The Dynamics of Optimal Taxation when Human Capital is Endogenous

Marek Kapicka*

Department of Economics

University of California, Santa Barbara

February 12, 2006

Abstract

This paper characterizes Pareto efficient income taxes in a dynamic economy with human capital accumulation. I extend the tools and insights developed by Mirrlees (1971) into a dynamic framework. I follow Diamond (1998) by assuming that there are no income effects on labor supply. If the government can freely borrow and save, I show that i) the problem of finding efficient allocation can be decomposed into two relatively simple stages and ii) if agents have access to capital market (with zero tax on capital), the efficient allocations may be in some cases implemented in a competitive equilibrium by using history independent income taxes. I compute the sequence of optimal income taxes that implement the optimum *(to be verified)* and show that they marginal income taxes tend to decrease over time and that the gains from adjustment of human capital are about 12 times larger than the static gains from labor supply.

J.E.L Codes: E6, H2

Keywords: optimal taxation, income taxation, human capital

*E-mail: mkapicka@econ.ucsb.edu. Very preliminary and incomplete.

1 Introduction

In this paper I study Pareto efficient allocations and optimal taxes in a dynamic economy with permanent skill shocks and endogenous human capital which is unobservable by the government. I provide a solution method for this class of problems and solve numerically for the dynamics of efficient allocations and optimal taxes.

Recent research on dynamic optimal taxation with private information followed in the footsteps of Mirrlees ([14],[15],[16]). It has focused primarily on cases when the dynamics of efficient allocation is driven by the fact that private information is revealed only gradually over time (Albanesi and Sleet [1] and Battaglini and Coate[2]) or when the driving force is the aggregate state variable (Werning [19]) or both (Golosov, Kocherlakota and Tsyvinski [5], Kocherlakota [12]). In contrast, this paper focuses on a case when the driving force behind the dynamics is an unobservable individual specific state variable, in this case human capital. By assuming that private information shocks are permanent, it also differs from the previously cited literature (with the exception of Werning [19]) by focusing on taxation as a tool of redistribution, rather than the insurance against dynamically evolving shocks. In this sense, it is probably closer to the original Mirrleesian idea of optimal taxation.

A dynamic private information environment where the dynamics is driven by individual state variable rather than by private information is an alternative that has not been much studied so far. One reason is that such problems have been relatively hard to solve. In this paper I provide a key result that makes the analysis and computation of such models tractable. It is the Decomposition theorem of Section 3. I show that the social planner's problem can be conveniently separated into two subproblems: a problem of redistribution between the agents and the problem of finding the efficient labor supply and human capital allocations. This division is interesting theoretically but also provides an algorithm how to solve for the efficient allocations numerically.

This paper is closely related to Kapicka [10] where I analyze optimal steady state allocations when human capital is unobservable, the government is restricted to use current income taxes and agents cannot borrow or save. This paper extends these results in two ways. First, no exogenous restrictions are imposed. Second, I now solve for the whole transitional dynamics of efficient allocations and not just for a steady state. These additional results come at a cost, however. First, I assume that the government can freely borrow and save at an exogenously given interest rate rather than assuming that the resource constraint must clear in each period. Second, I assume that preferences take a very particular form: there are no income effects on labor supply. Such preferences were used recently by Diamond [4] to gain insights into a static optimal taxation problem. Here I show that this specification brings even more benefits in a dynamic setting: it simplifies both the computation of efficient allocations and the problem of their implementation in a competitive equilibrium. In addition, Saez [17] shows that, at least in a static setting, the optimal tax is not so different from the case when income effects are present.

The presence of private information implies that different agent end up with different marginal utilities: they are equal to the shadow price of resources only on average. I introduce a cumulative distortion function to be a function that gives, for each agent, the average distortion for all agents with lower skills. I show that the cumulative distortion function plays a pivotal role in the analysis of the private information economy. The reason is provided in the Decomposition theorem in section 3: For a given agent, all one needs to know to solve for the whole time path of labor supply and human capital allocation is his cumulative distortion. Moreover, the problem of finding these allocations can be conveniently written recursively and solved for numerically. One can thus think of the social planners problem in the following way. The social planner chooses the cumulative distortion function. The agent then solves an individual distorted problem of finding optimal labor, schooling, and human capital allocations for a given cumulative distortion. This decomposition is very convenient since it is the second stage that is the most complicated to analyze. In fact, while the choice of the cumulative distortion function captures all the redistribution aspects of the model, the individual's distorted problem captures all the dynamics that is present in the model.

There are many ways to implement the efficient allocations in a competitive equilibrium. I argue that one particularly appealing implementation, where agent can freely borrow and save and the government uses income taxes that depend only on current income, may be available. The problem with this implementation is that the agent may be able to jointly deviate in labor supply and human capital investment. The answer to whether this is profitable or not is a quantitative one. For some specification, this may be profitable and for some it may not. The no income effect specification of the utility plays an important role in ruling out "triple" deviations in savings, human capital, and labor supply simultaneously. (to be completed): I construct a verification procedure that checks if a particular allocation may indeed be implementable by this simple scheme.

Interestingly, if the implementation with history independent taxes works, the pattern of government policies is much different than in the case when the dynamics of allocations is driven by private information ([12], [1]). Besides using a very simple income tax function, the government's role is not to prevent excessive saving by imposing capital taxes. On the contrary, unrestricted borrowing and saving is essential in that it provides the agents with the ability to transfers resources across time.

The paper is organized as follows. The next section sets up the model. Section 3 characterizes the efficient allocations. Section 4 analyzes how to implement the efficient allocations in a competitive equilibrium. Numerical simulations and computed optimal tax codes are presented in Section 5. Section 6 concludes. The Appendix contains most of the proofs.

2 The Model

Time is discrete, $t \geq 0$. There is a measure 1 of agents in the economy. Each individual is associated with a skill level $\theta \in [0, \bar{\theta}] = \Theta$.

I will assume this skill level does not change over time. This is certainly a very restrictive condition, but is necessary to keep the model tractable. Distribution of skills is given by a distribution F. I assume that F is twice differentiable and has density $f(\theta)$. The skills are private information of each agent: only she knows her own ability. The skills affect earnings of the agent in a way specified below.

Each agent is endowed with one unit of time. At each period, time can be divided between leisure, work, and time spent by human capital accumulation. Denote working time as l_t and time spent by accumulating human capital by s_t.[1]

[1]It is necessary to interpret time spent by accumulating human capital quite extensively. This does not

Period utility of each person depends on consumption c_t and leisure $1 - l_t - s_t$. I assume that the utility function is such that there income effects on leisure are zero: an individual evaluates consumption and leisure sequences according to

$$\sum_{t \geq 0} \beta^t U(c_t + v(1 - l_t - s_t)) \qquad 0 < \beta < 1.$$

where $U : \mathbf{R}_+ \to \mathbf{R}$ is period utility and $v : [0,1] \to \mathbf{R}$ is utility from leisure. I assume that U and v are continuously differentiable on $\mathbf{R}_{++} \times (0,1]$, strictly increasing and strictly concave.

Each individual starts with initial human capital h_0. To reduce the complexity of the model, I assume that h_0 is identical for all people and is observed by the government. Agent's human capital at the beginning of period $t+1$ is denoted h_{t+1}. It depends on time spent accumulating it in previous period s_t, previous level of human capital h_t and is given by a human capital accumulation function $G : \mathbf{R}_+ \times [0,1] \to \mathbf{R}_+$:

$$h_{t+1} = G(h_t, s_t)$$

I assume that G is continuously differentiable on $\mathbf{R}_{++} \times (0,1]$, strictly increasing and strictly concave and that $\lim_{s \to 0} G_s(s,h) = +\infty$ for $h > 0$. Moreover, I assume that $G(h,0) \in (0,h)$ if $h > 0$ and that there is $\bar{h}$ such that $G(\bar{h},1) \leq \bar{h}$.

Human capital affects production abilities of the agent. I assume an efficiency unit specification: a person with human capital h_t, skills θ and working l_t hours produces $y_t = \theta h_t l_t$ at time $t \geq 0$.

It is assumed that the government can borrow or lend at an interest rate $\frac{1}{\beta} - 1$. The government has to finance a sequence of expenditures and the present value of the expenditures is E. The government is supposed to maximize expected discounted utility of an agent that is yet to draw his ability level from the distribution F.

I define an allocation to be a sequence of functions $\sigma = \{c_t, y_t, h_{t+1}\}_{t=0}^{\infty}$ where $c_t : \Theta \to \mathbf{R}_+$ specifies consumption in period t, $y_t : \Theta \to \mathbf{R}_+$ specifies output in period t and $h_{t+1} : \Theta \to$

include only time spent in schools but also other activities that increase individual's human capital, i.e. on-the-job training.

$\mathbf{R}_+$ specifies human capital at the beginning of period $t+1$.[2]. The allocation must satisfy the feasibility constraint:

$$E + \int_{\Theta} \sum_{t \geq 0} \beta^t c_t(\theta) f(\theta) d\theta \leq \int_{\Theta} \sum_{t \geq 0} \beta^t y_t(\theta) f(\theta) d\theta$$

Since agent's skill level is a private information, the social planner needs to elicit agent's type from her. At the beginning of period 0 the agent is asked to report his type to the social planner. The utility of a θ - type agent who reports $\hat{\theta}$ is given by

$$\begin{aligned} V(\theta;\hat{\theta}) &= \max_{\{h_{t+1}\}} \sum_{t \geq 0} \beta^t U[c_t(\hat{\theta}) + v(1 - \frac{y_t(\hat{\theta})}{\theta h_t} - g(h_t, h_{t+1})] \\ & \quad s.t. \quad 0 \leq g(h_t, h_{t+1}) \leq 1 - \frac{y_t(\hat{\theta}_t)}{\theta h_t} \quad \forall t \geq 0 \end{aligned}$$

taking h_0 as given.[3] The incentive compatibility constraint requires the allocation to be such that θ - type agent prefers to report his own type to any other report: For all $\theta \in \Theta$ it is required that

$$V(\theta;\theta) \geq V(\theta;\hat{\theta}) \quad \forall \hat{\theta} \in \Theta \tag{1}$$

An allocation that is both feasible and incentive compatible will be called *incentive-feasible*. Denote the set of all incentive feasible allocations by Σ^{IF}. The social planner chooses an incentive feasible allocation to maximize the expected utility of an agent:

$$W(E, h_0) = \max_{\sigma \in \Sigma^{IC}} \int_{\Theta} \sum_{t \geq 0} \beta^t U[c_t(\theta) + v(1 - \frac{y_t(\theta)}{\theta h_t(\theta)} - g(h_t(\theta), h_{t+1}(\theta))] f(\theta) d\theta. \tag{2}$$

3 Characterizing Efficient Allocations

In this section I will characterize incentive feasible allocations by using the first order and envelope conditions. Before doing so, I will show one important feature of the solution to the social planner's problem: period utility of each agent will be constant over time. This fact will help to simplify the structure of the first order conditions significantly.

[2]Schooling s_t and labor supply l_t can be recovered by inverting the human capital production function and output production function respectively.

[3]The dependence of V on h_0 is kept implicit.

Define period utility $u_t(\theta) = c_t(\theta) + v(1 - \frac{y_t(\theta)}{\theta h_t(\theta)} - g(h_t(\theta), h_{t+1}(\theta))$. Next proposition states that, in the optimum, the social planner will never want period utility of a given agent to vary across time. The reason is that, with quasilinear utility and interest rate equal to the discount rate there are no costs of transferring utility across time in terms of consumption and everyone prefers constant flow of utility to a time varying one. The proof is omitted, because it is straightforward.

Proposition 1 *If an allocation σ solves the social planner's problem then for all $\theta \in \Theta$, $u_t(\theta) = u(\theta)$ for some function $u(\theta)$.*

From now on, I will think of period utility to be the choice variable of the social planner. I will also restrict attention to incentive feasible allocations that exhibit the property that period utility is constant over time and call them *constant utility incentive feasible.* I will also think about allocation in terms of labor supply $l_t(\theta) = \frac{y_t(\theta)}{\theta h_t(\theta)}$ rather than in terms of output. To sum up, an allocation σ will now consist of period utility $u(\theta)$ and sequences of labor supply and human capital $\sigma_{l,h} = \{l_t, h_{t+1}\}_{t=0}^{\infty}$.

I will now derive two sets of necessary conditions for incentive compatibility: the first order condition w.r.t. h_{t+1} and the envelope condition w.r.t. θ. I use the envelope condition rather than the first order condition because it is more general and applies even in the case when allocations are not differentiable in θ.

Proposition 2 *If an allocation is constant utility incentive feasible then*

$$u(\theta) = (1-\beta) \int_0^{\theta} \sum_{t \geq 0}^{\infty} \beta^t v_t' l_t \frac{d\varepsilon}{\varepsilon} + u_0. \tag{3}$$

and

$$\frac{v_t'}{G_{s_t}} \geq \beta(v_{t+1}' \frac{G_{h_{t+1}}}{G_{s_{t+1}}} + v_{t+1}' \frac{l_{t+1}}{h_{t+1}}) \qquad = \textit{if } g(h_t, h_{t+1}) > 0 \tag{4}$$

Proof. See the Appendix. ■

The first equation (3) shows how agent's period utility varies with his type. The variation in period utility is proportional to the informational rent an agent obtains from having a certain type. u_0 is the utility of the the lowest type agent, who gets no informational rent. The second equation (4) is the Euler equation in human capital.

An allocation that exhibits constant period utility, satisfy the resource constraint and the constraints (3) and (4) as *constant utility first order incentive feasible.* The social planner's problem of maximizing expected utility by choosing a constant utility first order incentive feasible allocation will be called a relaxed social planner's problem.

The two constraints (3) and (4) are necessary for an allocation to be incentive compatible, but not sufficient. Thus, the solution to the relaxed social planner's problem may not be identical to the solution of the social planner's problem. There are two reasons why this may be true. First, an individual might find it profitable to deviate jointly in his choice of human capital and in his choice of report. Kocherlakota [13] argues that similar joint deviations are profitable. In the appendix I show a simple two period example that this may happen in my setting as well (to be added). Second, even if human capital sequence were fixed at a given level and joint deviation were not profitable, the envelope condition (3) might still not be enough to prevent deviation in the report itself. This is a well known problem from the static optimal taxation literature.

Therefore, I will consider an ex-post incentive compatibility verification procedure suggested by Abraham and Pavoni. After the optimum is found using only the necessary conditions, I will allow the agents to reoptimize and choose a different report and a different human capital sequence. If the agents decide not to change their behavior, the allocation is incentive compatible and an efficient welfare maximizing allocation is found.

3.1 The Lagrangean

Let λ be the Lagrange multiplier on the resource constraint and $\mu(\theta)f(\theta)$ be the Lagrange multiplier on the incentive compatibility constraint. Define the Lagrangean for the social planner's problem

$$\mathcal{L}(\sigma,\lambda,\mu)=\int_{\underline{\theta}}^{\overline{\theta}}\sum_{t=0}^{\infty}\beta^t\{U(u)+\lambda[\theta h_t l_t-v(l_t+s_t)-u]-\mu[(1-\beta)\int_{\underline{\theta}}^{\theta}v_t' l_t\frac{d\varepsilon}{\varepsilon}+u_0-u]\}fd\theta-\lambda E$$

The efficient allocation is the saddle point of the Lagrangean:

$$W(E,h_0)=\max_{\sigma}\min_{\lambda,\mu}\mathcal{L}(\sigma,\lambda,\mu)\qquad s.t.\ (4) \tag{5}$$

I will now partially characterize the optimum. First order condition in λ is just the requirement that the resource constraint must be satisfied. The resource constraint is now written in terms of u and $\sigma_{l,h}$:

$$\int_{\underline{\theta}}^{\overline{\theta}} uf d\theta \leq (1-\beta)\int_{\underline{\theta}}^{\overline{\theta}}\sum_{t=0}^{\infty}[\theta h_t l_t - v(l_t+s_t)]f d\theta. \tag{6}$$

The Lagrange multiplier λ is therefore such that (6) holds and I will denote such value λ^*. First order condition in $u(\theta)$ and in u_0 are

$$U'(u) = \lambda-\mu \tag{7}$$

$$\int_{\Theta}\mu f d\theta = 0 \tag{8}$$

Combining both first order conditions together, one obtains that the efficient allocations satisfy

$$\int_{\Theta}[U'(u)-\lambda]f d\theta = 0 \tag{9}$$

This equation says that, on average, marginal utility must be equal to the shadow price of resources. It holds because increasing or decreasing utility uniformly for all agents is feasible for the social planner and has no effect on the incentive compatibility constraint. Unlike models with no private information, this equation does not hold for each agent. On the contrary, private information introduces distortions in a sense that marginal utility of almost all agents is not equal to the shadow price of resources.

Instead of taking first order conditions in $\sigma_{l,h}$ I will define a partially optimized Lagrangean $\hat{\mathcal{L}}(\sigma_{l,h})$ where (6), (7) and (8) are assumed to hold. I will denote the value of the

Lagrange multiplier λ such that (6) holds by λ^* and use (7) and (8) to eliminate μ :

$$\begin{aligned}\hat{\mathcal{L}}(\sigma_{l,h}) &= \int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=0}^{\infty} \beta^t \{U(u) + \lambda^*[\theta h_t l_t - v_t - u] + [U'(u) - \lambda^*][\int_{\underline{\theta}}^{\theta} v_t' l_t \frac{d\varepsilon}{\varepsilon} - u]\} f d\theta - \lambda E \\ &= \int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=0}^{\infty} \beta^t \{\lambda^*[\theta h_t l_t - v_t] - v_t' l_t \frac{1}{f\theta} \int_{\underline{\theta}}^{\theta} [U'(u) - \lambda^*] f(\varepsilon) d\varepsilon\} f d\theta + \xi(u) - \lambda E \qquad (10)\end{aligned}$$

where I have defined $\xi(u) = \frac{1}{1-\beta} \int_{\underline{\theta}}^{\overline{\theta}} \{U(u) - U'(u)u\} f d\theta$. The equality uses integration by parts and equation (8) to reverse the order of integration. For a given $\sigma_{l,h}$ the optimal utility allocation u is mechanically given by (3) and so the choice variables of the partially optimized Lagrangean are only $\sigma_{l,h}$.

Define a *cumulative distortion function* $X_{u,\lambda}(\theta)$ to be a function giving, for each agent, the average percentage deviation of marginal utility from the shadow price, the average being taken across all agents with lower skills. The cumulative distortion function is

$$X_{u,\lambda}(\theta) = \frac{1}{\lambda} \int_{\underline{\theta}}^{\theta} [U'(u) - \lambda] f d\varepsilon.$$

The cumulative distortion function will play an important role later on and so it is worthwhile to analyze its properties.

Lemma 3 *Suppose that u is increasing. Then $X_{u,\lambda}(\theta)$ is positive for all $\theta \in \Theta$ In addition, $X_{u,\lambda}(\underline{\theta}) = X_{u,\lambda}(\overline{\theta}) = 0$.*

Proof. $X_{u,\lambda}(\overline{\theta}) = 0$ is obvious. $X_{u,\lambda}(\underline{\theta}) = 0$ follows from (9). Since u is increasing in θ and U is strictly concave, U' is decreasing in θ. Equation (9) then implies that $U'(u) - \lambda$ is first positive and then negative. Consequently, $\int_{\underline{\theta}}^{\theta} [U'(u) - \lambda] f d\varepsilon$ is always positive. ■

The cumulative distortion function is thus is a hump-shaped nonnegative function that starts and ends at 0. The assumption that u is increasing is rather innocuous - it follows directly from the envelope condition (3).

The cumulative distortion function appears directly in the Lagrangean (10). One can thus write is as follows:

$$\hat{\mathcal{L}}(\sigma_{l,h}) = \lambda^* \int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=0}^{\infty} \beta^t \{\theta h_t l_t - v_t - v_t' l_t \frac{1}{f\theta} X_{u,\lambda^*}(\theta)\} f d\theta + \xi(u) - \lambda^* E$$

I will now conjecture and later prove that the social planner's problem can then be written as

$$\begin{aligned} W(E, h_0) &= \max_{\sigma_{l,h}} \hat{\mathcal{L}}(\sigma_{l,h}) \qquad s.t.(4) \\ &= \max_{\sigma_{l,h}} \lambda^* \int_{\underline{\theta}}^{\overline{\theta}} \sum_{t=0}^{\infty} \beta^t \{\theta h_t l_t - v_t - v_t' l_t \frac{1}{f\theta} X_{u,\lambda^*}(\theta)\} f d\theta + \xi(u) - \lambda^* E \qquad s.t.(4) \\ &= \int_{\underline{\theta}}^{\overline{\theta}} \max_{\sigma_{l,h}(\theta)} \sum_{t=0}^{\infty} \beta^t \{\theta h_t l_t - v_t - v_t' l_t \frac{1}{f\theta} X_{u,\lambda^*}(\theta)\} f d\theta + \xi(u) - \lambda^* E \qquad s.t.(4) \end{aligned}$$

Thus, one can break the maximization of the partially optimized Lagrangean into a continuum of separate maximization problems. These maximization problems are interconnected - but only through the cumulative distribution function. In other words, knowledge of $x = X_{u,\lambda}(\theta)$ is sufficient to compute the optimal sequence of labor supply and human capital stock for a $\theta-$type individual. For an arbitrary distortion x, the value of this problem is given by

$$\begin{aligned} Q_0(x, \theta, h_0) &= \max_{\{l_t, h_{t+1}\}_{t\geq 0}} \sum_{t=0}^{\infty} \beta^t \{\theta h_t l_t - v_t - v_t' l_t \frac{x}{f\theta}\} \\ & \quad s.t.(4), \\ & \quad\quad 0 \leq g(h_t, h_{t+1}) \leq 1 - l_t \quad \forall t \geq 0 \end{aligned}$$

Denote the solution to this problem by a sequence of functions $\sigma_{l,h}^d = \{l_t^d, h_{t+1}^d\}_{t\geq 0}$. One way to look at this problem is to view it as an individual's problem, which is affected by a cumulative distortion.[4] I will therefore call this problem an individual's distorted problem. This problem is much simpler than the original problem of finding the utility maximizing report - and that is the main benefit of this formulation. Moreover, it will be shown that the

[4]If $x = 0$ then this problem is identical to the individual's problem with no government intervention.

individual's distorted problem can be conveniently written recursively. Before that, I will summarize and formally prove the results found so far.

Theorem 4 (The Decomposition Theorem) *An allocation σ^*, together with the Lagrange multiplier on the resource constraint λ^* solves the relaxed social planner's problem if and only it is first order incentive feasible and satisfies*

$$\sigma^*_{l,h}(\theta, h_0) = \sigma^d_{l,h}(X^*(\theta), \theta, h_0)$$

for all $t > 0$ and for all $\theta \in \Theta$, where $X^*(\theta)$ is the cumulative distortion function:

$$X^*(\theta) = \frac{1}{\lambda^*}\int_{\underline{\theta}}^{\theta} [U'(u^*) - \lambda^*] f d\varepsilon.$$

and it satisfies $X^*(\bar{\theta}) = 0$.

Proof. See the Appendix. ■

Although the social planner's problem is not a convex problem, sufficiency is obtained even with reliance on the first order condition in u. The reason is that all the nonconvex elements of the problem appear in the individual's distorted problem where I did not rely on the first order conditions.

The theorem is called the Decomposition theorem, because of its main finding - one can decompose the social planner's problem into two related but distinct problems. One problem is to find the utility allocation u. This is, essentially, the problem of redistribution between agents of different skills. The second problem is to find the sequence of labor supply and human capital allocations. This is the dynamic problem where the evolution of human capital plays the major role. These two problems are interrelated. The individual's distorted problem is connected with the problem of redistribution in a very simple way - only through the cumulative distortion function. On the other hand, the redistribution problem is affected by the individual's problem in a much more complex way. In particular, the whole solution to the distorted problem matters for both the resource constraint and the envelope condition (3). What is important, however, is that the individual's distorted problem is so simple, because it is the dynamic element of the model which is by far the hardest to solve for.

I will now turn to the recursive characterization of the individual's distorted problem.

3.2 A Recursive Characterization

The individual's distorted problem involves one endogenous state variable - human capital and two exogenous state variables - individual's skill level θ and the distortion x. Both of these variables work like a fixed effect - they are constant over time. To ensure that the Euler equation (4) holds over time, I introduce an additional co-state variable. Define $d_t = \frac{v_t'}{G_{s_t}}$. It is the marginal disutility from increasing next period human capital by one unit. It is also the left-hand side of the Euler equation (4) and so it becomes a convenient co-state variable. In a recursive representation of the individual's distorted problem, the agent is thus required to choose only allocations such that the benefits from investing in human capital, i.e. the right-hand side of (4) are equal to d. The Bellman equation is

$$\begin{aligned} Q_{x,\theta}(h,d) &= \max_{l,h',d'}\{\theta hl - v - v'l\frac{x}{f\theta} + \beta Q_{x,\theta}(h',d')\} \\ s.t.\ d &= \beta v'(\frac{l}{h} + \frac{G_h}{G_s}) \end{aligned} \tag{11}$$

where h' and d' satisfy their respective laws of motion

$$\begin{aligned} h' &= G(h,s) \\ d' &= \frac{v'}{G_s} \end{aligned}$$

It is straightforward to show that there is an equivalence between this dynamic program and the original sequence program:

$$Q_0(x,\theta,h_0) = \max_d Q_{x,\theta}(h_0,d)$$

since in the first period the allocation does not have to deliver any particular d. Similarly, the solution to the sequence individual's distorted problem can be generated from the solution to the recursive individual's distorted problem.

4 Implementation in a Competitive Equilibrium

I now turn attention to the implementation of the efficient allocations in a competitive equilibrium. One way for the government to do it is to impose income taxes that depend on the whole lifetime profile of incomes. This implementation is a relatively straightforward extension of Werning's [19] result. I will instead use the fact that the utility function is such that there are no income effects on labor supply. I will argue that much simpler implementation may be available. In particular, if people can borrow and save at the interest rate $\frac{1}{\beta} - 1$, one may be able to implement the efficient allocations by using income taxes that depend only on current income.

In Kapicka [10] I show that allocations implementable by current income taxes can be conveniently represented by a mechanism where the social planner has no memory. The agent is asked to report his type each period and the allocations depend only on current report. In general, the mechanism with no memory imposes much more severe restrictions on the allocations than the efficient mechanism introduced in section 2: reporting agent's true type must be preferred to any alternative sequence of reports. This stands in contrast with the requirement (1) that reporting true type must be preferred to any report that is constant over time. By requiring that the agent does not want to deviate in any single period the mechanism with limited memory restricts the ability to transfer informational rents of an agent across time. That's why taxes that depend only on current income are in general inferior to taxes that depend on the whole sequence of incomes.

An efficient allocation will therefore in general not satisfy the constraints of the mechanism with limited memory. Can one make-up for this deficiency with unlimited borrowing and saving? That is, can private markets give the agents the ability to transfer resources across time that was taken away from the government by the restriction of history independent income taxes?

The answer to this question is a qualified yes. To show where the problem lies, consider first a case of exogenous human capital. In this case, using current income taxes together with borrowing and saving will indeed implement the optimum. The argument goes as follows. For any constant utility incentive feasible allocation, one can find another allocation which delivers the same present value of consumption, the same labor supply and satisfies the restrictions of the mechanism with limited memory. That is, this allocation can be implemented with history independent taxes. Allowing the agents to borrow and save will then allow them to choose the efficient sequence of consumption. The assumption of no income effects is the second step of the argument: with income effects the agent might be able to jointly deviate in the sequence of reports and in savings. Next lemma shows the first part of the result formally. For the purpose of this lemma I define an allocation to be a pair of utility and income.

Lemma 5 *Suppose that human capital is exogenous. Then for any constant utility incentive feasible allocation* $\{u, l_t\}_{t\geq 0}$ *there exists a utility sequence* $\{\tilde{u}_t\}_{t\geq 0}$ *such that* $u = (1-\beta)\sum_0^\infty \beta_t^t \tilde{u}_t$ *and an allocation* $\{\tilde{u}_t, l_t\}_{t\geq 0}$ *satisfies for all* $t \geq 0$*, all* $\theta, \hat{\theta}$

$$\tilde{c}_t(\theta) - v(1 - \frac{y_t(\theta)}{\theta h_t} - S(h_t, h_{t+1})) \geq \tilde{c}_t(\hat{\theta}) - v(1 - \frac{y_t(\hat{\theta})}{\theta h_t} - S(h_t, h_{t+1}))$$

where $\tilde{c}_t(\theta) = \tilde{u}_t(\theta) + v(1 - \frac{y_t(\theta)}{\theta h_t} - S(h_t, h_{t+1}))$.

Proof. See the Appendix. ■

The idea of the proof is that with no income effects, one can rearrange the utility sequence in such a way that the present value of utility is the same but, taking any period separately, the agents prefer to tell the truth. That is, offering the agents the option to change her report in each period will not make her better off. But that essentially transforms the problem into a series of static problems. That is, one can induce the agent to supply the same labor supply with mechanism that has no memory and consequently with taxes that depend only on current income. The cost, however, is that the utility sequence is no longer constant.

Suppose now that the agents can borrow or save at the interest rate $\frac{1}{\beta} - 1$. I want to show that when agents are allowed to trade, they will choose the efficient allocation $\{u, l_t\}_{t\geq 0}$. To show this, two things needs to be proven. First, agent's trading must not affect her labor supply. This result is a consequence of the fact that there are no income effects. Second, the agents must choose constant utility u. This result is also easy to show. The budget constraint of a $\theta-$type agent can be written as

$$\sum_{t=0}^{\infty} \beta^t \tilde{u}_t = \sum_{t=0}^{\infty} \beta^t u_t^*.$$

where $\{u_t^*\}_{t\geq 0}$ is the sequence of utility allocations chosen by the agent. Is not hard to see that the agents will choose a constant sequence of utility: $u_t^* = u^*$. Therefore u^* is equal to u. The utility will be the same as under the efficient mechanism. Thus, I conclude that to implement any constant utility incentive feasible allocation $\{u, l_t\}_{t\geq 0}$ one can use income taxes that depend only on current income and allow agent borrow and save. They will end up choosing labor supply $\{l_t\}_{t\geq 0}$ and have period utility u.

With human capital is endogenous, the above result does not hold: an agent my benefit by jointly deviating in the sequence of reports and the sequence of human capital. However, the set of first order incentive feasible allocations that can be implemented with history independent mechanism is the same as the set of first order incentive feasible allocations that can be implemented with efficient mechanism. If there is a profitable deviation in the history independent mechanism, it is in this sense a second order deviation.

Ultimately, the question whether history independent taxes can be used is therefore a quantitative one. In some cases it may be true while in some cases it may not. I will therefore construct an ex-post implementation verification procedure. After the efficient allocation is found, I will check if there is a sequence of reports that may improve upon truthtelling. If there is none, the allocation can be decentralized with history independent income taxes and zero capital taxes. If there is a profitable deviation, one needs to use history independent taxes to implement the efficient allocations.

The implementation verification procedure should not be confused with the incentive compatibility verification procedure. It is much stronger one. If an allocation passes the ex post implementation verification procedure, it passes the ex post incentive compatibility verification procedure. If not, one needs to verify incentive compatibility directly.

5 Numerical Example

In this section I will numerically compute the optimal allocations. I will compare the optimal tax code with the U.S. tax code and show how a potential tax reform should look like. The utility function is assumed to be logarithmic and to exhibit a constant elasticity of labor supply,

$$\begin{aligned} U(u) &= \log(u) \\ v(n) &= \frac{n^{1+k}}{1+k} \end{aligned}$$

where the elasticity of labor supply is given by $\frac{1}{k}$. The human capital production function is assumed to be Cobb-Douglas:

$$G(h,s) = (1-\delta)h + \delta h^{\alpha} s^{1-\alpha}.$$

For the utility function, I choose $k = 2$ so that the elasticity of labor supply is 0.5. For the human capital production function I assume Ben-Porath specification with $\alpha = 0.5$. There is a very diverse evidence regarding both depreciation δ. The evidence ranges from 0.0016 to 0.089, with most of the estimates concentrated around 0.04.[5] This is also the value I have chosen for parameter δ. The time period is one year and so the discount factor β was set equal to 0.96.

I will calibrate the distribution of skills in such a way that the resulting distribution of earnings, assuming steady state, will resemble the empirical distribution of earnings. There is several problems with this approach. First, to meaningfully calibrate the model, one needs to assume that the initial human capital stock varies across people. But, why can't the social planner then use this information or any other previous information to deduce the distribution of skills? This problem is hard to escape with permanent type of shocks. I suggest the following resolution. Suppose the tax reform takes place at time 0. The world has started in period $-T$, when human capital was identical across population. At this time a constitution was written saying that only income taxes that depend on current income may be used. In the light of the previous section, this provision was not restrictive. Neither the shape of the income tax function nor capital taxes were specified in the constitution and were chosen by previous governments. In particular, current U.S. income tax schedule was chosen and savings were not allowed. When the tax reform takes place in period 0, the government must follow the constitution and so must use current income taxes.

Tax return data for 1992 are used for the empirical distribution of earnings. Since the data are hence the implied skills levels not very smooth, I use a double Pareto-Lognormal

[5] See the evidence in Browning, Hansen and Heckman [3] or Trostel [18].

distribution to approximate the empirical distribution of skills. This distribution combines lognormal distribution with heavy Paretian tails and replicates the empirical distribution reasonably well.[6] To calibrate the distribution of skills I first construct the U.S. income tax code. I use the NBER TAXSIM program to construct the effective federal marginal income tax schedule for calendar year 1992. The agent in the model is supposed to represent a household. Thus I restrict attention to married couples with two children. I adjust the income tax schedule by adding state income tax and sales tax as a linear tax.[7] What needs to be determined are government transfers, which are supposed to be independent of income and thus correspond to the negative of income tax at income equal to 0. Their value will be determined endogenously to balance the government budget. It is supposed that consumption to income ratio is 0.75. Thus, government expenditures are equal to 25% of total income.

To solve for the efficient allocations I follow the following procedure. I fix the utility allocation u and the Lagrange multiplier on the resource constraint λ. I then compute the cumulative distortion function $X_{u,\lambda}$ and adjust u_0 until this function satisfies $X_{u,\lambda}(\overline{\theta}) = 0$. I then solve the dynamic program 11 for each agent and generate the sequence of human capital and labor supply allocations for 250 periods. The envelope condition (3) is then used to compute a new utility allocation Tu. I repeat the procedure until $\|Tu - u\| \leq \varepsilon$ for some error tolerance ε. After that, I check if the resource constraint holds with equality (up to an error tolerance). If not, I update the Lagrange multiplier λ until the resource constraint is satisfied with equality.

The results are shown in Figures 1 and 2. Figure 1 depicts the marginal tax rates in periods 1, 50, and 250 (when the economy is more or less in a steady state)[8]. The Figure shows that marginal income tax rates decrease over time. They decrease more or less uniformly, by the same amount for each income level. The largest decreases are concentrated in the initial periods. The intuition is that it is the future marginal tax rate that drives

[6]See Jorgensen and Reed [9] for the definition of the distribution.

[7]Tax rates for the state income tax and sales tax were obtained by dividing government receipts from these taxes by labor income and consumption respectively. Their values were 2.78% and 7.06% .

[8]Marginal income taxes in period 1 appear not to be smooth. This is because the current marginal tax rates are not smooth as well and this property is inherited by the initial distribution of human capital.

investment in human capital. Current marginal tax rates are therefore less important for the investment and the social planner can afford to set them higher.

[INSERT FIGURE 1 HERE]

The overall gain in the production can be decomposed into two elements. First, there is a static gain in the first period, when only labor supply can be adjusted. This gain turn to be fairly small, about 2.2% of aggregate income. Then there is the dynamic gain in the long run, when the human capital is adjusted. Figure 2 shows that this gain is much larger: in the long run, the aggregate income increases by as much as 28% of initial production.

[INSERT FIGURE 2 HERE]

Figure 2 also shows how the government expenditures are financed. The government runs a deficit in the first 18 periods. After that, the production abilities of the economy are improved and the government starts paying off the deficit.

5.1 Implementation Verification Procedure

TO BE COMPLETED

6 Conclusions

This paper analyzes the efficient allocations in a dynamic economy where private information skill shocks are permanent and human capital is endogenous and unobservable by the government. The main contribution is to provide a tractable framework which can be used to analyze these allocations as well as to provide an algorithm to compute them numerically. I also discuss the problem of implementation of the efficient allocations in a competitive equilibrium and show it may be possible to do so in a very simple way: with history independent taxes and unlimited borrowing and saving. The features of government policies are thus very different from the case when the dynamics of allocations is driven by private information: the tax function is much simpler and borrowing and lending by an individual is not distorted. Numerical simulations reveal several interesting results: Marginal income taxes should decrease over time and this decrease is more or less uniform over all ranges of income. The production gains from the adjustment of human capital are much larger than the gains from adjustment of labor supply. Government should first run deficit to finance the tax reform.

The results were found under the assumptions that government can freely borrow and save at a given interest rate, and utility function is such that there are no income effects. What happens if these assumptions, as well as the assumption of permanent skills, are relaxed? The assumption of permanent shocks is crucial for one reason: it ensures that the cumulative distortion is constant over time. If the private information about the shocks is revealed over time, this will no longer hold. In this case, there are two sources of dynamics: the evolution of private information and the evolution of individual specific state variable. Both elements are probably important in determining the efficient allocations and government policies. Whether such problem can be successfully solved depends on the ability to write the Decomposition theorem recursively. One faces similar complications when the assumption that government can borrow or save is relaxed. Again, the main implication is that the time profile of cumulative distortions will no longer be constant.

When the assumption of no income effects is relaxed, one faces two types of problems. First, the individual's distorted problem is complicated by the fact that it also depends on individual's utility. Thus, the dimension of the state space increases. Second, one may face additional complications with the implementation since now a joint deviation in savings, human capital investment and labor supply may be profitable. In such case one needs to use more complicated tax functions to implement the optimum.

References

[1] Albanesi, S. and Sleet, Ch. (2003). "Dynamic optimal Taxation with Private Information." Discussion Paper 140, Federal Reserve Bank of Minneapolis.

[2] Battaglini, M. and Coate, S. (2003). "Pareto Efficient Income Taxation with Stochastic Abilities." Princeton University, mimeo.

[3] Browning, M., Hansen, L.P. and Heckman, J.J. (2000). "Micro Data and General Equilibrium Models." (In J. B. Taylor and M. Woodford, eds, *Handbook of Macroeconomics* Vol 1a, Chapter 8).

[4] Diamond, P. (1998). "Optimal Income Taxation: an Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *The American Economic Review,* **88**, 83-95.

[5] Golosov, M., Kocherlakota, N. and Tsyvinski, A. (2003). "Optimal Indirect and Capital Taxation." *Review of Economic Studies*, 70, 569-588.

[6] Heckman, J.J. (1975). "Estimates of a Human Capital Production Function Embedded in a Life-Cycle Model of Labor Supply." (In N. Terlekyj, ed., *NBER Studies in Income and Wealth*, **40**, New York, Columbia University Press).

[7] Heckman, J.J., Lochner, L. and Taber, Ch. (1998). "Explaining rising wage inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents." *Review of Economic Dynamics* **1**, 1-58.

[8] Heckman, J.J., Lochner, L. and Taber, Ch. (1998). "Tax Policy and Human-Capital Formation." *The American Economic Review*, **88**, Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association, 293-297.

[9] Jorgensen, M. and Reed, W.J. (2004). "The Double Pareto-Lognormal Distribution - A New Parametric Model for Size Distributions." *Com. Stats -Theory & Methods*, **33**.

[10] Kapicka, M. (2004). "Optimal Taxation and Human Capital Accumulation." UCSB, unpublished manuscript.

[11] Kotlikoff, L., Smetters, K. and Walliser, J. (2001). "Finding a Way Out of America's Demographic Dilemma." Boston University, unpublished manuscript.

[12] Kocherlakota, N. (2005). "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation." *Econometrica*, 73, 1587-1621.

[13] Kocherlakota, N. (2004). "Figuring out the impact of hidden savings on optimal unemployment insurance." *Review of Economic Dynamics*, **7**, 541-554.

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

[15] Mirrlees, J.A. (1976). "Optimal Tax Theory: A Synthesis." *Journal of Public Economics*, **6,** 327-358.

[16] Mirrlees, J.A. (1986). "The Theory of Optimal Taxation." (In K.J. Arrow and M.D. Intriligator, eds., *Handbook of Mathematical Economics*, vol. III, chapter 24).

[17] Saez, E. (2001). "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies,* **68**, 205-229.

[18] Trostel, P A. (1993). "The Effect of Taxation on Human Capital." *Journal of Political Economy*, **101**, 327-350.

[19] Werning, I. (2001). "Dynamic Optimal Taxation." MIT, unpublished manuscript.

7 Appendix

Proof of Proposition 2. I will omit the proof that (4) is necessary. Applying the envelope theorem on the incentive compatibility constraint on (1) gives

$$V(\theta,\theta) = \frac{U[u(\theta)]}{(1-\beta)} = \int_0^\theta \sum_0^\infty \beta^t U'(u(\varepsilon)) v_t'(\varepsilon) l_t(\varepsilon) \frac{d\varepsilon}{\varepsilon} + V_0. \tag{12}$$

where $v_t'(\varepsilon) = v'(l_t(\varepsilon) + g(h_t(\varepsilon), h_{t+1}(\varepsilon))$. I will also skip the proof that the envelope theorem applies. It is a fairly straightforward modification of the proof that can be found in ([10]) or in Mirrlees ([16]).

Rewriting (12), I get that

$$u(\theta) = U^{-1}[(1-\beta)\int_0^\theta \sum_0^\infty \beta^t U'(u(\varepsilon)) v'(l_t(\varepsilon) + s_t(\varepsilon)) l_t(\varepsilon) \frac{d\varepsilon}{\varepsilon} + U_0].$$

where $U_0 = (1-\beta)V_0$. I will use the fact that, for any differentiable function $\varphi(\theta)$ one has

$$U^{-1}(\varphi(\theta)) = \int_0^\theta U^{-1'}(\varphi(\varepsilon))\varphi'(\varepsilon)d\varepsilon + U^{-1}(\varphi(0)) = \int_0^\theta \frac{\varphi'(\varepsilon)}{U'[U^{-1}(\varphi(\varepsilon))]} d\varepsilon + U^{-1}(\varphi(0))$$

I apply this formula for $\varphi(\theta) = (1-\beta)\sum_0^\infty \beta^t \int_0^\theta U'(u(\varepsilon))v'(l_t(\varepsilon) + s_t(\varepsilon))l_t(\varepsilon)\frac{d\varepsilon}{\varepsilon} + U_0$ which is clearly differentiable in θ. Note that $U'[U^{-1}(\varphi(\theta))] = U'(u(\theta))$ and so both terms involving $U'(u(\theta))$ cancel out. What remains is equation (3), where $u_0 = U^{-1}(\varphi(0)) = U^{-1}(U_0)$. ■

Proof of Theorem 4. Necessity of the conditions is obvious. For sufficiency, let $\sigma^* = \{u^*(\theta), l_t^*(\theta), h_{t+1}^*(\theta)\}$, together with $\hat{\lambda}^*$ and $\mu^* = \lambda^* - U'(u^*)$ be an allocation that satisfies the conditions of the theorem. Let $\sigma = \{u_t(\theta), l_t(\theta), h_{t+1}(\theta)\}$ be an alternative allocation that solves 2. I will prove sufficiency if I show that allocation σ^* delivers expected utility at least as high as σ.

Since U is increasing and concave, I have $U(u^*) - U(u) \geq U'(u^*)(u^* - u)$. Hence

$$\int_\Theta [U(u^*) - U(u)] f d\theta \geq \int_\Theta U'(u^*)(u^* - u) f d\theta$$

and so it will suffice to show that the right hand side of the inequality is positive.

To show this, substract first the two corresponding envelope conditions. I have

$$u(\theta) - u^*(\theta) = u(0) - u^*(0) + (1-\beta)\int_0^\theta \sum_{t\geq 0}^\infty \beta^t (v_t' \frac{l_t}{\varepsilon} - v_t'^* \frac{l_t^*}{\varepsilon}) d\varepsilon.$$

Multiply both sides by $\mu^*(\theta)f(\theta)$ and integrate over Θ :

$$\int_\Theta \mu^*(u-u^*)fd\theta = \int_\Theta \mu^* f \int_{\underline{\theta}}^\theta \sum_{t\geq 0}^\infty \beta^t (v_t' \frac{l_t}{\varepsilon} - v_t'^* \frac{l_t^*}{\varepsilon}) d\varepsilon$$

since $[U(u(0)) - U(u^*(0))] \int_\Theta \mu^* f d\theta = 0$. Reversing the order of integration, I get

$$\begin{aligned}
\int_\Theta \mu^*(u^*-u)fd\theta &= (1-\beta)\lambda^* \int_\Theta X^*(\theta)\{\sum_{t\geq 0}^\infty \beta^t (v_t'^* \frac{l_t^*}{f\theta} - \frac{v_t' l_t}{f\theta})\} f d\theta \\
&= (1-\beta)\lambda^* \int_\Theta \sum_{t\geq 0}^\infty \beta^t \{[\theta h_t l_t - v_t - v_t' l_t \frac{X^*(\theta)}{f\theta}] - [\theta h_t^* l_t^* - v_t^* - v_t'^* l_t^* \frac{X^*(\theta)}{f\theta}]\} f d\theta \\
&+ (1-\beta)\lambda^* \int_\Theta \sum_{t\geq 0}^\infty \beta^t [\theta h_t^* l_t^* - v_t^* - (\theta h_t l_t - v_t)] f d\theta \\
&\leq \lambda^* \int_\Theta (u^*-u) f d\theta.
\end{aligned}$$

The inequality follows from two facts. First, $\{l_t^*(\theta), h_{t+1}^*(\theta)\}$ maximizes individual's distorted problem and so the first expression on the right hand side of the second equality is nonpositive. Second, the resource constraint holds with equality and so the second expression on the right hand side of the second equality is equal to $\lambda^* \int_\Theta (u^*-u)fd\theta$. I will now use the fact that $\mu^* = \lambda^* - U'(u^*)$ and so

$$\int_\Theta [\lambda^* - U'(u^*)](u^*-u)fd\theta \leq \lambda^* \int_0^{\bar{\theta}} (u^*-u)fd\theta.$$

Upon cancelling terms and rearranging the terms, I get that

$$\int_\Theta U'(u^*)(u^*-u)fd\theta \geq 0$$

which completes the proof. ■

Proof of Theorem 5. I will assume that human capital is exogenous and identical for everyone. The extension for the case when human capital varies across people can be treated in a similar way.

If an allocation is incentive compatible then it satisfies (1). Taking l_t as given, define a consumption sequence $\tilde{c}_t$ by

Since

$$\tilde{c}_t(\theta) - v(1 - \frac{y_t(\theta)}{\theta h_t} - S) \geq \tilde{c}_t(\hat{\theta}) - v(1 - \frac{y_t(\hat{\theta})}{\theta h_t} - S) \tag{13}$$

If a pair of consumption $\tilde{c}_t(\theta)$ and $y_t(\theta)$ satisfies this condition, I will say it is incentive compatible in period t. There is at least one consumption function $\tilde{c}_t(\theta)$ that satisfies (13) and so is well defined. One way to show it is to use the fact, that if $\tilde{c}_t(\theta)$ and $y_t(\theta)$ is period t incentive compatible, it must satisfy the envelope theorem

$$\tilde{c}_t(\theta) = \int_0^\theta v_t' l_t \frac{d\varepsilon}{\varepsilon} + v(1 - l_t(\theta) - S(h_t, h_{t+1})) + u_0.$$

for some u_0. I will now normalize the consumption sequence in such a way that $\sum \beta^t \tilde{c}_t = \sum \beta^t c_t$. This can be done by increasing or decreasing $\tilde{c}_t(\theta)$ by a constant amount for all t.

The allocation $\{\tilde{c}_t, y_t\}_{t \geq 0}$ is thus incentive compatible in all periods. It is therefore incentive compatible. It also has the same present value as $\{c_t, y_t\}_{t \geq 0}$ and is therefore incentive feasible. ■

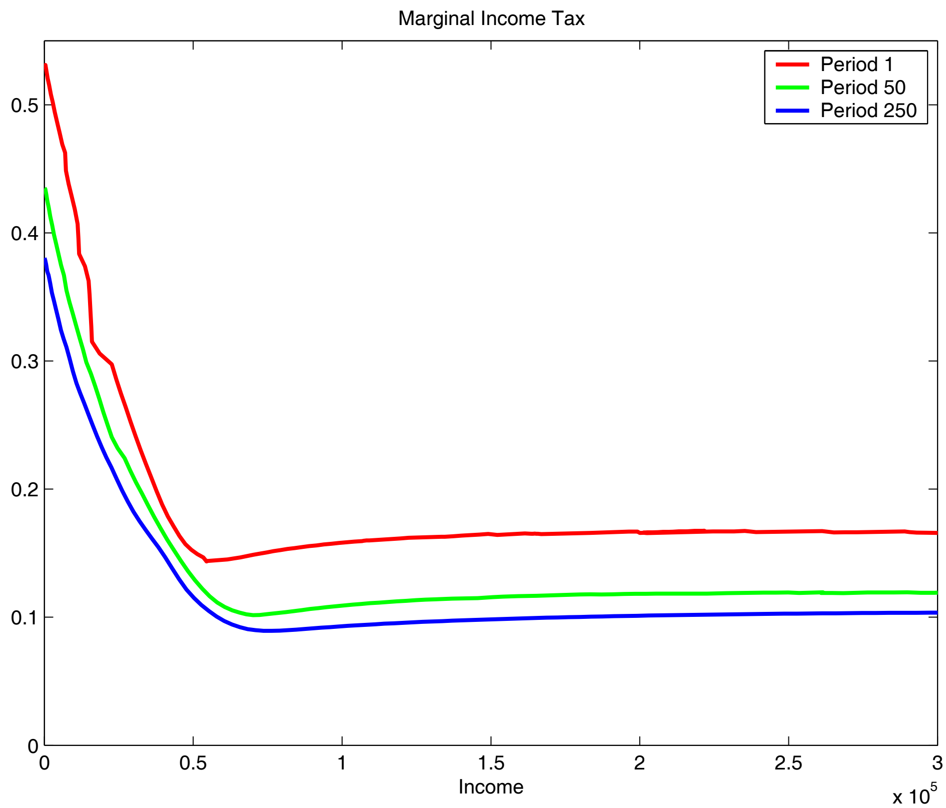

Figure 1: Marginal Income Taxes

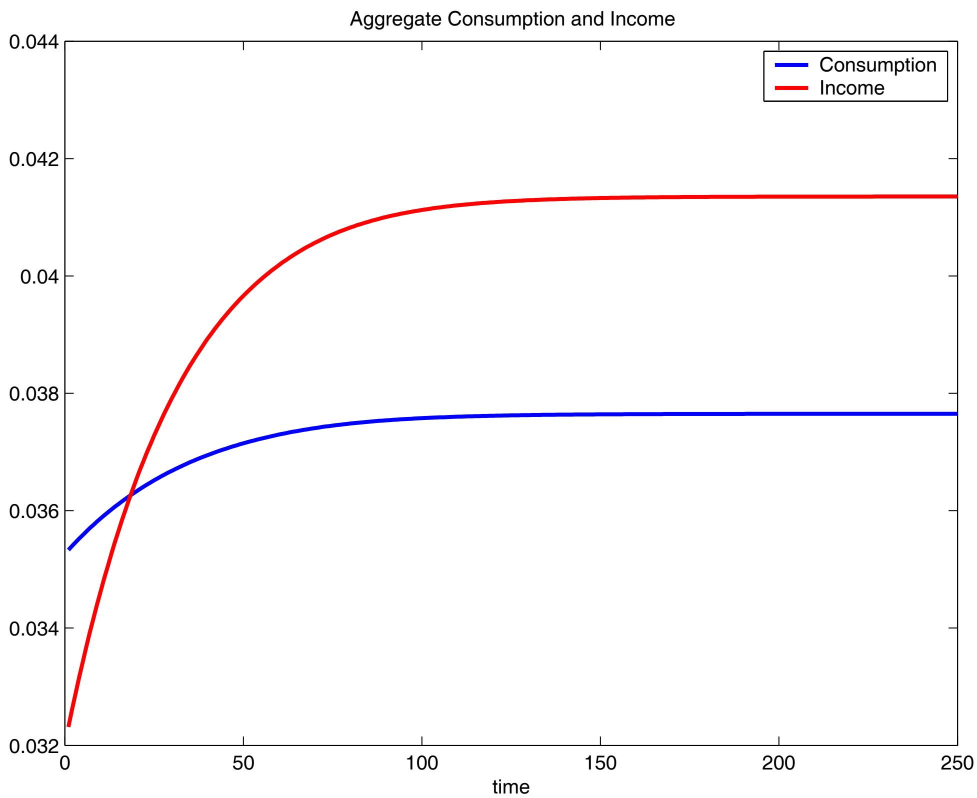

Figure 2: Aggregate Consumption and Income