Endogenous Skill Acquisition and Taxation*

James Bullard†
Federal Reserve Bank of St. Louis

Yibo Zhang‡
Washington University in St. Louis

This version: 26 October 2009

Abstract

We study dynamic Mirrlessian-style taxation in a lifecycle economy. In contrast to the recent Mirrleesian dynamic optimal taxation literature in which individual skills are subject to shocks but otherwise fixed over time, in our model agents make a conscious decision about human capital acquisition given their own aptitude for learning. This aptitude is private information. Human capital accumulation is the engine of growth in our model. We find that there will be no human capital accumulation, and hence no growth in the economy when there is no taxation of any sort. We suggest a taxation scheme which will induce human capital accumulation and hence economic growth in this stylized environment. The key feature of the tax scheme is to provide incentives for human capital accumulation for those that have high aptitude by credibly transferring resources to them later in life, after they have revealed their aptitude. We show that a moderate transfer will induce growth. In general the tax-transfer scheme is highly non-linear.

Keywords: taxation, private information, endogenous growth, human capital.

JEL codes: E62, H20, O4

---

*We thank Marcus Berliant, John Nachbar, Ping Wang, Eric Young for helpful comments, and participants of Midwest Macro Meetings, Midwest Theory Conference and the Graduate Student Conference at Washington University in St. Louis. We also thank German Cubas for a helpful discussion. All errors remain our own.

†Email: bullard@stls.frb.org. Any views expressed are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

‡Corresponding author. Email: yzhang@artsci.wustl.edu

1 Introduction

1.1 The recent taxation literature

Recent research in dynamic optimal taxation that provides an extension of the Mirrlees (1971) model in a general equilibrium framework has been fruitful.[1] As in the original Mirrlees problem, skills and work effort are private information and the government may only know the labor income, the product of skill level and work effort, but not each separately. Based on the assumption that a main risk in economic life is skill risk, this new line of research commonly assumes that agents face skill risk all their lives. This assumption guarantees that the skill level and work effort at each point of time are private information and hence there is no information revelation over time. The unobservable skills are allowed to evolve stochastically overtime with very few restrictions on the evolution. In this context, the optimal taxation scheme will strive to strike a balance between social insurance and incentives.

1.2 Endogenous skill acquisition

Idiosyncratic skill risk is an important aspect of life and it may well be the government's goal to provide a certain degree of insurance against it. However, we wish to focus on an aspect that is usually neglected in this line of literature: people are not born with the skills they use in their daily job; most of them need to devote some time early in their lives to formal training to gain those skills. And, it is also a fact of life that some people benefit more from training than others. In other words, people are born with different aptitude for learning rather than skills.[2] We view this aptitude for learning as private information. We expect that taxation would have an important influence on agents' skill acquisition decision.

[1] See Kocherlakota (2006) for a survey. See also Golosov, Tsyvinski and Werning (2007).

[2] Huggett, Ventura and Yaron (2006) show that differences in learning ability account for the bulk of the variation in the present value of earnings across agents.

In particular, we wish to study how taxation influences agents' human capital accumulation decision in a context in which this decision influences the growth rate of the economy. For this purpose, we allow for endogenous skill acquisition of agents with different aptitudes for learning. The key problem is to design a taxation scheme that will potentially generate a positive growth rate of the economy by inducing different types of agents to make socially desirable skill acquisition decisions.[3] Once the agents make their skill acquisition decisions, their types are fully revealed. The life cycle structure means that there is always private information among the new entrants into the economy. For tractability, we do not consider the possibility of slacking when working and also we do not consider the uncertainty associated with human capital investment.[4]

1.3 What we do

We study a lifecycle economy with human capital accumulation decisions and a government that can tax. The engine of growth in this economy is human capital accumulation, so that if agents do not decide to invest in human capital, the economy will not grow and all future generations will be impacted. Two types of agents are born at each date t, those with a high aptitude for learning, and those with a low aptitude for learning. The aptitude is private information. High types can costlessly pretend to be low types, not invest in human capital, simply working and holding assets as in a lifecycle model without human capital accumulation.

We study two economies in this environment. The first is a *laissez-faire* economy in which the government does not intervene in any way. In this economy, we will show that the high type agents are better off pretending to be low types, and so all agents work and hold assets as in an ordinary lifecycle economy, and since no human capital is accumulated the economy does not grow. In the second economy, we show that the government can

[3] This will improve upon the laissez-faire outcome. We will discuss social welfare later in the paper.

[4] Although we acknowledge that this is an interesting aspect of human capital investment. See Grochulski and Piskorski (2005) and Singh (2008) among others.

improve on this outcome through revenue-neutral taxation. The essence of the tax-transfer scheme is to tax low types and transfer to high types, encouraging the high types to make an investment in human capital when young which then benefits the entire economy by inducing economic growth. Under this tax and transfer scheme, lifetime utility between high and low types is equalized, but the two agent types live different lives. The high types will accumulate human capital and earn relatively high labor income later in life, but will also work longer (retire later). Low types will work with their inherited human capital instead of investing in additional human capital, hold more assets, have lower labor income later in life, and retire earlier.

1.4 Main findings

We find that the tax and transfer scheme that induces human capital accumulation is relatively modest for our calibrated economy. For instance, in the baseline case, the amount of the tax that has to be levied on low types is less than one-half of one percent of a period's labor income. We view this as encouraging, as it may indicate that it is not too difficult to provide appropriate incentives in economies in this class.

The tax scheme that induces growth appears to be regressive, taxing lower consumption agents to subsidize higher consumption agents and hence has a similar flavor to Kocherlakota (2006). Whether the tax is actually regressive depends on how the tax is implemented. We discuss this issue in the main text below. Tax schedules as a function of either labor or capital income are nonlinear.

The *laissez-faire* economy we study has the virtue of life cycle income and wealth equality. High types pretend to be low types, all agents earn the same lifecycle labor income and capital income based on their life cycle asset-holding. This equality is upset by the tax and transfer scheme which induces human capital accumulation and growth, and in this sense the tax scheme induces income inequality beyond the inequality inherent in the lifecycle economy. In particular, in the after-tax economy, labor income of high types is higher in middle and later age. This is because the purpose of the tax scheme is to promote the investment in human capital which is driving growth in these economies. On the other hand, the tax scheme mitigates the wealth inequality inherent in the lifecycle economy.

We think this set of results helps to relate ideas from the more recent dynamic taxation literature to widely-studied frameworks in macroeconomics. Our simple framework also allows us to provide some closed form solutions for optimal choices made by the agents in the economy.

1.5 Related literature

We provide one answer to the question: what is the smallest distortion that will induce positive growth in this stylized economy? We consider the results presented here as a first step in exploring the intriguing relationship between taxation, endogenous skill acquisition and growth.

Our approach is related to work by Kocherlakota (2005, 2006) and Golosov, Kocherlakota, and Tsyvinski (2003). These authors focus on informational or enforcement frictions that are somewhat different from ours. They are able to characterize taxation schemes that implement the constrained Pareto optimal allocation in economies with stochastically evolving skills. We show, in contrast, what a growth-inducing tax scheme would look like in a calibrated economy. We do not discuss optimal taxation.[5] However, our results can be more easily related to taxation schemes existing in the real world.

Along the lines of studying dynamic taxation when human capital is endogenous, our paper is most closely related to papers by Kapicka (2006, 2009) and Bohacek and Kapicka (2008). These authors study the dynamics of optimal taxation when human capital is endogenous. Our study differs in the following respects. First, human capital is unobservable[6] by the government in their models while in ours, learning aptitude will be revealed once

[5] See our discussion about social welfare and also calibration results later in the paper.

[6] In Kapicka (2009) and Bohacek and Kapicka (2008) both agents' productivity and human capital are unobservable.

individual human capital accumulation decisions have been made. Secondly, there is no growth aspect in their model. Thirdly, we study in an overlapping generation framework where agents live for a finite time while they use an infinite horizon framework.[7] Our framework allows us to calibrate our model to lifecycle facts and match some important dimensions such as schooling choices and retirement.

Werning (2007) also studies dynamic taxation but in a Ramsey framework with workers heterogenous in skills.[8] To be exact, he studies the optimal taxation of labor and capital in a dynamic economy subject to government expenditure and technology shocks. Werning finds that the skill distribution plays a crucial role in the determination of Ramsey tax rates. What we do, on the other hand, is to flip the question, what role does taxation play in the determination of the skill distribution given that people make conscious decisions about skill acquisition?

Erosa and Koreshkova (2007) provides a novel approach to studying the impact of income taxation on human capital accumulation where there is a production technology for human capital which takes expenditure and (quality-adjusted) parental time as inputs. The model is also built with heterogeneous individuals, which allows the cross-sectional implications of the theory to be used in parameterizing the human capital technology. They develop a quantitative theory of economic inequality to investigate the effects of replacing the current U.S. progressive income tax system with a proportional one.

Many early studies have also focused on the effects of taxation on growth. For instance, Lucas (1990) calculates that eliminating the capital tax and raising the labor tax in a revenue-neutral way would have a trivial effect on the U.S. growth rate while others strongly disagree. Stokey and Rebelo (1995) find that Lucas's conclusion is robust in a continuous time environment with infinitely lived representative agents.

[7] Kapicka (2009) starts with finite horizen in model setup but sets the time horizen to infinity in calibration.

[8] Heterogeneity is ex-ante, which is also the case in our model. See discussions about this in Werning (2007, p. 5).

Berliant and Leyard (2005) focus on information revelation. They show that the possibility of information revelation and no commitment makes indirect mechanism design (that is, optimal income taxation) extremely complicated even in a two period model.

2 Environment

2.1 Model background

Time t is discrete. There is a single good that can be consumed, and that is considered capital if held as assets. At each date t, a new generation of identical agents is born. Each agent lives for five periods and the economy continues forever. Each generation has a population of mass one, and there is no population growth.

Agents differ in their aptitude for learning, or absorbing knowledge, when they are born. For tractability, we assume there are two types, high ability learners (the high type) and low ability learners (the low type). A fraction p of each generation is of high type and a fraction $1-p$ is of low type. We assume the distribution of types is time invariant and the spread of the distribution is also constant over time. In what follows, we will frequently use the notation $x_t^i(t)$, where subscripts denote birth dates, parenthesis denote real time, and the superscript denotes within the generation agent "type", in particular $i \in \{H, L\}$.

Each agent is endowed with one unit of time in each period of his life. In the first period, agents may choose to invest some fraction of time $\tau_t^i(t)$ in improving labor quality, that is, in formal education or training; we wish to follow Azariadis and Drazen (1990) and think of this as something like postsecondary education. There is no leisure alternative in the first three periods. Then $\tau_t^i(t) \in [0,1)$, and agents devote the rest of their time endowment in the first period, $1-\tau_t^i(t)$, to working. Agents devote all the time in second and third periods working. In the fourth period, households may choose to devote a fraction of their time endowment, $\ell_t^i(t+3) \in [0,1]$, to leisure and work the rest of the time, $1-\ell_t^i(t+3)$. Agents do not work in the last period of their life.[9]

2.2 The training technology

We assume that when agents are born, they inherit the average level of efficiency units in the economy at that date. This might be thought of as "common knowledge" or the average level of human capital in the economy. We denote this quantity by $x(t)$.

All households have access to a training technology in youth. The training technology can be used to augment the agent's efficiency units or productivity, which can then be sold on a competitive labor market in later periods to generate a larger labor income than the agent would otherwise enjoy. The type-specific training technology is specified as follows[10]

$$x_t^i(t+1) = \left[1 + a^i \tau_t^i(t)\right] x(t) \tag{1}$$

where $x(t)$ is the average quality of labor in the economy at t. Thus the ability to increase human capital, or become a "high-skilled worker," depends both on a^i and on the time devoted to training in youth.

We think of a^i as an individual's aptitude for learning. Some would benefit more from formal training than others. To keep the analysis relatively simple, we analyze an extreme case by setting $a^H > 0$ and $a^L = 0$. By pushing a^L all the way to zero we are assuming the low types have no ability at all to absorb knowledge, while the high types have some ability to do so. Of course, the low types still inherit human capital, and they can use these

[9] In constructing the model, what we have in mind is a life span starting from age 18 and ending at age 78. Each period lasts 12 years. So the first period is from age 18 to age 30, the last period is from age 66 to age 78. This will allow us to calibrate to the life-cycle facts in US in the dimension of post-secondary education and retirement. It is also worth noting that the CIA world Factbook 2008 estimate of the US overall life expectancy is about 78.

[10] Our specification of training technology bears a resemblance to the one in Azariadis and Drazen (1990). However, ours differs from theirs in two important respects. First of all, we have a type specific training technology rather than a common training technology. Secondly, the learning rate does not increase over time since poverty trap is not of our concern in this paper.

efficiency units to gain labor income. It is just that they will not benefit from *augmenting* their stock of human capital, while the high types will.

A critical component of our analysis is that the agents know their own type, and this aptitude is private information. In particular, the government or a social planner cannot observe which agents among the young generation are the high ability agents. This information may or may not be revealed in later periods once youthful decisions have been carried out.

Given our assumptions about a^i, $i \in \{L, H\}$, the low type household will never have an incentive to use the training technology, and hence will choose not to devote any time to training. Because of this fact we can simplify the notation somewhat by denoting $a^H = a$ and $\tau_t^H(t) = \tau_t(t) = \tau(t)$.

2.3 Growth

It follows that the economywide average level of efficiency units can be expressed as

$$x(t+4) = \frac{1}{5}\sum_{i=0}^{3}\{p\left[1 + a\tau(t+i)\right]x(t+i) + (1-p)x(t+i)\} + \frac{1}{5}x(t+4). \quad (2)$$

Rearranging and simplifying gives the human capital accumulation equation as

$$x(t+4) = \frac{1}{4}\sum_{i=0}^{3}\left[1 + ap\tau(t+i)\right]x(t+i). \quad (3)$$

This implies that different values of τ will yield different *rates of growth* in labor quality x. It can be shown that the growth rate of human capital is a complicated function of $a\tau p$ along a balanced growth path. Thus the growth rate of human capital, and ultimately all real quantities in this economy, will depend on the fraction of high type agents p, the efficiency of the training technology a, and the time devoted to training by the high types τ.

2.4 Production

Production is carried out by a large number of competitive firms all of which have access to a standard Cobb-Douglas production technology with no technological progress. We analyze these firms as if there were only one of them. The technology takes physical capital and efficiency units—quality-adjusted labor—and combines them to produce a unit of output. The technology is

$$Y(t) = \Lambda K(t)^{\alpha} L(t)^{1-\alpha} \tag{4}$$

where $Y(t)$ is aggregate output, $K(t)$ is aggregate physical capital, α is capital share, Λ is a scale parameter, and

$$\begin{aligned} L(t+1) = {} & p\left[1-\tau(t+1)\right]x(t+1)+(1-p)x(t+1) \\ & + p\left[1+a\tau(t)\right]x(t)+(1-p)x(t) \\ & + p\left[1+a\tau(t-1)\right]x(t-1)+(1-p)x(t-1) \\ & + \left\{p\left[1-\ell_{t-2}^{H}(t+1)\right]\left[1+a\tau(t-2)\right]x(t-2)+(1-p)\left[1-\ell_{t-2}^{L}(t+1)\right]x(t-2)\right\}, \end{aligned} \tag{5}$$

or,

$$\begin{aligned} L(t+1) = {} & [1-p\tau(t+1)]x(t+1)+[1+ap\tau(t)]x(t)+[1+ap\tau(t-1)]x(t-1) \\ & + x(t-2)\times\left\{(1+a\tau(t-2)p)-p\ell_{t-2}^{H}(t+1)(1+a\tau(t-2))-(1-p)\ell_{t-2}^{L}(t+1)\right\}, \end{aligned} \tag{6}$$

is the aggregate level of efficiency units in the economy at date $t+1$. The intensive form is then

$$f\left[k(t)\right] = \Lambda k(t)^{\alpha} \tag{7}$$

where $k(t) = K(t)/L(t)$ is the capital-efficiency units ratio, or the *capital-effective labor* ratio. Competitive firms pay inputs their marginal products. Hence the rental rate on capital is $r(t+1) = \alpha\Lambda k(t)^{\alpha-1}$ and the wage rate per efficiency unit is $w(t) = (1-\alpha)\Lambda k(t)^{\alpha}$. We assume full depreciation of physical capital, implying households face a gross rate of return between t and $t+1$, denoted $R(t)$, which is given by $R(t) = r(t+1)+1-\delta = r(t+1)$.

2.5 Household optimization

Each type i agent solves an optimization problem, choosing $\{c_t^i(t),\, c_t^i(t+1),\, c_t^i(t+2),\, c_t^i(t+3),\, c_t^i(t+4),\, \ell_t^i(t+3),\, \tau_t^i(t)\}$ to maximize

$$U^i = \ln c_t^i(t) + \beta \ln c_t^i(t+1) + \beta^2 \ln c_t^i(t+2) + \beta^3 \left[\ln c_t^i(t+3) + \gamma \ln \ell_t^i(t+3)\right] + \beta^4 \ln c_t^i(t+4). \quad (8)$$

Utility is maximized taking the wage per efficiency unit, $w(t)$, the interest rate, $R(t)$, and the average human capital level $x(t)$ as given. The budget constraint is given by

$$c_t^i(t) + \frac{c_t^i(t+1)}{R(t)} + \frac{c_t^i(t+2)}{R(t)R(t+1)} + \frac{c_t^i(t+3)}{R(t)R(t+1)R(t+2)} + \frac{c_t^i(t+4)}{R(t)R(t+1)R(t+2)R(t+3)} \leq \left[1-\tau_t^i(t)\right] x(t)w(t) + \frac{[1+a^i\tau_t^i(t)]x(t)w(t+1)}{R(t)} + \frac{[1+a^i\tau_t^i(t)]x(t)w(t+2)}{R(t)R(t+1)} + \frac{\left[1-\ell_t^i(t+3)\right][1+a^i\tau_t^i(t)]x(t)w(t+3)}{R(t)R(t+1)R(t+2)}. \quad (9)$$

Low types set $\tau_t^i(t) = 0$, but high types may wish to choose a positive value for $\tau_t^i(t)$. We only consider interior solutions, and we show later in the paper that our calibrated cases have a unique interior solution.

2.6 First order conditions

The first order conditions yield a set of consumption smoothing conditions

$$c_t^i(t+1) = \beta R(t) c_t^i(t), \quad (10a)$$

$$c_t^i(t+2) = \beta^2 R(t)R(t+1) c_t^i(t), \quad (10b)$$

$$c_t^i(t+3) = \beta^3 R(t)R(t+1)R(t+2) c_t^i(t), \quad (10c)$$

and

$$c_t^i(t+4) = \beta^4 R(t)R(t+1)R(t+2)R(t+3) c_t^i(t), \quad (10d)$$

where $i \in \{H, L\}$. The conditions also imply a leisure choice

$$\ell_t^i(t+3) = \frac{\gamma\beta^3 R(t)R(t+1)R(t+2)c_t^i(t)}{[1+a^i\tau_t^i(t)]x(t)w(t+3)}, \tag{11}$$

where $i \in \{H, L\}$, provided $\ell_t^i(t+1) \in (0,1]$. The high type agents will choose

$$\tau_t^H(t) = -\frac{1}{a} + \frac{\gamma\beta^3 R(t)\, R(t+1)R(t+2)c_t^H(t)}{\{a[R(t+2)(R(t+1)w(t+1)+w(t+2))+w(t+3)] - R(t)\, R(t+1)R(t+2)w(t)\}x(t)} \tag{12}$$

provided $\tau_t^H(t) \in (0,1]$. The quantity $c_t^H(t)$ will be linear in $x(t)$ implying $\tau_t^H(t)$ will be stationary.

These conditions in conjunction with the budget constraint imply that solutions depend on first period consumption for the high and low types, which are given by

$$c_t^H(t) = \frac{(1+a)w(t)}{a\left(1+\beta+\beta^2+\beta^3+\beta^4\right)}x(t) \tag{13}$$

and

$$c_t^L(t) = \frac{x(t)}{\left(1+\beta+\beta^2+\beta^3+\beta^4+\beta^3\gamma\right)R(t)\,R(t+1)R(t+2)R(t+3)} \times [R(t)\,R(t+1)R(t+2)R(t+3)w(t) + R(t+1)R(t+2)R(t+3)w(t+1) + R(t+2)R(t+3)w(t+2) + R(t+3)w(t+3)]. \tag{14}$$

Along a balanced growth path, the gross rate of return to physical capital $R(t) = R\ \forall t$, the wage per efficiency unit $w(t) = w(t+1) = w$, and let us label the term $wx(t)$ as potential first period labor income. Then high-type agents will consume a fraction of potential first period labor income, with the fraction given by $\frac{(1+a)}{a\left(1+\beta+\beta^2+\beta^3+\beta^4\right)}$. The low type agents will consume a fraction given by $\frac{\left(1+R+R^2+R^3\right)}{R^3\left(1+\beta+\beta^2+\beta^3+\beta^4+\beta^3\gamma\right)}$. The fraction of time spent in training can be written as

$$\tau_t^H(t) = -\frac{1}{a} + \frac{\gamma\beta^3(1+a)}{a(1+\beta+\beta^2+\beta^3+\beta^4)} \times \\ \frac{R(t)\,R(t+1)R(t+2)w(t)}{a[R(t+2)(R(t+1)w(t+1)+w(t+2))+w(t+3)] - R(t)\,R(t+1)R(t+2)w(t)}, \tag{15}$$

or, along a balanced growth path

$$\tau_t^H(t) = \frac{(1+a)\,\gamma\beta^3 R^3}{a\left(1+\beta+\beta^2+\beta^3+\beta^4\right)[a(R^2+R+1)-R^3]} - \frac{1}{a}. \tag{16}$$

This will tend to be positive for sufficiently large a, that is, a sufficiently productive training technology, given R. However the balanced growth path interest rate itself will change as the incentives for the high type to invest in training change, so that R will also be a function of a in the general equilibrium.

Asset holding in the first period is

$$s_t^i(t) = \left[1-\tau_t^i(t)\right]x(t)w(t) - c_t^i(t) \tag{17a}$$

for $i \in \{H, L\}$. Thus high type agents choosing to devote positive time to training will have lower income in the first period. We may expect that low type agents will hold more assets than high types, both absolutely and as a fraction of first period earnings. The high type agents who have invested a positive amount of time in training are expecting a large labor income payday in the second, the third and part of the fourth periods of life, and so they have less incentive to save. This will tend to bid up the rate of return to asset holding, benefitting the low type agents who rely more on asset holding to finance consumption in the later periods of life. This mechanism will be a crucial feature of the model.

The asset holding in the second, third, and fourth periods will respectively be

$$s_t^i(t+1) = R(t)[(1-\tau_t^i(t))x(t)w(t) - c_t^i(t)] \\ + w(t+1)x(t)(1+a\tau_t^i(t)) - c_t^i(t+1), \tag{17b}$$

$$s_t^i(t+2) = R(t+1)\times$$
$$\{R(t)[(1-\tau_t^i(t))x(t)w(t) - c_t^i(t)] + w(t+1)x(t)(1+a\tau_t^i(t)) - c_t^i(t+1)\}$$
$$+ w(t+2)x(t)(1+a\tau_t^i(t)) - c_t^i(t+2), \quad (17c)$$

and

$$s_t^i(t+3) = R(t+2)s_t^i(t+2) + w(t+3)x(t)(1+a\tau_t^i(t)) - c_t^i(t+3). \quad (17d)$$

In the fourth period of life, leisure choices will be

$$\ell_t^H(t+3) = \frac{\gamma\beta^3 R(t)\, R(t+1) R(t+2)}{w(t+3)\, x(t)(1+a\tau(t))} c_t^H(t) \quad (18)$$

for the high types, and

$$\ell_t^L(t+3) = \frac{\gamma\beta^3 R(t)\, R(t+1) R(t+2)}{w(t+3)\, x(t)} c_t^L(t) \quad (19)$$

for the low types. Hence, we have

$$\frac{\ell_t^H(t+3)}{\ell_t^L(t+3)} = \frac{c_t^H(t)}{(1+a\tau(t))c_t^L(t)}. \quad (20)$$

For a fixed ratio of first period consumption, a higher value of τ implies that high type agents will retire later than the low type.

3 The role of private information

3.1 Overview

The high type agents can pretend to be low type agents for their entire life cycle if it is to their advantage. The reverse case is not possible, so the low types will simply work and hold assets, taking as given the wages, rates of return, and the level of human capital. The high type agents will consider whether it is to their advantage to pretend to be low types; they can emulate low types by simply working the whole time in the first period and forgoing the opportunity to invest in training.

An important feature of this economy is that in the *laissez-faire*, zero tax equilibrium, the high type agents will never choose to invest a positive amount of time in training. Instead, they will pretend to be low types. As a result, no additional human capital will be accumulated and the economy will not grow. Furthermore, the government will not be able to distinguish high types from the true low type agents at any point in their life cycle. We now turn to developing this result.

3.2 General equilibrium

The general equilibrium condition is

$$K(t+1) = L(t+1)k(t+1) = S(t) \tag{21}$$

where

$$\begin{aligned} S(t) = {} & (1-p)s_t^L(t) + ps_t^H(t) + (1-p)s_{t-1}^L(t) + ps_{t-1}^H(t) \\ & + (1-p)s_{t-2}^L(t) + ps_{t-2}^H(t) + (1-p)s_{t-3}^L(t) + ps_{t-3}^H(t) \end{aligned} \tag{22}$$

is aggregate asset holding at t.

We can use the fact that along the balanced growth path, $x(t)$ grows at a constant rate g to further simplify $L(t+1)$ as well as to show that $S(t)$ is linear in $x(t)$. From equation (3), we can solve analytically the constant growth rate g.[11] Further substituting the terms in (6) we can show that $L(t+1)$ is linear in $x(t)$ as well. These facts imply that we can factor $x(t)$ out of equation (21) and substitute appropriately to obtain an equation describing the general equilibrium which depends on $k(t-3)$, $k(t-2)$, $k(t-1)$, $k(t+3)$, $k(t+2)$, $k(t+1)$ and $k(t)$ only. Let us denote this equation by

$$F\left[k(t-3), k(t-2), k(t-1), k(t), k(t+1), k(t+2), k(t+3)\right] = 0. \tag{23}$$

3.3 Incentives to invest in training

We are concerned that the high types have sufficient incentive to invest time in training. These agents undertake a thought experiment at the beginning

[11] The expression for g is too complicated and cumbersome to list explicitly here.

of their first period. We view this thought experiment as local in nature. We imagine a "separation equilibrium" in which all other high type agents are investing positive time in training according to their optimality conditions. An individual agent then contemplates a deviation from this equilibrium. The deviation is to pretend to be a low type agent, ignoring the training technology altogether, and simply work and hold assets in the first period and behave the same way the rest of his lifetime as the low type. If this life cycle choice leads to greater utility at the factor prices dictated by the separation equilibrium, then we say that the separation equilibrium is not robust to this contemplated deviation, or is not *implementable* when we allow for private information. In this economy, such a deviation would then be followed by all high type agents, no human capital accumulation would occur at all, and the economy would cease to grow. One way to generate an equilibrium with growth is to introduce a tax policy to provide a large enough incentive for the high type to invest in human capital.

The incentive condition we have just described can be written as

$$\begin{aligned} &U^H\left[c_t^H(t), c_t^H(t+1), c_t^H(t+2), c_t^H(t+3), c_t^H(t+4), \ell_t^H(t+3)\right] \\ &\quad \geq U^H\left[c_t^L(t), c_t^L(t+1), c_t^L(t+2), c_t^L(t+3), c_t^L(t+4), \ell_t^L(t+3)\right]. \quad (24) \end{aligned}$$

Since preferences are identical this amounts to saying that the high types must receive life cycle utility greater than or equal to the low types. All of the arguments in this inequality are functions of the capital effective labor ratio. Let us write this incentive condition as

$$H\left[k(t+3), k(t+2), k\left(t+1\right), k\left(t\right)\right] \geq 0. \quad (25)$$

3.4 Social welfare

A social welfare function would put some weight on the lifetime utility of each type of agent, discounting the sum of this weighted utility over the infinite future. In steady states of economies with no human capital accumulation and hence no growth, the steady state utilities of each type will be equal and constant. In economies with human capital accumulation and hence positive growth, consumption will grow and hence standards of living will rise over time.

The nature of the model is that, in order to obtain the greatest level of physical capital accumulation, all agents should work and hold assets during their first period of life. Human capital accumulation interferes with this mechanism since agents have to take time away from market work to devote to training. A myopic social planner that only cares about the generation born today would prefer this solution, as it generates the most physical capital for the current generation. However, a sufficiently patient social planner, and in particular one who does not discount the future at all, would always choose a growing economy. We think these points are well understood and so we will assume a sufficiently patient social planner, allowing us to focus on generating equilibria which are associated with human capital accumulation and growth.

4 Two economies

4.1 Calibration

Our calibration strategy involves choosing parameters so that the steady state implications of the benchmark economy presented above are consistent with or close to observations for the United States. We then calculate whether the equilibrium without the no-deviation constraint is implementable.

For convenience, $x(t)$ is fixed at 1. We set the capital share of income, α, equal to 1/3, consistent with evidence from the U.S. NIPA. We set β, the discount factor, equal to 0.737998. This value corresponds to an annual discount factor of 0.975.[12] We set the proportion of high ability learners in each generation, p, equal to 0.5. Calibrating this parameter requires knowledge of a distribution of innate ability, and so there is no clear counterpart in the data. We take the stand that all the people having some college or more are from the high learning ability group and those that take education

[12] Recall that each period corresponds to 12 years.

no further than high school are from the low learning ability group. About 52.5% of the population in the U.S. has at least some college,[13] so choosing 0.5 for p seems reasonable.

We are left with values of three parameters to pin down: Λ, the scale factor in the production technology, γ, the weight on leisure and a, the individual aptitude for learning. We choose the values of these parameters to match three facts about the U.S. economy. First, according to Gordon (198x), the value of the stock of physical capital is similar to the value of the stock of human capital in the U.S. This suggests that we should study equilibria in which the capital to effective labor ratio, k, is near unity. Second, the average retirement age in the U.S. is about 62. This means that average of low type leisure and high type leisure is 1/3 in the fourth period of life. With the benchmark value of γ we chose, people retire at the age of 62.5 on average. Third, the average years of schooling in the U.S. is 14 (for the cohort born 1980), according to Jones and Romer (2009). This suggests that we should study equilibria in which τ is about 1/3.[14] These considerations suggest that we set $\Lambda = 4.2$, $\gamma = .35$, and $a = .67$.

4.2 Laissez-faire

We have not said anything about a government, even though we are posing a tax problem. Following the recent dynamic taxation literature, we start from asking the question: what kind of taxation will take care of the incentive problem. In our setting, we ask what kind of taxation scheme will be able to induce training effort among high ability types.

We first study a *laissez-faire* economy. We want to understand whether the incentive condition would be met in a candidate equilibrium in which the government chooses to play no role whatsoever. In keeping with our concept

[13] Data source: OECD Briefing: Note for US (Education at a Glance 2007).

[14] Education Attainment of the Population 25 years and Over in 2008 also suggests that the average years of post-high school schooling is about 2. This implies in our model τ should be around 0.35. See www.census.gov/population/www/socdemo/educationkps2008.htm

Table 2 Both Sexes

of a local deviation of a single high type agent from a separation equilibrium in which all high type agents are devoting positive time to training, we examine a situation in which such a local deviation can be calculated. We restrict attention to balanced growth path equilibria, in which k solves

$$F\left[k(t-3), k(t-2), k(t-1), k\left(t\right), k\left(t+1\right), k\left(t+2\right), k(t+3)\right] = F\left[k\right] = 0 \tag{26}$$

and human capital grows at a constant rate dictated by the value of k. Solving this equation yields a unique value of $k \approx 0.999477$. The main point concerning this calculation is that the utility of the low types is $U^L \approx 2.48891$ while utility of the high types is $U^H \approx 2.48615$. Thus the condition

$$H\left[k\right] \geq 0 \tag{27}$$

fails. All high types would deviate and become low types; they would invest zero time in training, and the economy would not grow.[15]

Table 1 shows some basic outcomes for the *laissez-faire* economy. In the table, $y_{L,t}(.)$ denotes labor income of generation t, $y_{K,t}(.)$ represents capital income, and $y_t(.)$ is total income. In addition, $c_t(.)$ is consumption, $s_t(.)$ is asset holding, and $s_t(.)/y_t(.)$ is assets as percentage of total income at that stage of the lifecycle.

For future reference we record some key features of the no growth equilibrium here: (1) leisure of the generation t agents at time $t+3$ is 0.276939 implying agents retire at age 62.7; (2) wage per efficiency unit is 2.91105; (3) rate of return on capital is 1.29522 which corresponds to an annual rate of 1.02179; (4) physical capital per efficiency unit of labor along the balanced growth path: 1.12377; and (5) lifetime utility for each generation t agent is 2.52005.

Figure 1 shows[16] consumption, asset holding and total income for the *laissez-faire* economy. Figure 2 shows lifecycle leisure and Figure 3 shows labor income, capital income and total income. These figures can be com-

[15] We also found this result to be quite robust to alternative calibrations.

[16] We begin with Figure 8 here because we use Figures 1-7 later in the text.

Table 1.
Lifecycle results for *laissez-faire* economy.

	t	t+1	t+2	t+3	t+4
$c_t(.)$	2.63736	2.52098	2.40973	2.30339	2.20174
$s_t(.)$	0.273691	0.744567	1.46571	1.69989	0
$y_{L,t}(.)$	2.91105	2.91105	2.91105	2.10487	0
$y_{K,t}(.)$	0	0.35449	.964378	1.89841	2.20174
$y_t(.)$	2.91105	3.26554	3.87543	4.00328	2.20174
$s_t(.)/y_t(.)$	9.4%	22.8%	37.8%	42.5%	0

Table 1: Lifecycle results for the laissez-faire economy.

pared to the case with taxation which we develop below. We stress that there is only one type of life cycle behavior in the *laissez-faire* economy.

4.3 A tax scheme that induces training

We now turn to a second economy, one with a tax-transfer scheme. A successful tax may be able to induce high types to invest in training. We now introduce a government with a commitment technology. They commit or promise to levy a tax on low types and give a transfer to high types in old age. Let us start with a scenario where the government chooses to make the transfer between the two types in their fourth period.[17,18]

The nature of the tax is to levy an amount of tax ξ scaled by the level of human capital $x(t)$ in the fourth period of life on the low type agents, and to make a transfer, also equal to ξ scaled by $x(t)$, to the high type agents.

[17] We will also study scenarios where the government makes the transfer between the two types in other periods (2nd, 3rd or 5th).

[18] We would not consider making transfer in the first period because that would be quite a different problem because of the private information. Hence we are not studying education subsidy. For education subsidy in the context of multidimensional screening see Brett and Weymark (2003).

With this scheme, the budget constraint becomes

$$\begin{aligned}
&c_t^i(t)+\frac{c_t^i(t+1)}{R(t)}+\frac{c_t^i(t+2)}{R(t)R(t+1)} \\
&\quad+\frac{c_t^i(t+3)}{R(t)R(t+1)R(t+2)}+\frac{c_t^i(t+4)}{R(t)R(t+1)R(t+2)R(t+3)} \\
&\quad\le\left[1-\tau_t^i(t)\right]x(t)w(t)+\frac{[1+a^i\tau_t^i(t)]x(t)w(t+1)}{R(t)} \\
&+\frac{[1+a^i\tau_t^i(t)]x(t)w(t+2)}{R(t)R(t+1)}+\frac{\left[1-\ell_t^i(t+3)\right][1+a^i\tau_t^i(t)]x(t)w(t+3)+\xi^i x\left(t\right)}{R(t)R(t+1)R(t+2)},
\end{aligned} \tag{28}$$

for $i\in\{H,L\}$, where $\xi^H>0$, $\xi^L<0$, and $\xi^H=-\xi^L=\xi$. This tax scheme raises no revenue.[19] All equilibrium choices except $\ell_t^H(t+3)$ will be influenced by this tax scheme. In particular,

$$\tau_t^H(t)=\frac{\gamma\beta\left[a\xi+R\left(t\right)\left(1+a\right)w\left(t\right)\right]}{a\left(1+\beta\right)\left[aw\left(t+1\right)-R\left(t\right)w\left(t\right)\right]}-\frac{1}{a}, \tag{29}$$

and thus ξ $\left(=\xi^H>0\right)$ will tend to raise the time devoted to training and so acts as a subsidy to training.

We now solve $F=0$ and $H=0$ for k and ξ. The solution is unique with $k\approx 0.999502$ and $\xi\approx 0.00881286$. At this pair (k,ξ), all conditions for a separation equilibrium are met and in addition, high type agents have no incentive to deviate—lifetime utility of generation t is equalized for both types at approximately 2.48753. The economy grows along a balanced growth path, and the growth rate is about 0.372 percent at an annual rate. This is not as large as the growth rate of the US economy, but is close to the part of the growth rate contributed by human capital accumulation.[20]. The return to physical capital investment is about 2.846% at an annual rate. The wage per efficiency unit is 2.79954.

[19] This is characteristic of the Mirrleesian style taxation literature. If the government wants to raise some tax revenue, then it has to resort to some other sources of taxation. This requires further analysis which we did not undertake here.

[20] Jones and Romer (2009) suggests that annual growth rate contributed by human capital accumulation is .6% or less. Our calculation is consistent with their conclusions.

Table 2.
Lifecycle facts for the growth economy.

	t	t+1	t+2	t+3	t+4
$c_t^H(.)$	2.34171	2.42025	2.50143	2.58533	2.67204
$c_t^L(.)$	2.31972	2.39753	2.47794	2.56105	2.64695
$s_t^H(.)$	$-.515284$	.309626	1.38371	1.89916	0
$s_t^L(.)$	.479812	1.07397	1.82565	1.89887	0
$Ly_t^H(.)$	1.82643	3.45152	3.45152	2.54665	0
$Ly_t^L(.)$	2.79954	2.79954	2.79954	1.90317	0
$Ky_t^H(.)$	0	$-.721638$	.433621	1.93783	2.6597
$Ky_t^L(.)$	0	.67196	1.50405	2.55675	2.6593
$y_t^H(.)$	1.82643	2.72988	3.88514	4.4933	2.6597
$y_t^L(.)$	2.79954	3.4715	4.30359	4.45111	2.6593
$s_t^H(.)/y_t^H(.)$	-28.2	11.3	35.6	42.3	0
$s_t^L(.)/y_t^L(.)$	17.1	30.9	42.4	42.7	0

Table 2: Lifecycle facts for the growth economy.

There are also differences between types in the equilibrium. The high type agents devote about 34.8 percent of their first period time to education, about 4.2 years. High types retire at the age of 63 while low types retire at the age of 62. In the U.S. data, there is a clear correlation between educational attainment and retirement age, with less educated workers tending to retire earlier. For instance, for men aged $60-64$ years, about 45 percent of those with a high school education or a GED equivalent were retired in the year 2000, while only 30 percent of those with a college degree were retired.[21]

The lifecycle patterns also differ by types. Figures 4 through 10 show aspects of the growth economy.

1. Consumption grows over the lifecycle for both types and the high types always consume more than the low types.[22] The low types are

[21] Source: Health and Retirement Survey as cited by Kopecky (2005).

[22] There is no hump-shape consumption path here. Even though our model has labor-leisure choice in the fourth period, the leisure enters the utility function additively so it is like life-cycle models with additively separable utility function defined over consumption

"consumption poor" throughout their lives but they are paying the tax. The extra leisure they enjoy compensates them for the lower consumption profile, but leisure is not taxed by any direct method.

2. A low type holds more assets than a high type in all but the fourth period of the lifecycle. This is intuitive since the high type receives the reward for investing in training so they invest in physical capital less than they would if they did not invest in human capital.

3. Labor income is higher for a high type in all but the first period. This is also intuitive. In the first period, the high type devotes time to training rather than work. The investment in human capital only pays off in the second period and beyond. Both types work the whole time in the second and the third period so the high type earns more labor income than the low type during these periods. In the fourth period, the high type not only earns higher wage but also works more (retires later) than the low type. As a result, the labor income is also higher for the high type in the fourth period.

4. A low type has a higher capital income than a high type in all but the last period. Furthermore in the fifth period, a high type has only slightly more capital income. These follow directly from observation 2.

5. Overall, a low type has a higher amount of total after-transfer income in the first three periods and only lightly less total income in the last two periods.

6. Throughout life a low type devotes a higher portion of his total income to asset holding than his high type counterpart.

7. Labor income of the high type is roughly hump-shaped.

alone and the ratio between next period consumption and current period consumption equals βR, which is greater than one in our calibration.

8. The college premium is about 20%. This is because the high types invest in human capital while the low types simply use their inherited human capital, and the ratio of efficiency units of a high type over a low type is 1.2.

At a first glance, observation 1 and 5 seem to be at odds. How could a high type be able to maintain a higher consumption profile if his total income is constantly lower than a low type? The answer is that the low type has to hold more assets as a fraction of income. The high type can rely on human capital accumulation to generate higher labor income over the lifecycle, but the low type does not have as much human capital and so must save more by holding physical capital.

What is the effect of this separation of skills on the low type? To some extent they also benefit from the separation of skills. They have become a larger provider of physical capital in the economy. With high types devoting time to human capital accumulation and more consumption in each period, there is a relative scarcity of physical capital. This boosts the return to physical capital. In fact, the return to physical capital is higher in the separation equilibrium than the non-separation equilibrium.

So far we have only discussed the lifecycle dimension of the economy. We can also discuss the distributional aspects. Although our model is too simple to do a serious study of sources of inequality in the economy, one insight we can get is that the transfer scheme that induces human capital accumulation and thus growth tends to generate more income inequality beyond the one inherent in the life cycle economy but mitigate the wealth inequality. A simple calculation shows that in the *laissez-faire* economy, the total income gini coefficient is 0.112381 and the bias-corrected one is 0.140476; the wealth gini is 0.298753 and the bias-corrected one is .398337. In the growth economy however, the after-tax total income gini is .140572 and the bias-corrected one is .156191. The wealth gini coefficient is .245 and the bias-corrected one is .280. [23]

[23] For a serious (more computationally intensive) study of sources of income inequality,

4.4 Comparison between the two economies

By computing the total asset holdings in period t for both economies, we see that the total asset holding is much smaller in the economy with positive growth. On the other hand, the total efficiency units of labor are slightly more in this economy. It follows that in this economy the physical capital per efficiency unit is lower, there is a relative scarcity of capital, and the return to capital is higher. This is so because the taxation scheme discourages the high type from holding too much assets. They have to work hard to make a good living.

It is interesting to note that the wage per efficiency unit is lower in the growth economy. This does not directly imply that the low type is worse off since they also benefit from the relative scarcity of physical capital in the growth economy.

The lifetime utility of a current generation agent in the economy with growth is about 1.3% lower than that of a same generation agent in the *laissez-faire* economy. This suggests that the growth is at the expense of sacrificing (albeit little) the welfare of the current generation in return for improvement upon the welfare of the future generations.

Lastly, the high type would retire later than they would in a no growth economy (63 as compared with 62.7) and the low type would retire earlier (enjoy more leisure) than they would in the *laissez-faire* economy (62 vs 62.7). What is more interesting is that the average retirement age in the economy that grows is 62.5, lower than 62.7 in the laissez-faire economy. This is to say, in the economy where the government steps in to promote human capital accumulation, the population on average enjoys more leisure than they would if the government chooses not to intervene.

---

see Hugget et al (2009). Our model does not have uncertainty aspect of human capital investment. Our compuation of these income/wealth gini coefficients are far off from realistic values. This may imply that human capital investment risk is very important in explaining inequality.

Table 3.
Different timing of the taxation-transfer.

	t+1	t+2	t+3	t+4
k	.999502	.999502	.999502	.999502
ξ	.00449337	.00629281	.00881286	.0123421
u	2.48753	2.48753	2.48753	2.48753
$c_t^H(t)$	2.34171	2.34171	2.34171	2.34171
$c_t^L(t)$	2.31972	2.31972	2.31972	2.31972

Table 3: Values of key variables at different timing of the taxation-transfer.

5 Robustness

5.1 Variations of the second economy

We now show that the timing of the transfer makes no difference only the magnitude of the transfer changes.

Our guess is that the transfer needs to be larger if made later. The intuition is we are changing the lifetime discounted utility of both types by making the transfer. Because the discount factor is less than one, a larger transfer is needed for a later one.

A robustness check confirmed this conjecture. Table 3 shows that as the transfer is delayed, capital per efficiency unit of labor along the balanced growth path is the same while the magnitude of the transfer needed increases. Furthermore, the lifetime utility of both types will be equalized at the same level no matter when the transfer is to be made. Finally it is also worth noting that the consumption for both types in the first period of their life are the same irrespective of the timing of the transfer. Since all the other variables such as consumption in later periods, leisure choice and human capital choice depend on these two variables and the return on physical capital only, this would imply that all the other choice variables will have the same value no matter when the transfer is made.

The above results also suggest we can smooth the transfer throughout period 2 to period 5 of agents' lives.

Table 4.
Different timing of taxation.

	t+1	t+2	t+3	t+4
T_k^H	N/A	-1.45%	$-.345\%$	$-.4640\%$
T_k^L	N/A	.418%	.455%	.4641%
T_l^H	$-.130\%$	$-.182\%$	$-.427\%$	N/A
T_l^L	1.26%	.225%	.463%	N/A

Table 4: Effective income tax rates at different timing of the taxation.

5.2 How to make the transfers

In principle, we can make the transfer through a labor income tax, a capital income tax, a combination of both, or a total income tax.

When implemented as a labor income tax, we are essentially making transfers from the lower labor income group to the higher labor income group. This sounds regressive. A similar argument does not apply, however, if we levy a capital income tax. This is because in our calibrated example, the low type almost always earns more capital income than the high type agent at the same stage of their lifecycle.

Notice that capital income taxes are only feasible in one of the last three periods due to the fact that the high types borrow in the first period and hence they do not earn capital income in the second period. On the other hand, labor income taxes are only feasible in the second, third or fourth period because nobody works in the last period in our model.

Table 4 shows that the magnitude of this redistributive taxation is quite small irrespective of the form it takes and the timing when it is levied.

6 Evaluation of conventionally studied taxes

With the results from the previous section, we are able to evaluate conventionally studied taxes such as progressive taxes and flat rate taxes. Again, we start with labor income taxes followed by capital income taxes.

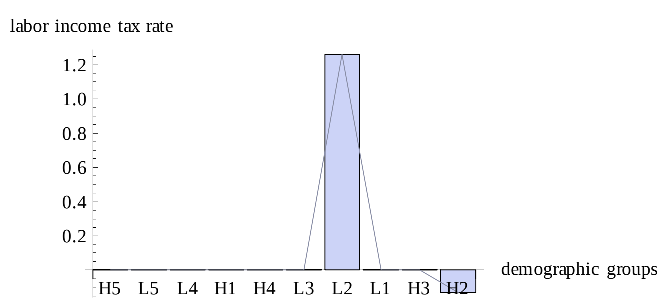
6.1 Labor income taxes

We know the effective labor income tax rate given different timing of the transfer. We also know from a previous section that the timing of the transfer does not matter for the allocation. These suggest that we could evaluate conventionally studied labor income taxes in two ways.

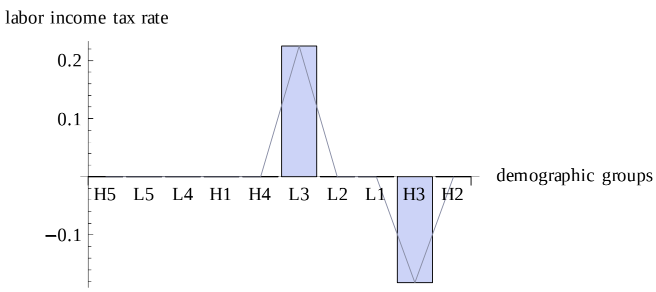
First of all, we can take a snapshot of the labor income of the ten demographic groups at any given point in time. We could plot the effective labor income tax on the two groups that are affected by the transfer and compare with conventionally studied taxes.

In the following graphs, the demographic groups are listed in ascending order of labor income. In particular, the mapping is as follows:

	group	labor income	labor income tax
1	H_{t-4}	0	0
2	L_{t-4}	0	0
3	L_{t-3}	1.66504	0
4	H_t	1.82643	0
5	H_{t-3}	2.22802	0
6	L_{t-2}	2.56086	0
7	L_{t-1}	2.67754	1.26%
8	L_t	2.79954	0
9	H_{t-2}	3.15725	0
10	H_{t-1}	3.30111	$-.130\%$



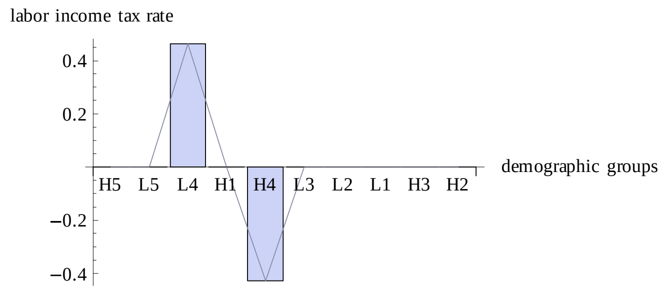
	group	labor income	labor income tax
1	H_{t-4}	0	0
2	L_{t-4}	0	0
3	L_{t-3}	1.66504	0
4	H_t	1.82643	0
5	H_{t-3}	2.22802	0
6	L_{t-2}	2.56086	.225%
7	L_{t-1}	2.67754	0
8	L_t	2.79954	0
9	H_{t-2}	3.15725	$-$.182%
10	H_{t-1}	3.30111	0



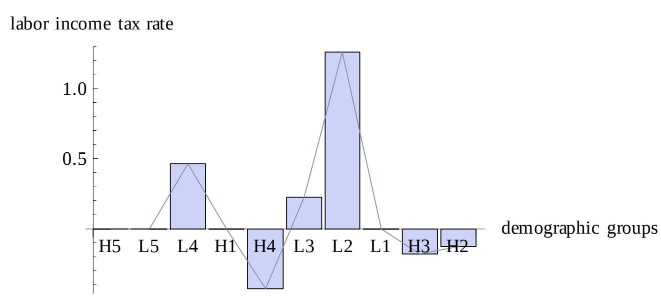
	group	labor income	labor income tax
1	H_{t-4}	0	0
2	L_{t-4}	0	0
3	L_{t-3}	1.66504	.463%
4	H_t	1.82643	0
5	H_{t-3}	2.22802	$-$.427%
6	L_{t-2}	2.56086	0
7	L_{t-1}	2.67754	0
8	L_t	2.79954	0
9	H_{t-2}	3.15725	0
10	H_{t-1}	3.30111	0

Secondly and more interestingly, those effective labor income tax rates given different timing of the transfer provide the upper (lower) bound of the taxes when all the six groups (those in the second, third and fourth period of their life) are affected by the transfer.[24]

[24] Computationally, it will be more difficult to see the exact rate of taxes needed if the



	group	labor income	labor income tax
1	H_{t-4}	0	0
2	L_{t-4}	0	0
3	L_{t-3}	1.66504	.463%
4	H_t	1.82643	0
5	H_{t-3}	2.22802	−.427%
6	L_{t-2}	2.56086	.225%
7	L_{t-1}	2.67754	1.26%
8	L_t	2.79954	0
9	H_{t-2}	3.15725	−.182%
10	H_{t-1}	3.30111	−.130%



In summary, no matter how the transfer is made (be it in one period or smoothed in more than one period), the tax scheme seems to be highly non-linear. This is very different from a flat labor income tax schedule or from a progressive (monotonically increasing) tax schedule.

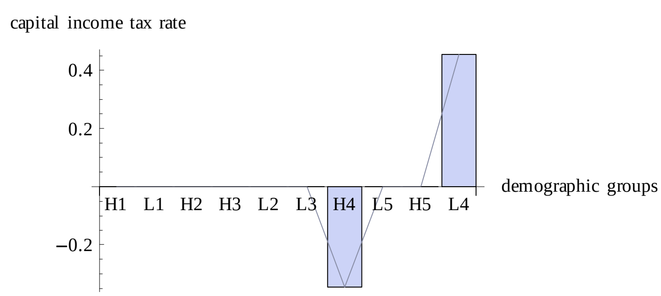
government makes transfer between all the 3 pairs.

6.2 Capital income taxes

Similar to the previous section, we could apply this analysis to capital income taxes. Differently from labor income taxes, however, we can not discuss the capital income tax smoothing case because the capital income distribution changes when the timing of the tax-transfer scheme changes.

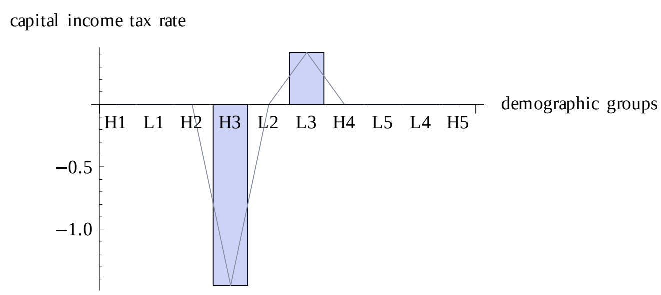
In the following graphs, the demographic groups are listed in ascending order of capital income when the taxation is levied in the fourth period. In particular, the mapping is as follows:

	group	capital income	capital income tax
1	H_t	0	0
2	L_t	0	0
3	H_{t-1}	$-.69019$	0
4	H_{t-2}	.396652	0
5	L_{t-1}	.642678	0
6	L_{t-2}	1.37582	0
7	H_{t-3}	1.69537	$-.345\%$
8	L_{t-4}	2.21485	0
9	H_{t-4}	2.23585	0
10	L_{t-3}	2.23685	$.455\%$



When the taxation is levied in the third period, the mapping is as follows:

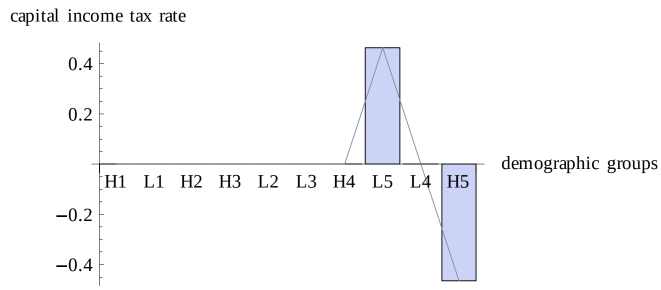
	group	capital income	capital income tax
1	H_t	0	0
2	L_t	0	0
3	H_{t-1}	$-.69019$	0
4	H_{t-2}	.396652	-1.45%
5	L_{t-1}	.642678	0
6	L_{t-2}	1.37582	.418%
7	H_{t-3}	1.70308	0
8	L_{t-4}	2.21485	0
9	L_{t-3}	2.22914	0
10	H_{t-4}	2.23585	0



When the taxation is levied in the fifth period, the mapping is as follows:

	group	capital income	capital income tax
1	H_t	0	0
2	L_t	0	0
3	H_{t-1}	$-.69019$	0
4	H_{t-2}	.396652	0
5	L_{t-1}	.642678	0
6	L_{t-2}	1.37582	0
7	H_{t-3}	1.70308	0
8	L_{t-4}	2.22518	.4641%
9	L_{t-3}	2.22552	0
10	H_{t-4}	2.23585	$-.4640\%$

In each case, the tax schedule is nonlinear.



7 Conclusion

We have studied dynamic taxation in a five period lifecycle economy with endogenous human capital accumulation. In contrast to the recent Mirrleesian dynamic optimal taxation literature in which individual skills are subject to shocks but otherwise fixed over time, in our model agents make a conscious decision about human capital acquisition given their own aptitude for learning. This aptitude is private information. Human capital accumulation is the engine of growth in our model.

We find that there will be no human capital accumulation, and hence no growth in the economy when there is no tax-transfer scheme. This provides a case for redistributive taxation. We suggest a taxation scheme which will induce human capital accumulation and hence economic growth. The key feature of the tax scheme is to tax the low type at some point in their life and make transfer to the high type.

We find the timing of the taxation does not matter for the class of schemes we discuss, and only a very moderate transfer is called for to induce growth. This class of taxation schemes can be implemented in a variety of ways which can be related fairly directly to an actual macroeconomy.

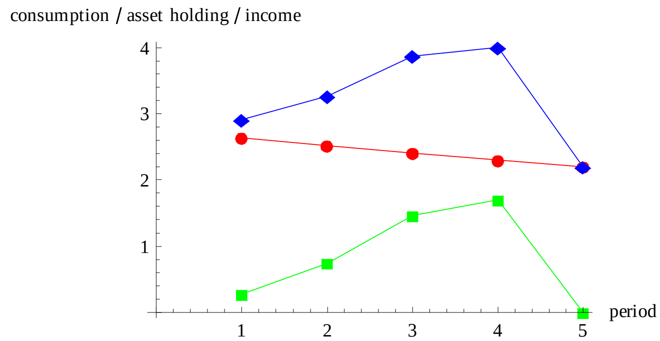


Figure 1: Lifecycle consumption, asset holding and income in the laissez-faire economy.

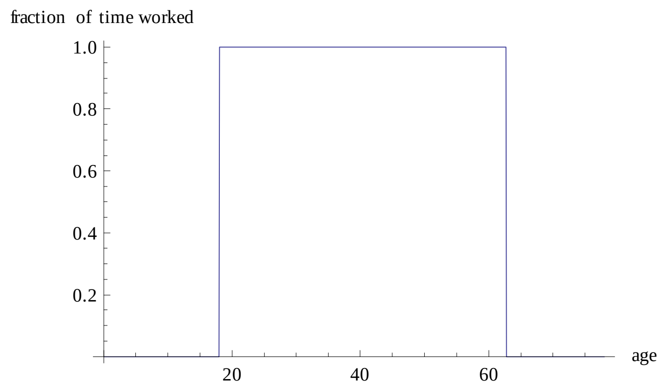


Figure 2: Lifecycle leisure results in the laissez-faire economy.

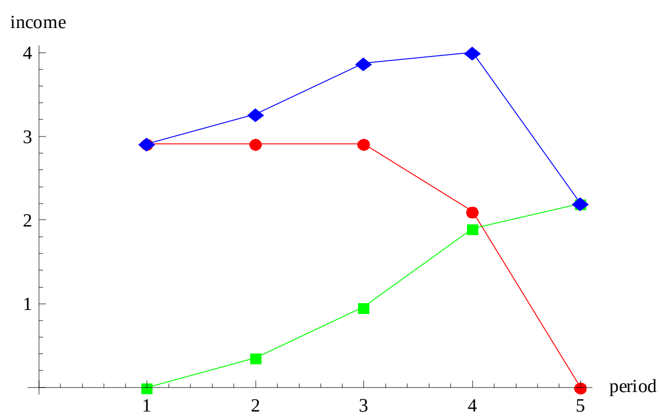


Figure 3: Lifecycle labor income, capital income and total income in the laissez-faire economy.

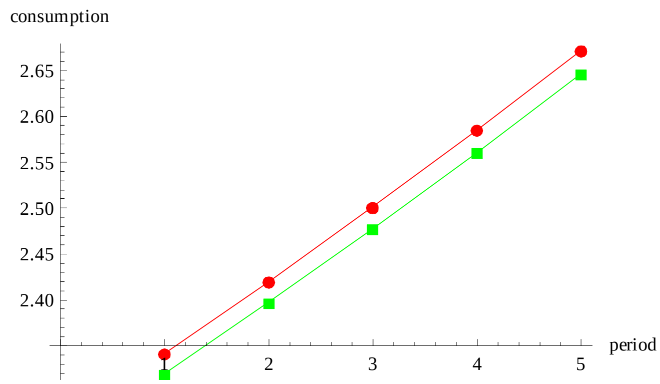


Figure 4: Lifecycle consumption paths of both types in an economy with taxation-transfer.

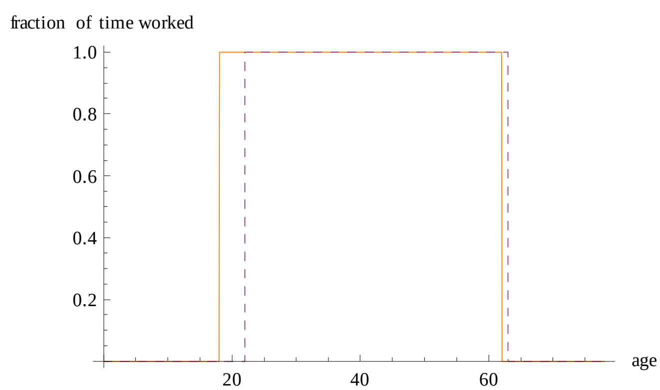


Figure 5: Lifecycle leisure results in the laissez-faire economy.

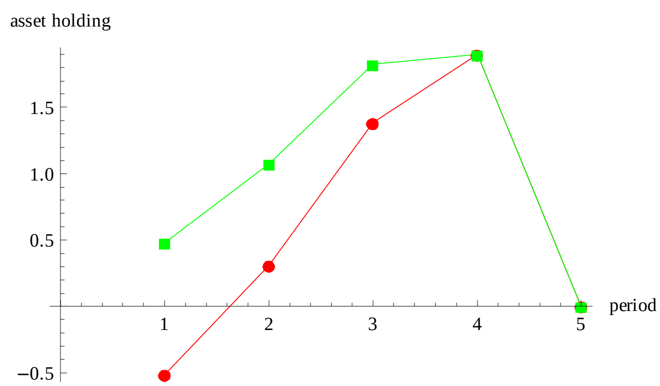


Figure 6: Lifecycle asset holding paths of both types in an economy with taxation-transfer.

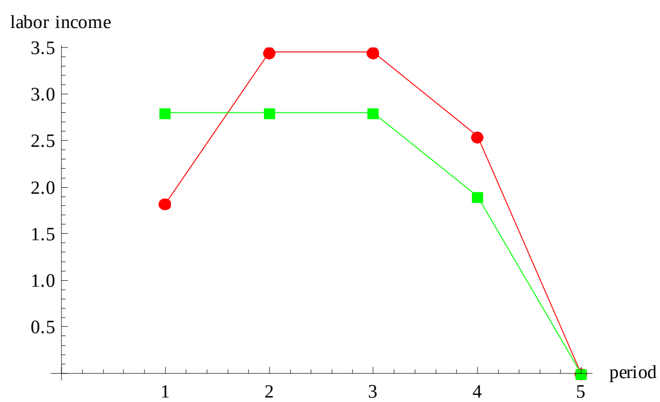


Figure 7: Lifecycle labor income of both types in an economy with taxation-transfer.

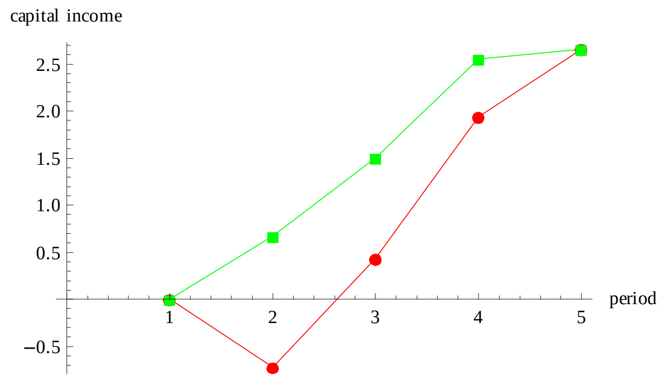


Figure 8: Lifecycle capital income of both types in an economy with taxation and transfer.

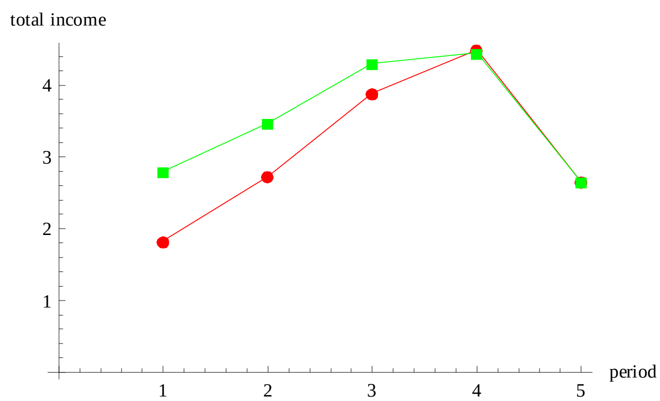


Figure 9: Lifecycle total income of both types in an economy with taxation-transfer.

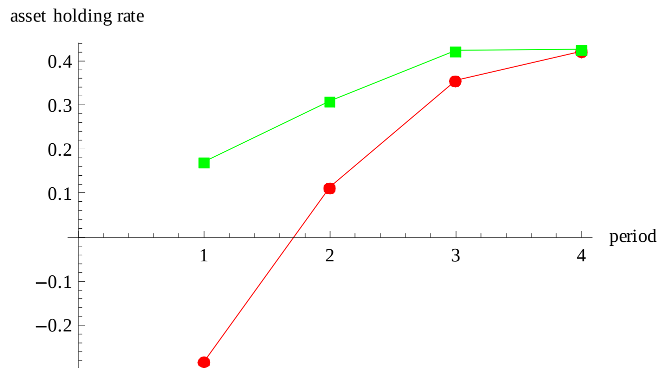


Figure 10: Lifecycle asset holding rates of both types in an economy with taxation-transfer.

References

[1] Atkeson, A.,Chari, V. V. and Kehoe, P. (1999): "Taxing Capital Income: A Bad Idea," *Federal Reserve Bank of Minneapolis Quarterly Review.*

[2] Azariadis, C. and Drazen A. (1990): "Threshold Externalities in Economic Development," *Quarterly Journal of Economics*, 105,501-526.

[3] Boldrin, M. and Montes, A. (2005): "The Intergenerational State: Public Education and Pensions", *Review of Economic Studies*, 72, 651-664.

[4] Berliant, M. and Ledyard J. (2005): "Optimal Dynamic Nonlinear Income Taxes with No Commitment," Working Paper.

[5] Brett, C. and J. A. Weymark (2003): "Financing Education Using Optimal Redistributive Taxation," *Journal of Public Economics,* 87, 2549-2569.

[6] Bohacek, R. and Kapicka, M. (2008):"Optimal Human Capital Policies", *Journal of Monetary Economics*, 55,1-16.

[7] Diamond, Peter A. and Mirrlees, James A. (2002): "Optimal Taxation and The Le Chatelier Principle," MIT Department of Economics Working Paper No. 02-30. Available at SSRN: http://ssrn.com/abstract=331300

[8] Erosa A. and Gervais M. (2001): "Optimal Taxation in Infinitely-Lived Agent and Overlapping Generations Models: A Review," Federal Reserve Bank of Richmond *Economic Quarterly*, Vol 87/2, 23-44.

[9] Erosa A. and Gervais M. (2002): "Optimal Taxation in Life-Cycle Economies,"*Journal of Economic Theory.*

[10] Erosa A. and Koreshkova (2007): "Progressive Taxation in Dynastic Model of Human Capital," *Journal of Monetary Economics,* 54, 667-685.

[11] Fender, J. and Wang, P. (2003): "Educational Policy in a Credit Constrained Economy with Skill Heterogeneity," *International Economic Review*, 44, 939-964.

[12] Gendell, Murray. (2001):"Retirement Age Declines Again in 1990s," *Monthly Labor Review*, October, 12-21.

[13] Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies 70*, 569-88.

[14] Golosov, M., Tsyvinski, A. and Werning, I. (2007): "New Dynamic Public Finance: A User's Guide," the *NBER Macroeconomic Annual 2006,* MIT Press.

[15] Gordon, R. (198x).

[16] Grochulski, B. and T. Piskorski (2005): "Optimal Wealth Taxes with Risky Human Capital," Federal Reserve Bank of Richmond Working Paper No. 05-13.

[17] Hugget, M., Ventura G. and Yaron A. (2006): "Human Capital and Earnings Distribution Dynamics," *Journal of Monetary Economics*, 53, 265-90.

[18] Hugget, M., Ventura G. and Yaron A. (2009): "Sources of Lifetime Inequality," manuscript.

[19] Jones, Charles I. and Romer Paul M. (2009): "The New Kaldor Facts: Ideas, Institutions, Population, and Human Capital," the *American Economic Journal: Macroeconomics,* forthcoming.

[20] Jones, Larry E., Manuelli, Rodolfo E. and Rossi, Peter E. (1997): " On the Optimal Taxation of Capital Income," *Journal of Economic Theory,* 101, 485-517.

[21] Kapicka, M. (2006): "Optimal Income Taxation with Human Capital Accumulation and Limited Record Keeping," *Review of Economic Dynamics*, 9(4), 612-639.

[22] Kapicka, M. (2009): "The Dynamics of Optimal Taxation when Human Capital is Endogenous," UCSD, Working Paper (Revision requested at RED).

[23] Kocherlakota, N. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, Vol. 73, No. 5, 1587-1621.

[24] Kocherlakota, N. (2006): "Advances in Dynamic Optimal Taxation," *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Volume I, 2 .

[25] Kopecky, K. (2005): "The Trend in Retirement." Manuscript, Western Ontario.

[26] Lucas, Robert E., Jr. (1990): "Supply-Side Economics: An Analytical Review," *Oxford Econ. Papers,* 42, 293-316.

[27] Mirrlees, James A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, **38**, 175-208.

[28] Rampini, A. and Bisin A. (2006): "Markets as Beneficial Constraints on the Government," *Journal of Public Economics,* forthcoming.

[29] Singh, A. (2008): "Human Capital Risk in Life-cycle Economies," MPRA Paper No. 10292.

[30] Stokey, N. and Rebelo S. (1995): "Growth Effects of Flat-Rate Taxes," *Journal of Political Economy*, 103, 519-550.

[31] Werning, I. (2007): "Optimal Fiscal Policy with Redistribution," *Quarterly Journal of Economics*, Vol 122, No. 3, 925-967.

[32] Zhang, Y. (2007): "A Hitchhiker's Guide to Optimal Taxation (with annotated bibliography to new advances in Dynamic Optimal Taxation in the end)," mimeo, Washington University in St. Louis.