \geq 0 \), the labor demand \( n_t^d(\cdot) \) satisfies (3), the implicit rental price \( r_t \) of human capital satisfies (5), the breakeven skill \( x_t \) satisfies (9), the sequence \( \{c_t(h, \tau), s_t(h, \tau), n_t(h, \tau)\}_{t=0,1,2,\ldots} \) satisfies (15), and \( \{H_{mt}, m_t\}_{t=0,1,2,\ldots} \) coincides with the same generated by the optimal sequence \( \{s_t(h, \tau), n_t(h, \tau)\}_{t=0,1,2,\ldots} \), such that

(16) \[ m_t = \int d\Psi_t(h, \tau), \quad \{h : n_t(h, \tau) = 0\} \]

(17) \[ H_{m,t+1} = (1 - \delta) \int h d\Psi_t(h, \tau) + \int s_t(h, \tau) d\Psi_t(h, \tau), \]
\[ H_0 = \int h d\Psi_0(h, \tau), \]

the labor market clears such that at each date \( t=0, 1, 2, \ldots, \)

(18) \[ \{n_t^d(h, w_t; H_{mt}, m_t) d\Psi_t(h, \tau) = 1 - m_t \}
\[ \{h : n_t(h, \tau) = 0\} \]

Notice that the labor demand function does not depend on the redistributive tax rate because the tax is based on indirect profit not on the output of the firms. On the other hand, the labor supply or equivalently occupational choice as characterized in (16)
depends on the after tax real wage. Nevertheless, the market clearing wage does not depend on \( \tau \) because the profit maximizing firm equates the before tax real wage to the marginal product of labor. Figure 2 illustrates the labor market equilibrium in a situation where subsidy \( z_t \) is negative. Note that because of the discrete occupational choice, the labor supply curve (called \( L' \) schedule) is a step function. At \( \bar{w}_t = 0 \), the basic skill, \( x_t \) equals zero which means \( L' \) equals zero, because everybody chooses to be a manager. At \( \bar{w}_t \) which is the post tax wage when \( x_t = h_0 \), an adult is indifferent between two occupations. This explains horizontal segment BC of the labor supply function over the range \( 1 - m \leq L' \leq 1 \). The labor market equilibrium condition (18) holds at the point D where the MPL schedule intersects the labor supply schedule corresponding \( L' = 1 - m \), as shown in (18).

<Figure 2 comes here>

The goods market clears such that at each date \( t = 0, 1, 2, \ldots \),

\[
(19) \quad \int_{h \geq 0} (c_t(h, \tau) + s_t(h, \tau)) d\Psi_t(h, \tau) = \int_{\{h; n_t(h, \tau) = 0\}} q_t(h, n_t^d(h); H_{mt}, m_t) d\Psi_t(h, \tau_t). \]

The above definition yields a sequence \( \{m_t, H_{mt}\}_{t \geq 0} \) of state variables that characterize the equilibrium, where \( H_{mt} \) denotes the total human capital of managers such that

\[
(20) \quad H_{mt} = \int_{\{h; n_t(h, \tau) = 0\}} h d\Psi_t(h, \tau) \quad \text{for} \ t = 0, 1, 2, \ldots
\]

By (4), (16) and (18) the equilibrium wage rate \( w_t \) is given by

\[
(21) \quad w_t = aA_0 m_t^0 H_t (1 - m_t)^{\alpha-1} \quad t = 0, 1, 2, \ldots
\]

By (21) the wage rate of workers increases with the economy's average human capital, \( H_t \) and the relative proportion of managers, \( m_t \). The former positively influences the productivity of workers through an external effect while the latter augments the relative
scarcity of workers. By (5) and (21) the implicit rental \( r_t \) price of human capital is given by

\[
(22) \quad r_t = (1 - a)A\theta m_t^\theta (1 - m_t)^{\alpha} \quad t = 0, 1, 2,
\]

By (22) the price \( r_t \) of human capital and hence the gross rate of return \( r_t + 1 - \delta \) from the investment in schooling is an inverted-U shaped function of \( m_t \). A new manager generates an external benefit to other managers with her innovative activities. She, however, contributes to the relative scarcity of workers and hence boosts the wage rate thus raising her cost of production. For a low value of \( m_t \), additional benefits are higher than additional costs and, therefore, returns to schooling increases with additional managers in the economy. A high value of \( m_t \), however, turns the balance in the opposite direction. At \( m_t = \theta/(\theta + \alpha) \), the return to schooling reaches its maximum. By (9), (21) and (22) we obtain the following expression for the breakeven skill.

\[
(23) \quad x_t = \frac{aH_t + \tau(1 - a)(1 - m_t)H_t^{1 - a}H_t^a}{(1 - a)(1 - m_t)(1 - \tau)} \quad t = 0, 1, 2,
\]

3. Steady State

We shall restrict our attention to steady state equilibria, which preserve the initial distribution of human capital, set forth in (1). In such a steady state there are two distinct dynasties: workers and managers. Workers starting with zero human capital stay as workers and managers starting with \( h_0 \) units of human capital continue as managers.

There is endogenous growth in such a steady state but the growth occurs in such a way that it perpetuates the initial inequality in the distribution of wealth. In other words, \( m_t \) must remain constant in the steady state such that \( \psi_t(0) = \psi_t(0) = 1 - m \) for all \( t \). By (20) in a steady state equilibrium, the average human capital grows at a constant rate \( \gamma \) starting from its initial state \( mh_0 \).
In order to characterize the steady state, it is important to understand what incentive compatibility conditions will guarantee that managers and workers do not switch their occupations so that $m$ is constant in the steady state. To answer this question we need to analyze the property of the optimal investment function of the managers in the steady state. We have the following proposition:

**Proposition 1:** If the utility function is of the constant relative risk aversion class meaning $U(C_i) = C_i^{1-\lambda} / (1-\lambda)$, in the steady state, a manager invests a constant fraction $i(m, \tau)$ of her human capital in her child's education such that $\delta \leq i(m, \tau) < (1 - \tau)r(m)$. The steady state growth rate is given by:

$$\gamma(m, \tau) = i(m, \tau) - \delta$$

where

$$i(m, \tau) = (\beta(1 - \tau)r(m) + 1 - \delta)^{1/\lambda} - 1 + \delta$$

Proof: Appendix.

Notice that the balanced growth rate $\gamma$ is directly related to the steady state rate of investment, which in turn depends positively on the after-tax return on schooling. Since return on schooling reaches its maximum at $m^* = \theta(\theta + a)$, for a given $\tau$, the balanced growth rate $\gamma$ also attains its maximum at the same $m^*$. In Figure 3, we have illustrated this by drawing $i(m, \tau)$ and $\delta$ schedules. The steady state growth rate is the vertical difference between $i(m, \tau)$ and $\delta$, which reaches its maximum at $m^*$.

<Figure 3 comes here>

We are now ready to characterize the incentive compatibility condition for the managers to invest in human capital. Managers find it incentive compatible to invest in human capital if the indirect life time utility for being a manger exceeds the indirect life time utility for being a worker. In other words, the incentive compatibility condition for being a manger is:
(25) \[ \sum_{i=0} \beta^t U((1-\tau)r(m) - i(m,\tau))h_i \geq \sum_{i=0} \beta^t U(w_i + z_i) \]

Since in the steady state there is balanced growth meaning \( h_t, w_t \) and \( z_t \) all grow at the same rate, it is straightforward to verify that (25) holds if

(26) \[ h_0 \geq h^*(m,\tau,h_0) \]

where

(27) \[ h^*(m,\tau,h_0) = \frac{w(m,h_0) + \tau r(m) \frac{m}{1-m} h_0}{(1-\tau)r(m) - i(m,\tau)} \]

where \( w(m,h_0) \) is the market clearing wage rate at date 0. Given an initial distribution of human capital parameterized by \( (h_0,m) \) and a tax rate \( \tau \), \( h^* \) defines the minimum level of human capital that a manager must possess to attain the same consumption as the worker at the steady state. In other words, if the initial human capital \( h_0 \) is just equal to \( h^* \), all the managers are marginal managers in the sense that they are indifferent as to whether they will invest in human capital or not.

Our next task is to characterize the range of \( m \) over which a steady state equilibrium exists. In others words, we need to know the admissible range of \( m \), over which neither managers nor workers switch their occupations. To obtain such an admissible range, we need to know the properties of \( h^*(m,\tau,h_0) \). Using (27) it is straightforward to verify that \( h^*(m,\tau,h_0) \) is linear homogenous in \( h_0 \) meaning

(28) \[ h^*(m,\tau,h_0) = \xi(m,\tau)h_0 \]

where
\[
\xi(m, \tau) = \frac{am^{1-\theta}A_0}{(1-m)^{1-\theta}} + \frac{\pi(m)m}{1-m} - \frac{1}{(1-\tau)r(m) - i(m, \tau)}
\]

Using (26) and (28), it is straightforward to check the following lemma:

**Lemma 1:** An adult manager will invest in human capital if

\[(30) \quad \xi(m, \tau) \leq 1\]

Next we consider the incentive compatibility condition for workers not to invest in human capital. Along a balanced growth path, at date \(t\) an adult worker may become a manager if and only if she invests \(h_t\) units in human capital. A worker will not undertake such an investment decision if the opportunity cost falls short of the future return meaning

\[(31) \quad U'(w_t + z_t h_t) > \beta U'(C_{t+1}) \{(1-\tau)r_m + 1-\delta\}\]

Next we have the following lemma:

**Lemma 2:** If the utility function is of the constant relative risk aversion class meaning \(U(C_t) = C_t^{1-\lambda}/(1-\lambda)\), inequality (30) is sufficient to guarantee (31).

**Proof:** Appendix

Proposition 1 together with Lemmas 1 and 2 have the following steady state implication. If we consider growing economies with \(\gamma > 0\), in the steady state, the investment fraction, \(i(m, \tau)\) must be positive. This requires that the managers must find it incentive compatible to invest positive amount in the steady state. In such a case, from Lemma 3 it follows that workers do not find it incentive compatible to invest in education. The immediate implication is, therefore, that along a steady state path only managers invest in education and workers do not. Therefore, in the steady state \(H_t = H_m\).
We are now ready to characterize the admissible range of $m$ over which a steady state equilibrium exists. Notice that if the inequality (30) holds, it is incentive compatible for the managers to invest in human capital and stay as managers and workers not to invest in human capital and thus stay as workers. The inequality (30) thus defines the range of $m$ over which steady state equilibrium exists. The steady state equilibrium is thus state dependent in the sense that the initial proportion of managers, $m$ determines the steady state proportion of managers.

For analytical tractability, henceforth we consider a logarithmic utility function for which $λ=1$. In case of a logarithmic utility function, $ξ(m, τ)$ in (30) can be written as:

$$ξ(m, τ) = \frac{am}{(1 - a)(1 - m)} + \frac{dm}{1 - m} + \left[ \frac{1 - \beta}{1 - \alpha} \frac{1}{1 - \tau} + (1 - \beta) \frac{1 - \delta}{r(m)} \right]^{-1}$$

We have the following two lemmas:

**Lemma 3:** For a given $τ$, $ξ(m, τ)$ is monotonically increasing in $m$.

Proof: Appendix.

**Lemma 4:** For any given $τ$, there exists a unique $m_c(τ)$ as a function of $τ$ such that for all $0 < m ≤ m_c(τ)$, a steady state distribution of human capital exists.

Proof: Appendix.

Lemma 4 suggests that there exists a unique critical threshold $m_c(τ)$ at which $ξ(m, τ)$ equals unity. In Figure 4, $m_c(τ)$ is shown as the level of $m$ at which $ξ(m, τ)$ schedule intersects the unit line. The set of admissible steady states is, therefore, given by the set, $0 < m ≤ m_c(τ)$

<Figure 4 comes here>

From Figure 4, it is evident that if $m > m_c(τ)$, it is not incentive compatible for managers to stay as managers and therefore, $m$ declines until $m = m_c(τ)$.

---

8 The issue arises whether in case of $m > m_c(τ)$, workers find it optimal to switch their occupations. It is possible that when $ξ(m, τ) > 1$ (meaning $m > m_c(τ)$), (31) may not hold, which means workers may find it
hand by Lemmas 3 and 4, if \( m < m_c(\tau) \), there is no change in \( m \) because the inequality (30) holds at which both managers and workers do not find it optimal to switch by Lemma 3 and 4. We thus have the following lemma:

**Lemma 5:** For a given \( \tau \), in the steady state either \( m = m_c(\tau) \) or \( m < m_c(\tau) \). If \( m > m_c(\tau) \), \( m \) converges to \( m_c(\tau) \).

It is important to observe that the set of steady state proportion of managers thus has an upper bound \( m_c(\tau) \) which depends on the redistributive tax rate \( \tau \). By changing \( \tau \) the fiscal authority can alter the set of admissible proportions of managers. In order to analyze the optimal redistributive tax policy, it is necessary to determine the sign of the comparative statics derivative of \( dm_c(\tau)/d\tau \). It is straightforward to verify from (32) that \( \xi(m, \tau) \) is monotonically increasing in \( \tau \) meaning, \( \frac{\partial \xi(m, \tau)}{\partial \tau} > 0 \). In other words, \( \xi(m, \tau) \) function in Figure 4 shifts upward when \( \tau \) rises. The immediate implication is that \( m_c(\tau) \) is monotonically decreasing in \( \tau \). In other words, we have the following lemma:

**Lemma 6:** \( \frac{\partial m_c(\tau)}{\partial \tau} < 0 \).

4. **Optimal Redistributive Tax**

We now turn to the central comparative dynamics issue. What is the optimal long-run redistributive tax rate? The optimal tax rate \( \tau \) must maximize society's steady state utility. In case of a constant relative risk aversion class of utility functions maximizing steady state utility is equivalent to maximizing the balanced growth rate \( \gamma \).

Our goal in this paper is to characterize the optimal tax policy rule for the government. The task of finding such an optimal steady state tax policy rule is complicated by the fact

---

optimal to switch occupations. It can be shown that various dynamics for \( m \) are possible. If \( m > m_c(\tau) \), depending on the initial value of \( m \), workers might switch initially but eventually \( m \) declines to \( m_c(\tau) \). Since the primary focus of the paper is on the steady state, we do not get into the stability issues here.

* See Barro and Sala-i-Martin (1995) for a proof.
that the steady state itself depends on the initial distribution of human capital. The optimal tax policy rule is thus state dependent.

In order to determine the optimal tax policy, we use the following strategy. First, we solve the optimal tax policy function, which maximizes the growth rate $\gamma(m, \tau(m))$ subject to the restriction that $\xi(m, \tau)=1$. The motivation for solving this optimal tax problem arises from the fact that over a state space $m \geq m_\tau(\tau(m))$, we already know from Lemma 5 that $m$ falls until $m=m_\tau(\tau)$ at which $\xi(m, \tau)$ exactly equals unity. Second, we solve another optimal tax problem which maximizes $\gamma(m, \tau(m))$ over the state space, $m<m_\tau(\tau(m))$. Finally, we splice these two optimal tax policy rules to obtain the general optimal tax rule for any initial $m$.\footnote{Of course $m$ cannot take any arbitrary value. We restrict the domain of $m$ over which steady state growth as well as skill premium are non-negative. We address this issue in Lemma 7.}

In order to solve the first optimal tax problem we need to derive the range of $\tau$ for which in the steady state $m=m_\tau(\tau)$. We have the following lemma:

**Lemma 7:** The range of $\tau$ for which $m=m_\tau(\tau)$ is such that

$$
(33) \quad \tau_{\text{min}} \leq \tau < \tau_{\text{max}}
$$

where $\tau_{\text{max}}$ solves the following equation:

$$
(34) \quad \tau_{\text{max}} = 1 - \frac{m_\tau'(\tau_{\text{max}})}{1-a}
$$

and $\tau_{\text{min}}$ solves the following equation:

$$
(35) \quad \xi((1-a)(1-\tau_{\text{min}}), \tau_{\text{min}})=1
$$

Also $\tau_{\text{min}} < 0$.

Proof: . Appendix
Next we solve an optimal tax policy that solves the following growth maximization problem:

\[(36) \quad \text{Max } \gamma(m, \tau(m)) \]

s.t. \(\xi(m, \tau(m)) = 1\)

and \(\tau_{\text{min}} \leq \tau < \tau_{\text{max}}\) as specified in Lemma 7.

We have the following two lemmas:

**Lemma 8:** Define \(\tau^*(m)\) as the optimal tax policy that solves the growth maximization problem (36). Such a \(\tau^*(m)\) must be such that

\[(37) \quad \tau^*(m) = \max\{\tau_{\text{min}}, \tau_\gamma(m)\}\]

where \(\tau_\gamma(m)\) solves the root of the following equation

\[
(1-\tau)\frac{\xi_{\frac{\tau}{m}}}{\xi_m} = \frac{m(1-m)}{am-\theta(1-m)}
\]

**Proof:** Appendix.

We next have the following lemma.

**Lemma 9:** For all \(m\), \(\tau_\gamma(m) < 0\) for all \(m\) and \(\tau_\gamma(m) \geq 0\).

**Proof:** Appendix.

To get some intuition for the optimal tax policy function laid out in (37), notice that \(\tau_\gamma(m)\) maximizes growth over the range \(m > m^*\) where growth is decreasing in \(m\) as evident from in Figure 3. In this region an increase in \(\tau\) has two opposing effects on \(\gamma\). It has a distortionary effect that dampens growth by making the investment function \(i(m, \tau)\) shift downward and it has a growth enhancing effect by lowering the proportion of managers because \(m_\gamma(\tau)\) declines as \(\tau\) increases as seen in Lemma 6. The growth is maximized for \(\tau_\gamma(m)\) where these two opposing just balance each other. On the other hand, if \(\tau_{\text{min}}\)
>\tau_f(m)$, then $\tau_f(m)$ falls below the admissible range of $\tau$ and thus the optimal $\tau$ turns out to be $\tau_{\min}$.

As observed earlier the optimal tax function $\tau_f(m)$ spelled out in Lemma 8 and 9 applies only to a subspace of the admissible range of $m$ over which

$$(39) \quad m \geq m_\ell(\tau^*(m)).$$

We next need to formulate the optimal tax policy when $m < m_\ell(\tau^*(m))$. In this case, as seen in Lemma 1, $\xi(m, \tau) < 1$ for all admissible range of the tax rates. This means that a change in $\tau$ has neutral effect on the steady state $m$. Since for a given $m$, $\gamma(m, \tau)$ is inversely related to the tax rate $\tau$ (see equation 24(a)), the optimal tax rate must be the lowest admissible tax rate as per Lemma 7.

In light of this discussion we can now formally characterize the optimal tax rule as a function of the initial state. We have the following central proposition.

**Proposition 2:** Define the optimal tax policy function: $\tau_{\text{opt}}(m): m \rightarrow \tau_{\text{opt}}$. There exists a unique $\hat{m}$ such that the optimal tax policy function is:

$$(40) \quad \tau_{\text{opt}} = \begin{cases} 
\tau^*(m) & \text{if } m > \hat{m} \\
\tau_{\min} & \text{if } m \leq \hat{m}
\end{cases}$$

where $\hat{m}$ solves $m_\ell(\tau^*(\hat{m})) = \hat{m}$.

Proof: We first prove the uniqueness of $\hat{m}$ by noting that $\tau^*(m)$ is either monotonic or a constant function (see Lemma 8) and $m_\ell(\tau) < 0$. //

The crucial feature of this optimal tax policy is that the regardless of the state of $m$, $\tau_{\text{opt}}$ is always negative for all admissible range of the occupational state $(m)$. A positive $\tau$ has the usual distortionary effect by driving a wedge between before and after tax return.
to capital. On the other hand, it has an effect on the occupational distribution, \( m \). If \( m \) is too high (say \( m > m^* \)), by lowering \( m \) it may have a growth enhancing effect. However, the fact that the optimal redistributive tax rate, \( \tau \) is negative simply implies that the distortionary effect always dominates.

The model provides a strong case for a negative capital income tax for all admissible states. In the present context, it means an educational subsidy financed by wage income tax. The optimal redistributive tax rate thus involves taxation of wage income and subsidy to the managers who invest in human capital. Although the sign of the redistributive tax rate is state independent, the level of educational subsidy is not. Notice from Figure 8 that the optimal rate of educational subsidy (\( -\tau_{\text{opt}} \)) is given at the maximum level \( -\tau_{\text{man}} \) until the proportion of managers reaches the threshold \( m \) beyond which the rate of educational subsidy falls as \( m \) rises.

5. Conclusion

This paper addresses a topical issue in fiscal redistributive policy. What is the optimal redistributive capital income tax? In the context of a frictionless identical agents scenario, it is well known that the optimal capital income tax is zero in the steady state. The paper addresses this issue in a scenario with incomplete markets for human capital with endogenous occupational choice. We consider a model with discrete occupational choice with heterogeneity in the occupational distribution arising as an endogenous outcome. The steady state perpetuates the initial inequality in the distribution of human capital. In this environment, if the fiscal authority attempts to redistribute resources from the owners of human capital (who are managers) to the workers (who start off with zero human capital), it adversely impacts the growth rate in the long run. The optimal capital income tax rate is thus negative.

Our analysis provides a strong case for educational subsidy financed by taxing workers. The optimal educational subsidy is, however, dependent on the initial proportion of managers. The normative implication of the model is that economy with less managers warrants more subsidy than economies with excessively high proportion of managers. The optimal educational subsidy is thus state dependent. Our results broadly
agree with the Chamley-Judd finding that a positive capital income tax is an inefficient policy alternative. About the optimal tax policy, our results are stronger than Chamley (1986) and Judd (1985) because we find a rationale for negative capital income taxation or educational subsidy. The model implies that it is socially optimal for the government to perpetuate the income inequality through redistributive wage income taxation. This might give rise to alternative scenarios of endogenous growth and inequality relationship, which may be a topic for future research.

Appendix

Proof of Proposition 1: Define \( i_t = s_t / h_t \). Using equation (12), the law of motion for human capital, along a balanced growth path

\[
h_{i_t+1} / h_t = (1 + \gamma) \iff 1 + i_t - \delta = 1 + \gamma
\]

It immediately follows that \( i_t \) is a constant \( i \) such that

\[\gamma = i - \delta\]  

(A.1)

Next use the first order condition for (15) to obtain

\[\frac{C_{i+1}}{C_i} = \beta^{\lambda/\gamma} (1 - \tau) r_m + 1 - \delta \]^{\lambda/\gamma}

(A.2)

which means the balanced growth rate, \( \gamma \) is:

\[1 + \gamma = (\beta (1 - \tau) r_m + 1 - \delta)^{\lambda/\gamma}\]  

(A.3)

From (A.1) note that for \( \gamma \geq 0 \), it must be that \( i \geq \delta \) which establishes the lower bound for \( i \). Using (14), next verify that the budget constraint for the manager is:

\[c_t + s_t = (1 - \tau_t) r_t h_t\]  

(A.4)
Dividing through by $h_i$, one can immediately check that $i \leq (I-\tau)r(m)$.

Proof of Lemma 2: If $U(C_i) = \frac{C_i^{1-\lambda}}{1-\lambda}$, inequality (31) can be rewritten as:

$$
(A.5) \quad \frac{1}{(w_i + z_i - h_i)^{\lambda}} > \frac{\beta \left[ (1-\tau)r(m) + 1 - \delta \right]}{\left[ (1-\tau)r(m) - i(m, \tau) \right]^\lambda h_i^{\lambda}}
$$

Note that along a balanced growth path, $h_i = h_0 (1+\gamma)^i$

Next using (8) and the steady state balanced growth condition, we can write

$$
(A.6) \quad z_i = z_0 (1+\gamma)^i
$$

where

$$
(A.7) \quad z_0 = \frac{\tau \cdot r(m) \cdot m h_0}{(1-m)}
$$

Using (21) and the balanced growth condition we can write

$$
(A.8) \quad w_i = w_0 \cdot (1+\gamma)^i
$$

where

$$
(A.9) \quad w_0 = \frac{\alpha A_0 m^{1-\theta}}{(1-m)^{1-\delta}} h_0
$$

Using (A.6) through (A.9), the inequality (A.5) can be rewritten as:

$$
(A.10) \quad \frac{1}{w_0 + \frac{\tau \cdot r(m) \cdot m}{1-m} - 1} > \frac{\beta^{1/\lambda} \left[ (1-\tau)r(m) + 1 - \delta \right]^{1/\lambda}}{\left[ (1-\tau)r(m) - i(m, \tau) \right]^{1/\lambda} (1+\gamma)}
$$

Next note from (A.2) that along a balanced growth path

$$
(A.11) \quad \gamma = \beta^{1/\lambda} \left[ (1-\tau)r(m) + 1 - \delta \right]^{1/\lambda} - 1
$$

which means (A.10) reduces to
\[
(A.12) \quad \frac{[w(m, h_0) + \tau(r(m) \cdot m - 1)]}{[(1 - \tau)\tau(m) - i(m, \tau) - i(m, \tau)]} < 1
\]

Using the definition \( \xi(m, \tau) \), the above inequality can be rewritten as:

\[
(A.13) \quad \xi(m, \tau) \cdot \frac{1}{(1 - \tau)\tau(m) - i(m, \tau)} < 1
\]

Next since \( i(m, \tau) < (1 - \tau)\tau(m) \), the inequality (A.13) holds if \( \xi(m, \tau) < 1 \).

Proof of Lemma 3: Taking logarithm on both sides:

\[
(A.14) \quad \ln \xi(m) = \ln \left[ \frac{a}{1 - a} + \tau \right] + \ln \left( \frac{m}{1 - m} \right) - \ln(1 - \beta)(1 - \tau) - \ln \left( 1 + \frac{(1 - \delta)(1 - \tau)^{-1}}{\tau(m)} \right)
\]

Taking derivative with respect to \( m \) on both sides:

\[
(A.15) \quad \xi_m = \frac{\xi(m)}{m(1 - m)} \left[ 1 - (1 - \delta) \frac{am}{(1 - \tau)\tau(m) + 1 - \delta} + (1 - \delta) \frac{\theta(1 - m)}{(1 - \tau)\tau(m) + 1 - \delta} \right]
\]

Note that since \( 0 < a < 1 \) and \( 0 < m < 1 \), the second term in the square bracket is less than unity. Since \( \theta > 0, a > 0, 0 < m < 1 \), \( r(m) > 0 \) and \( \xi(m) > 0 \) implies that \( \xi'(m) > 0 \).

Proof of Lemma 4: Note from (33) that \( \lim_{m \to 0} \xi(m, \tau) = 0 \) and applying l’hopital’s rule

\[
\lim_{m \to 0} \xi(m, \tau) = \infty . \text{ Since by lemma 1, } \xi(m, \tau) \text{ is monotonic in } m, \text{ there exists a unique}
\]

\( m_e(\tau) \), at which \( \xi(m, \tau) = 1 \) holds. This immediately proves that for all \( 0 \leq m \leq m_e(\tau) \), a

steady state distribution of human capital exists.

Proof of Lemma 7: We proceed as follows. First, we determine the lower bound for \( m \).

From Figure 4 it follows that \( m \), \( m_{L1}(\tau) \leq m \leq m_{L2}(\tau) \) to ensure non-negative growth.

This immediately establishes that the lower bound for \( m \) is \( m_{L1}(\tau) \).
In order to determine the upper bound for $m$, we proceed as follows. Define the after
tax skill premium, $\pi_0$, as the difference the after tax (or subsidy) implicit rental income of
the managers and the after tax (or subsidy) wages at date 0. In other words,

$$\pi_0 = r(m)h_0 - \bar{w}(m)$$

which, after using (21) and (22) reduces to

$$\pi_0 = \frac{(1-a)(1-m) - am - \tau(1-a)}{(1-m)^{1-a}}A_0h_0m^{\theta}$$

In a steady state the skill premium must be non-negative otherwise nobody will choose to
be manager. From (A.17) it is straightforward to verify that for non-negative skill
premium (meaning $\pi_0 \geq 0$), $m \leq (1-a)/(1-\tau)$.

There are thus two possible upper bounds for $m$ which are $m^2_\tau(\tau)$ and $(1-a)/(1-\tau)$. We next show that the relevant upper bound for $m$ is $(1-a)/(1-\tau)$ not $m^2_\tau(\tau)$. From Lemma 5 we know that in a steady state, $m \leq m_\epsilon(\tau)$. Given that we focus our attention on
non-negative growth rate only, from Figure 4 it follows that $m_\epsilon(\tau) \leq m^2_\tau(\tau)$. Three
possibilities thus lend themselves:

(i) $m \leq (1-a)/(1-\tau) \leq m_\epsilon(\tau) \leq m^2_\tau(\tau)$: In this case, there is no change in $m$ and

$(1-a)/(1-\tau)$ remains the upper bound for $m$.

(ii) $m_\epsilon(\tau) \leq m \leq (1-a)/(1-\tau) \leq m^2_\tau(\tau)$: In this case, by Lemma 5, $m$ converges to
$m_\epsilon(\tau)$ which means that $(1-a)/(1-\tau)$ is the upper bound.

(iii) $m_\epsilon(\tau) \leq m \leq m^2_\tau(\tau) \leq (1-a)/(1-\tau)$: In this case also, $m$ converges to $m_\epsilon(\tau)$ and

$(1-a)/(1-\tau)$ still remains the upper bound.
We have thus established that the range of admissible \( m \) for which the steady state growth is non-negative and skill premium is positive is:

\[
(A.18) \quad m_L(\tau) \leq m \leq (1-a)(1-\tau)
\]

Next using the bounds for \( m \) as shown in (A.18), we can derive the admissible range of the tax rate, \( \tau \). To establish these bounds for \( \tau \), use equation (23) for the basic skill, \( x_t \). In the steady state, workers do not invest as per Lemma 2 which means \( H_t = H_m \). The basic skill \( x_t \) thus grows at the rate \( \gamma \) starting with the initial value \( x_0 \) given by

\[
(A.19) \quad x_0 = \frac{a + \tau(1-a)}{(1-a)(1-m)(1-\tau)} mh_0
\]

The upper bound for \( \tau \) comes from exploiting the fact that for any given \( m \), \( \frac{x_0}{h_0} < 1 \) \(^{11} \)

which means that

\[
(A.20) \quad \tau < 1 - \frac{m}{1-a}
\]

The upper bound for \( \tau \) in (A.18) is obtained when \( m \) reaches its lower bound \( m_L(\tau) \). This immediately establishes the upper bound for the tax rate \( \tau_{max} \) as set forth in (34).

To obtain the lower bound for \( \tau \), note from (A.18) that \( m \leq (1-a)(1-\tau) \) keeps the skill premium non-negative. Our goal is to find the minimum \( \tau \) that makes \( (1-\tau)(1-a) \) the maximum attainable steady state value of \( m \). In other words, using Lemma 2 it follows that \( \tau_{min} \) must be such that

\[
(A.21) \quad m_L(\tau_{min}) = (1-a)(1-\tau_{min})
\]

From (32) it follows that such a \( \tau_{min} \) must solve (35).

\(^{11}\) Note that we consider a strict inequality here because if \( x_0/h_0 = 1 \), the steady state investment rate equals zero in which case growth rate in (24) becomes negative. To see that \( i(m, \tau) = 0 \), use (27) to verify that in
We next establish that $\tau_{\min}$ must be negative. Define $\phi(\tau) = \xi((1-\tau)(1-a), \tau)$. We first show that there exists a $\bar{\tau}$ at which $\phi(\tau)$ reaches a unique maximum and $\bar{\tau}$ is negative. Using (32) one can verify that $\phi(\tau)$ can be rewritten as:

\[(A.22) \quad \phi(\tau) = \frac{1}{(1-\beta) \left[ 1 + \frac{(1-\delta)(1-a)}{xr(x)} \right]}\]

where $x=(1-a)(1-\tau)$. Next note that the sign of $\phi(\tau)$ depends on the elasticity of $r(x)$ with respect to $x$ which we denote by $e(x)$. Using (22), it is easy to verify that

\[(A.23) \quad e(x) = \frac{ax - \theta(1-x)}{1-x}\]

Notice that $\frac{d(xr(x))}{dx} \geq 0$ according as $e(x) \geq 1$. Next using the fact that

$r(x) = a A_0 x^a (1-x)^\theta$ (see equation 22) and the definition of $x$, it can be verified that

\[(A.24) \quad e(x) = 1 \quad \text{according as} \quad \tau = \bar{\tau}.\]

where $\bar{\tau} = \frac{-a(a+\theta)}{(1-a)(1+a+\theta)}$

Notice that $\bar{\tau}$ is strictly negative. We have thus established that $\phi(\tau)$ is monotonically increasing $\tau$ until it reaches its maximum at $\bar{\tau}$ and from then onward it starts falling with respect to $\tau$.

---

this case in the steady state, $h_0 = h^* = x_0$. Since $x_0 = \omega_0 / r_0$, it immediately means that $i(m, \tau) = 0$. Since we are interested in non-negative growth, we impose strict inequality on the bound for $x_0 / h_0$.
We are now ready to establish that $\tau_{\text{min}}$ in (33) is negative. To illustrate it we have drawn the $\phi(\tau)$ function in Figure 5. Notice that as long as $\phi(\tilde{\tau}) > 1$, there exists two solutions for $\phi(\tau) = 1$ meaning $\phi(\tau)$ will intersect the unit line in Figure 5 twice\(^{22}\) at $\tau_1$ and $\tau_2$ with $\tau_1 < \tau_2$. Since $\tau_1 < \tilde{\tau}$, $\tau_1$ must be strictly negative. Since our goal is to find the minimum $\tau$ for which $m_c(\tau_{\text{min}}) = (1-a)(1-\tau_{\text{min}})$, even though we get two solutions $\{\tau_1, \tau_2\}$ for $\tau_{\text{min}}$, $\tau_{\text{min}} = \tau_1$. This immediately establishes that $\tau_{\text{min}}$ is negative. The intuition for the lower bound for $\tau$ can be obtained by noting that once $m$ approaches its upper bound, $(1-a)(1-\tau)$, $\tau_{\text{min}}$ is the maximum possible subsidy the government can give to the managers to keep the skill premium non-negative and preserve the steady state for the managers.

*<Figure 5 comes here>*

We finally need to establish that the set of admissible tax rates is non-empty meaning $\tau_{\text{min}} < \tau_{\text{max}}$. Define the function $g(\tau)$ as follows:

\[
(A.25) \quad g(\tau) = \frac{a + \tau(1-a)}{(1-a)(1-\tau)} \cdot \frac{m_l^1(\tau)}{1 - m_l^1(\tau)}
\]

Using (A.19) and (34) one may observe that $\tau_{\text{max}}$ must satisfy the following equality

\[
(A.26) \quad g(\tau_{\text{max}}) = 1
\]

One can next verify that

\[
(A.27) \quad g(\tilde{\tau}) = \frac{a}{1 + \theta} \cdot \frac{m_l^1(\tilde{\tau})}{1 - m_l^1(\tilde{\tau})}
\]

\(^{22}\) We have verified numerically that for a reasonable parameter values (particularly for reasonably high value of $\beta$ above .9), $\phi(\tilde{\tau}) > 1$ ensuring two solutions for $\phi(\tau) = 1$.\]
Note that since $\theta > 0$ and $0 < a < 1$, $a/(1+\theta) < 1$. Since $m_\ell^1(\tau)$ is a very small number close to 0, $\frac{m_\ell^1(\tau)}{1-m_\ell^1(\tau)} < 1$. This proves that $g(\tau) < 1$. Next observe from Figure 3 that the $i(m, \tau)$ function shifts downward when $\tau$ rises making $m_\ell^1(\tau)$ increase meaning $m_\ell^1(\tau) > 0$.

Since $\tau_{\min} < \tau$, it follows that $\tau_{\min} < \tau_{\max}$. //

Proof of Lemma 8: Differentiating the function $\gamma(m_c(\tau), \tau)$ with respect to $\tau$, we obtain,

$$
\frac{d\gamma}{d\tau} = \frac{\partial \gamma}{\partial \tau} + \frac{\partial \gamma}{\partial m_c} \frac{dm_c}{d\tau}
$$

(A.28)

From (24) it is clear that the first term on the right hand side of (A.28) is negative. The second product term has a sign ambiguity. From Lemma 6, we know that $\frac{dm_c}{d\tau} < 0$.

From Figure 4 note that for a given $\tau$, $\gamma(m, \tau)$ peaks at $m^* = \theta/(\theta + a)$. This means that $\frac{\partial \gamma}{\partial m_c}$ is negative or positive depending on whether $m$ exceeds or falls short of $m^*$. If $m > m^*$, the growth maximization problem reduces to an unconstrained maximization problem with following first order condition

$$
-\beta r(m) + \beta(1-\tau)r'(m) \frac{dm_c}{d\tau} = 0
$$

(A.29)

Since $m_c(\tau)$ solves $\xi(m, \tau) = 1$, using the implicit function theorem we get,

$$
\frac{dm_c}{d\tau} = \frac{-\xi_{\tau}}{\xi_m}
$$

(A.30)
where \( \xi = \frac{\partial \xi}{\partial \tau} \) and \( \xi_m = \frac{\partial \xi}{\partial m} \). Solution to this solves \( \tau_r \).

Next using (22) observe that

\[
(A.31) \quad \frac{r'(m)}{r(m)} = \frac{\theta (1-m) - am}{m(1-m)}
\]

which means that (A.29) reduces to

\[
(A.32) \quad (1-\tau) \frac{\xi}{\xi_m} = \frac{m(1-m)}{am - \theta (1-m)}
\]

In other words, \( \tau_r(m) \) must solve (A.32).

We still need to show the optimal \( \tau_r \) satisfies the second order condition for growth maximization, use (A.28) and (A.29) to get the second derivative of \( \gamma \) with respect to \( \tau \). One gets

\[
(A.33) \quad \frac{d^2 \gamma}{d\tau^2} = -B \left[ r'(m) \frac{dm}{dt} + r''(m) \frac{dm}{dt} + r'(m) \frac{\partial T(\tau,m)}{\partial \tau} \right]
\]

where

\[
(A.34) \quad T(\tau,m) = (1-\tau) \frac{\xi}{\xi_m}
\]

Note that since \( m > m^* \), \( r'(m) > 0 \). Also, check from (22) that \( r''(m) < 0 \). From (37), \( dm/d\tau < 0 \). We have shown in the proof of Lemma 9 that \( \frac{\partial T(\tau,m)}{\partial \tau} < 0 \). The term in the square bracket in (A.33) is thus positive.

If \( m < m^* \), \( d\gamma/d\tau \) in (A.38) is negative which means that the government must set the tax rate \( \tau \) at the lower bound \( \tau_{\text{min}} \) which means

\[
(A.35) \quad \tau^*(m) = \tau_{\text{min}} \quad \text{if} \quad m \leq m^*
\]

If \( m > m^* \), then following three possibilities arise. (i) If \( \tau_{\text{min}} > \tau_r \), then \( \tau^* = \tau_{\text{min}} \) because \( \tau_r \) falls below the minimum admissible \( \tau \). (ii) If \( \tau_r(m) > \tau_{\text{min}} \) but \( m < m_r(\tau_r(m)) \), then \( \tau^*(m) = \tau_{\text{min}} \) because in this case, a change in \( \tau \) has neutral effect on \( m \). Otherwise, the optimal tax rate, \( \tau^*(m) = \tau_r \). This proves the optimal tax policy function \( \tau^*(m) \). \( Q.E.D. \)
Proof of Lemma 9: Using (A.32), one can verify that $\tau$ satisfies:

$$ (A.36) \quad T(\tau, m) = J(m, \theta) $$

where $T(\tau, m)$ is defined as in (A.34) and $J(m, \theta)$ is defined as:

$$ (A.37) \quad J(m, \theta) = \frac{m(1-m)}{am - \theta(1-m)} $$

Using (32), one gets the following expression for $\xi_{\tau}/\xi_{m}$:

$$ (A.38) \quad \frac{\xi_{\tau}}{\xi_{m}} = \frac{\left[ \frac{1-a}{a + \tau(1-a)} + \frac{1}{1-\tau} - \frac{(1-\delta) r(m)^{-1} (1-\tau)^{-2}}{1 + \frac{1-\delta}{r(m)^{-1}}} \right] m(1-m)}{\left[ 1-\frac{am}{(1-\delta) r(m) + 1-\delta} + \frac{(1-\delta)}{(1-\tau) r(m) + 1-\delta} \right]}$$

Next using (A.34) and (A.38), we get

$$ (A.39) \quad \lim_{\tau \to 0} T(\tau, m) = \frac{\left[ \frac{1-a}{1-\delta} - \frac{1-\delta}{1-\delta + r(m)} \right] [am - \theta(1-m)]}{\left[ 1-\frac{1-\delta}{1-\delta + r(m)} (am - \theta(1-m)) \right]} J(m, a, \theta) $$

Define $A_1 = \frac{1-\delta}{1-\delta + r(m)}$ and $A_2 = am - \theta(1-m)$. (A.39) can then be rewritten as:

$$ (A.40) \quad T(0, m) = \left[ \frac{a^{-1}A_2 - A_1 A_2}{1-A_1 A_2} \right] J(m, a, \theta) $$
Next note that since \(m > m^*\), the term \(A_2 > 1\). Since \(0 < a < 1\), from (A.40), it follows that \(T(0) < I(m, \theta)\). It is straightforward to verify from (A.34) and (A.38) that \(T(I) = 0\).

Next we show that \(T(\tau)\) is monotonically decreasing function of \(\tau\). To show this rewrite \(T(\tau)\) as follows:

\[
T(\tau, m) = \left[ \frac{1 + (1 - \tau)(1 - a)}{a + \tau(1 - a) - \frac{1}{1 + (1 - \tau)(1 - \delta)^{-1}r(m)}} \right] \cdot m(1 - m) \\
\left[ 1 + \frac{(1 - \delta)(\theta(1 - m) - am)}{(1 - \tau)r(m) + 1 - \delta} \right]
\]

(A.41)

It is now straightforward to verify from (A.41) that \(T(\tau)\) is monotonically decreasing in \(\tau\) because the numerator is decreasing and the denominator is increasing in \(\tau\). Thus we have established that \(T(\tau)\) is a continuous function with \(T(\tau) < 0\) and \(T(0) > I(m, a, \theta)\) and \(T(I) = 0\). Hence, \(\tau_r\), being the solution of (A.36) must be negative. Figure 6 illustrates the determination of \(\tau_r\). Note that the schedule \(T(\tau)\) intersects \(I(m)\) at the negative quadrant.

<Figure 6 comes here>

To prove that \(T(\tau)'(m) > 0\), note from (A.37) that \(I_m(m, \theta) < 0\) and

\[
(A.42) \quad |T_m'(0, m)| < I_m(a, m)
\]

because we have already established that the expression in the square bracket is less than unity. From Figure 6 we can inspect it graphically. When \(m\) increases \(I(m, \theta)\) shifts downward. \(T(m, \tau)\) schedule may shift either upward or downward but because of (A.42), \(T(m, \tau)\) shifts less in absolute magnitude than \(I(m, \theta)\). This establishes that the intersection point in Figure 6 must shift to the right meaning \(\tau_r\) must increase. Q.E.D
Figure 3

Figure 4
References


