

# On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System\*

Martin Gervais  
The University of Western Ontario  
gervais@uwo.ca

First Draft: March 2001

This draft: April 29, 2003

## Abstract

In life-cycle economies, where an individual's optimal consumption-work plan is almost never constant, the optimal marginal tax rates on capital and labor income vary with age. Conversely, the U.S. tax code implies marginal tax rates that vary with age because tax rates vary with earnings and earnings vary with age. A comparison of the optimal tax rates derived from the life-cycle model to those faced by an average individual in the U.S. economy indicates that whereas the age-profile of the labor income tax implied by the U.S. tax code is close to the optimal profile, the same cannot be said about the age-profile of the capital income tax. Nevertheless, if the tax authority is prevented from conditioning tax rates on age, a small degree of progressivity is desirable as progressive taxation better imitates optimal age-dependent taxes than an optimal age-independent tax system.

*Journal of Economic Literature* Classification Numbers: E62; H21

*keywords:* Optimal taxation; Progressive taxation; Life-cycle

---

\*I would like to thank Daniel Feenberg from the NBER for his help with TAXSIM from which tax rates from the data were computed, as well as Huberto Ennis, Andreas Hornstein, Igor Livshits, Miguel Molico, and seminar participants at various universities and the 2002 SED in New York for helpful comments.

# 1 Introduction

Since the seminal work of Mirrlees (1971), the trade off between equity and efficiency of progressive tax systems has received considerable attention—see Boadway (2000) for a review. On the one hand, progressive taxation is thought to give rise to more equitable allocations, but it does so at the cost of distorting the labor supply decision. Other authors have cast the trade off in terms of the implicit insurance provided by progressive taxation relative to its distortionary impact on labor/leisure and savings decisions—see Conesa and Krueger (2002). In this paper, I argue that progressive taxation may have a role purely on efficiency grounds, without relying on any redistribution arguments: A certain degree of progressivity in the tax system implies tax rates that are closer to optimal age-dependent tax rates. Thus, in a world in which the fiscal authority cannot condition tax rates on age, as is the case in the U.S., the optimal tax code involves progressive taxation.

The literature concerned with optimal dynamic fiscal policy has focused mainly on variants of the standard neoclassical growth model populated by infinitely-lived individuals.<sup>1</sup> In this context, Chamley (1986), Judd (1985), and others establish that an optimal income-tax policy entails taxing capital at confiscatory rates in the short-run and setting capital income taxes equal to zero in the long-run. Only labor income, if anything, should be taxed in the long run. The prescriptions from this literature are in sharp contrast with the observed tax codes of most countries. In the U.S. for instance, capital is taxed at non-trivial rates and the tax system is progressive. This paper investigates whether these observed features of actual tax codes can be rationalized by looking at dynamic fiscal policies in a life-cycle economy.<sup>2</sup>

In life-cycle economies both capital and labor income taxes are generally used by an optimizing government, even in the long run. When the government has access to a full set of proportional, age-conditioned, tax rates on capital and labor income, the

---

<sup>1</sup>Exceptions include Atkinson and Sandmo (1980), Escolano (1991), Ambler (2000) and Erosa and Gervais (2002).

<sup>2</sup>Chari and Kehoe (1999) review the literature on optimal taxation. Atkeson, Chari and Kehoe (1999) provide an introduction to the literature concerned with optimal taxation in infinitely-lived agent models, while Erosa and Gervais (2001) discuss both the infinitely-lived agent model and the life-cycle model.

optimal tax rates vary over the lifetime of individuals, that is, the optimal (marginal) tax rates are age-dependent. Under a Cobb-Douglas utility function, the shape of the optimal tax profiles on capital and labor income is determined by the labor supply profile chosen by individuals.<sup>3</sup> For instance, when individuals choose a hump-shaped labor supply profile, the optimal tax rate on capital income is negative until the labor supply peaks, becomes positive until individuals retire, and is zero thereafter. As for the labor income tax profile, it is hump-shaped and peaks at the same age as the labor supply. Conversely, because of the progressivity of the U.S. tax system, the marginal tax rates that individuals face vary with earnings. Since earnings vary over the lifetime of individuals, a progressive tax system implies that the marginal tax rates faced by the average U.S. tax payer also vary with age. The question that arises is how close the optimal marginal tax rates derived from a life-cycle model are to those faced by an average individual in the U.S. economy.

Given optimal tax rates generated by a parameterized version of the model, I evaluate how closely the U.S. tax code approximates the optimal one. To do so, the NBER TAXSIM model is used to calculate the implied marginal tax rates faced by individuals in the 1995 Current Population Survey.<sup>4</sup> The marginal income tax profiles differ from the optimal ones in two important ways. First and foremost, capital income is taxed too heavily, especially at young ages and during retirement. Second, although the labor income tax profile shares the hump shape of the Ramsey tax profile, the peak occurs later in the data than in the model. The second discrepancy occurs because optimal labor income tax rates are a function of the labor supply whereas the U.S. progressive tax system taxes labor income based on earnings, and earnings peak later in life than labor supply.

The above results suggest, at least as far as labor income taxes are concerned, that progressive taxation may provide a way for the government to implement optimal age-dependent taxes without observing individuals' age. As is well known, however, progressive taxation introduces a wedge between marginal and average tax rates that

---

<sup>3</sup>Note, however, that the labor supply profile that individuals choose is a function of the tax rates themselves.

<sup>4</sup>The data is available at <http://www.nber.org/~taxsim/byage/>. For more information about the TAXSIM model, see Feenberg and Coutts (1993) or <http://www.nber.org/~taxsim>.

is not present under age-dependent taxes. I evaluate the cost of this wedge by computing equilibria of the life-cycle model under different specifications of progressive tax systems.

An alternative to progressive taxation when the fiscal authority cannot tax individuals based on their age is of course a pure proportional tax system where individuals face age-independent tax rates on capital and labor income. In order to compare allocations under progressive tax codes to the allocation that obtains under an optimal age-independent tax code, an appropriate Ramsey problem needs to be formulated. And since I want to rule out re-distributional effects, this Ramsey problem must impose that the government balances its budget on a period by period basis.<sup>5</sup> In a formulation of the Ramsey problem in which the government chooses allocations rather than prices, this constraint essentially imposes that the capital stock must exactly match the savings of the entire population. Two additional constraints, which involve equalizing marginal rates of substitutions across individuals of all ages, are necessary for allocations to be implementable with an age-independent tax code. Allocations and welfare under progressive tax system are then compared to those obtained under optimal flat proportional tax rates.

Simulation results are as follows. First, the welfare cost of a progressive tax system where the tax base is total income is substantial. For an individual to be indifferent between the allocation under a progressive tax system meant to replicate the U.S. tax code and the allocation under optimal proportional tax rates, he/she would require consumption to be increased by over 6 percent in every period of his/her life. This result, however, is very sensitive to the degree of progressivity of the tax code. For low enough degrees of progressivity, the progressive tax system is preferred to the proportional tax system. Second, under a tax system where labor income is taxed at progressive rates and capital income is taxed at a constant proportional rate, a progressive tax system is preferred to constant proportional taxes on capital and labor even for relatively high degrees of progressivity.

The rest of the paper is organized as follows. The next section presents the eco-

---

<sup>5</sup>Stockman (2001) studies optimal taxation without government debt in a stochastic neoclassical growth model populated a representative infinitely-lived individual.

conomic environment, formulates a Ramsey problem, and characterizes optimal fiscal policies. Section 3 compares implied tax rates from the data to the optimal ones generated by a parameterized version of the model. The implications of progressive taxation are studied in section 5, where allocations are compared to allocations derived from a Ramsey problem which imposes zero government debt and age-independent tax rates, as formulated in section 4. Section 6 concludes the paper.

## 2 Economic Environment and Ramsey Taxes

Consider an economy populated by overlapping generations of individuals with finite lives.<sup>6</sup> Individuals make consumption and labor/leisure choices in each period so as to maximize their lifetime utility. Firms operate a neoclassical production technology: factors are paid their marginal products. The payments received by individuals on their factors (capital and labor) are subject to proportional taxes which can be conditioned on age. The government uses the revenues from taxation to finance an exogenously given stream of government purchases. In addition, the government absorbs any imbalance between tax revenues and public expenditures by issuing debt. Note that for any given fiscal policy, individual behavior (by consumers and firms) implies a particular allocation. The Ramsey problem consists of choosing, among all those allocations, the one that maximizes a particular utilitarian welfare function. This problem, the Ramsey problem, will be formally defined once the basic economic environment is introduced.

### 2.1 Economic Environment

**Households** Individuals live  $(J + 1)$  periods, from age 0 to age  $J$ . At each time period a new generation is born and is indexed by its date of birth. At date 0, when the change in fiscal policy occurs, the generations alive are  $-J, -J + 1, \dots, 0$ . To take these initial generations into account in what follows, it will prove convenient to denote the age of individuals alive at date zero by  $j_0(t)$ . For all other generations,

---

<sup>6</sup>This section draws from Erosa and Gervais (2002).

$j_0(t) = 0$ , so that for any generation  $t$ ,  $j_0(t) = \max\{-t, 0\}$ . One can thus think of  $j_0(t)$  as the first period of an individual's life which is affected by the date zero switch in fiscal policy. The population is assumed to grow at constant rate  $n$  per period, and  $\mu_j$  represents the share of age- $j$  individuals in the population. The labor productivity level of an age- $j$  individual is denoted  $z_j$ .

Let  $c_{t,j}$  and  $l_{t,j}$ , respectively, denote consumption and time devoted to work by an age- $j$  individual who was born in period  $t$ . Note that  $c_{t,j}$  and  $l_{t,j}$  actually occurs in period  $(t+j)$ . Similarly, the after-tax prices of labor and capital services are denoted  $w_{t,j}$  and  $r_{t,j}$ , respectively. Given a fiscal policy  $\pi$ , the problem faced by an individual born in period  $t \geq -J$  is to maximize lifetime utility subject to a sequence of budget constraints:

$$U^t(\pi) \equiv \max \sum_{j=j_0(t)}^J \beta^{j-j_0(t)} U(c_{t,j}, 1 - l_{t,j}), \quad (1)$$

$$\text{s.t. } c_{t,j} + a_{t,j+1} = w_{t,j} z_j l_{t,j} + (1 + r_{t,j}) a_{t,j}, \quad j = j_0(t), \dots, J, \quad (2)$$

where  $\beta > 0$  is a discount factor and  $a_{t,j}$  denotes total asset holdings by an age- $j$  individual who was born at date  $t$ . Initial asset holdings,  $a_{t,j_0(t)}$ , are taken as given for initial generations and are equal to zero for all other generations. The utility function  $U$  is assumed to be strictly increasing in both arguments, strictly concave, and satisfies standard Inada conditions. On the left hand side of equation (1),  $U^t(\pi)$  denotes the indirect utility function of a generation  $t$  individual, that is, the maximum lifetime utility an individual obtains under fiscal policy  $\pi$ . The budget constraint (2) expresses that individuals allocate their income, composed of labor and (gross) interest income, to consumption and saving.

Let  $p_{t,j}$  denote the Lagrange multiplier associated with the budget constraint (2) faced by an age- $j$  individual born in period  $t$ . The necessary and sufficient conditions for a solution to the consumer's problem are given by (2) and

$$\beta^{j-j_0(t)} U_{c_{t,j}} - p_{t,j} = 0, \quad (3)$$

$$\beta^{j-j_0(t)} U_{l_{t,j}} + p_{t,j} w_{t,j} z_j \leq 0, \quad \text{with equality if } l_{t,j} > 0, \quad (4)$$

$$-p_{t,j} + p_{t,j+1} (1 + r_{t,j+1}) = 0, \quad (5)$$

$$a_{t,J+1} = 0, \quad (6)$$

$j = j_0(t), \dots, J$ , where  $U_{c_{t,j}}$  and  $U_{l_{t,j}}$  denote the derivative of  $U$  with respect to  $c_{t,j}$  and  $l_{t,j}$  respectively.<sup>7</sup>

In principle, one could solve the above equations to obtain the optimal consumption and leisure profiles of individuals. Naturally, these demand functions would depend on the fiscal policy chosen by the government, and would represent the reaction functions that the government takes into account when choosing tax rates. Alternatively, these first order conditions can be used to substitute prices out of the budget constraint (2). The resulting equations, which only involve quantities, can then be imposed as constraints on an alternative formulation of the Ramsey problem where the government chooses allocations rather than tax rates.

**Technology and Feasibility** The production technology is represented by a neo-classical production function with constant returns to scale,  $q_t = f(k_t, l_t)$ , where  $q_t$ ,  $k_t$  and  $l_t$  denote the aggregate (per capita) levels of output, capital, and effective labor, respectively. Capital and labor services are paid their marginal products: before-tax prices of capital and labor in period  $t$  are given by  $\hat{r}_t = f_k(k_t, l_t) - \delta$ , where  $0 < \delta < 1$  is the depreciation rate of capital, and  $\hat{w}_t = f_l(k_t, l_t)$ .

Feasibility requires that total (private and public) consumption plus investment be less than or equal to aggregate output

$$c_t + (1 + n)k_{t+1} - (1 - \delta)k_t + g_t \leq q_t, \quad (7)$$

where  $c_t$  denotes aggregate private consumption at date  $t$ ,  $g_t$  stands for date- $t$  government consumption, and all aggregate quantities are expressed in per capita terms. Note that tomorrow's per capita capital stock needs to be multiplied by  $(1 + n)$  to account for population growth. Also, the date- $t$  aggregate levels of consumption and labor input, the latter being expressed in efficiency units, are obtained by adding up the weighted consumption (or effective labor supply) of all individuals alive at date  $t$ , where the weights are given by the fraction of the population that each individual

---

<sup>7</sup>The Inada conditions guarantee that consumption and leisure will be strictly positive in each period.

represents:

$$c_t = \sum_{j=0}^J \mu_j c_{t-j,j},$$

$$l_t = \sum_{j=0}^J \mu_j z_j l_{t-j,j}.$$

**The Government** To finance a given stream of government expenditures, it is assumed that the government has access to a set of fiscal policy instruments and a commitment technology to implement its fiscal policy. The set of instruments available to the government consists of government debt and proportional, age-dependent taxes on labor income and capital income. The date- $t$  tax rates on capital and labor services supplied by an age- $j$  individual (born in period  $(t-j)$ ) are denoted by  $\tau_{t-j,j}^k$  and  $\tau_{t-j,j}^w$ , respectively. In per capita terms, the government budget constraint at date  $t \geq 0$  is given by

$$(1 + \hat{r}_t)b_t + g_t =$$

$$(1 + n)b_{t+1} + \sum_{j=0}^J (\hat{r}_t - r_{t-j,j})\mu_j a_{t-j,j} + \sum_{j=0}^J (\hat{w}_t - w_{t-j,j})\mu_j z_j l_{t-j,j}, \quad (8)$$

where  $b_t$  represents government debt issued at date  $t$ ,  $w_{t,j} \equiv (1 - \tau_{t,j}^w)\hat{w}_{t+j}$ , and  $r_{t,j} \equiv (1 - \tau_{t,j}^k)\hat{r}_{t+j}$ . Equation (8) expresses that the government pays its expenditures, composed of outstanding government debt payments (principal plus interest) and other government outlays, either by issuing new debt (adjusted for population growth), by taxing interest income, or by taxing wage income.

In the spirit of Ramsey (1927), the government takes individuals' optimizing behavior as given and chooses a fiscal policy to maximize social welfare, where social welfare is defined as the discounted sum of individual lifetime welfares (Samuelson (1968) and Atkinson and Sandmo (1980)). In other words, the government chooses a sequence of tax rates in order to maximize

$$\sum_{t=-J}^{\infty} \gamma^t U^t(\pi), \quad (9)$$

where  $0 < \gamma < 1$  is the intergenerational discount factor and  $U^t(\pi)$  was defined earlier as the indirect utility function of generation  $t$  as a function of the government tax policy.

## 2.2 The Ramsey Problem

The Ramsey problem consists of choosing a set of tax rates so that the resulting allocation, when prices and quantities are determined in competitive markets, maximizes a given welfare function. Alternatively, a Ramsey problem where the government chooses allocations rather than tax rates can be formulated.<sup>8</sup> This is done by constructing a sequence of implementability constraints which guarantee that any allocation chosen by the government can be decentralized as a competitive equilibrium. The implementability constraints are obtained by using the consumers' optimality conditions (3)–(5) to substitute out prices from the consumer's budget constraints (2). After adding up these budget constraints, the resulting implementability constraint associated with the cohort born in period  $t$  is given by

$$\sum_{j=j_0(t)}^J \beta^{j-j_0(t)} (U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j}) = A_{t,j_0(t)}, \quad (10)$$

where  $A_{t,j_0(t)} = U_{c_{t,j_0(t)}}(1 + r_{t,j_0(t)})a_{t,j_0(t)}$ . It is important to note that these implementability constraints rely on the existence of age-dependent tax rates. Since factors are paid their marginal products, before-tax prices do not depend on age. It follows that after-tax prices can only depend on age if the government has access to age-dependent tax rates. Additional restrictions, which involve marginal rates of substitution over the lifetime of individuals, need to be imposed for an allocation to be implementable with age-independent taxes. In other words, the set of allocations from which the government can pick depends crucially on the instruments available to the government.

Since these implementability constraints are constructed from the optimality conditions of the consumers' problem, it is clear that any competitive equilibrium al-

---

<sup>8</sup>This is the formulation of the Ramsey problem generally used to study optimal taxation in infinitely-lived agent models. See Chari and Kehoe (1999) for a review.

location satisfies (10). Conversely, one can show that if an allocation satisfies the implementability constraints (10) as well as the feasibility constraint (7), then there exists a fiscal policy for which the allocation is a competitive equilibrium.<sup>9</sup> This is what allows us to set up a Ramsey problem in which the government chooses quantities rather than prices.

This Ramsey problem consists of choosing an allocation to maximize the discounted sum of successive generations' utility subject to each generation's implementability constraint as well as the feasibility constraint, that is,

$$\max_{\{c_{t,j}, l_{t,j}\}_{j=j_0(t), k_{t+J+1}}^J}_{t=-J} \sum_{t=-J}^{\infty} \gamma^t W_t \quad (11)$$

subject to feasibility (7) for  $t = 0, \dots$ . The function  $W_t$  is defined to include generation  $t$ 's implementability constraint in addition to generation  $t$ 's lifetime utility, where lifetime utility now refers to the direct utility function. Letting  $\gamma^t \lambda_t$  be the Lagrange multiplier associated with generation  $t$ 's implementability constraint (10), the function  $W_t$  is defined as

$$W_t = \sum_{j=j_0(t)}^J \beta^{j-j_0(t)} [U(c_{t,j}, 1 - l_{t,j}) + \lambda_t (U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j})] - \lambda_t A_{t,j_0(t)}. \quad (12)$$

It should be noted that since government debt is unconstrained, the government budget constraint (8) need not be imposed on the Ramsey problem. It can be shown that the government budget constraint holds if the implementability constraint (or the present value budget constraint of individuals) and the feasibility constraint are satisfied. Once a solution is found, the government budget constraint can be used to back out the level of government debt.

### 2.3 Optimal Tax Profiles

It will be shown in this section that if the Ramsey allocation converges to a steady state solution, optimal labor and capital income taxes will in general be different

---

<sup>9</sup>For details, see Atkeson, Chari and Kehoe (1999) or Erosa and Gervais (2002). A similar argument will be made formally in section 4 when I formulate a Ramsey problem without government debt.

from zero even in that steady state. Although the main results of this section hold more generally, attention will be restricted to steady states and to a particular class of utility functions.

Let  $\gamma^t \phi_t$  denote the Lagrange multiplier associated with the time- $t$  feasibility constraint (7). The steady state solution is characterized by the following equations:

$$1 - \delta + f_k = \frac{1 + n}{\gamma}, \quad (13)$$

$$(1 + \lambda)\beta^j U_{c_j} + \lambda\beta^j U_{c_j} H_j^c = \gamma^j \phi \mu_j, \quad j = 0, \dots, J, \quad (14)$$

$$(1 + \lambda)\beta^j U_{l_j} + \lambda\beta^j U_{l_j} H_j^l \leq -\gamma^j \phi \mu_j z_j f_l, \quad j = 0, \dots, J, \quad \text{with equality if } l_j > 0, \quad (15)$$

where

$$H_j^c = \frac{U_{c_j, c_j} c_j + U_{l_j, c_j} l_j}{U_{c_j}}, \quad (16)$$

$$H_j^l = \frac{U_{c_j, l_j} c_j + U_{l_j, l_j} l_j}{U_{l_j}}, \quad (17)$$

as well as the feasibility and implementability constraints (7) and (10). Atkeson, Chari and Kehoe (1999) refer to the terms  $H^c$  and  $H^l$  as *general equilibrium elasticities* since they capture the relevant distortions for setting the capital and labor income tax rates in general equilibrium.

The first order condition with respect to capital, equation (13), implies that the solution to this Ramsey problem has the modified golden rule property: the marginal product of capital (net of depreciation) equals the effective discount rate applied to successive generations  $[(1 + n)/\gamma] - 1$ . Equation (14), which corresponds to the first order condition with respect to consumption, expresses that the government equates the discounted marginal benefit of consumption to the marginal cost of producing extra consumption, which is given by the shadow value of production,  $\phi$ , appropriately weighted and discounted. Note that the marginal benefit includes the direct utility gain as well as the impact of a small change in consumption on the implementability constraint. Finally, the first order condition with respect to the labor decision (14) equates the benefit of an extra unit of leisure to the cost of having one less unit of

labor, which is given by the shadow value of production,  $\phi$ , appropriately weighted and discounted, times the marginal product of labor measured in efficiency units.

Erosa and Gervais (2002) show that capital and labor income will in general be taxed at non-zero rates in this environment. This follows from the fact that consumption and leisure are generally not constant over an individual's lifetime, even in steady state. Furthermore, the optimal tax rates will generally not be constant over an individual's life, as the general equilibrium elasticities  $H_j^l$  and  $H_j^c$  will change as individuals age.

In order to characterize the shape of the tax profiles over the lifetime of individuals, consider a utility function of the form

$$U(c_j, 1 - l_j) = \frac{c_j^{1-\sigma}(1 - l_j)^\eta}{1 - \sigma}, \quad (18)$$

where  $\eta = \theta(1 - \sigma)$ . Here,  $1/\sigma$  is the intertemporal elasticity of substitution, which measures the degree to which individuals are willing to substitute consumption over time, and  $\theta$  measures the intensity of leisure in individuals' preferences. The principles guiding the optimal manner in which to tax capital and labor over the lifetime of individuals are stated in the following Proposition.

**Proposition 1** *Assume that the utility function takes the form given by (18). Then (i) the tax rate on capital income at age  $j + 1$  is positive if and only if  $l_{j+1} < l_j$ , and (ii) the tax rate on labor income at age  $j$  is higher than at age  $j + 1$  if and only if  $l_{j+1} < l_j$ .*

**Proof.** See Appendix.

By taxing or subsidizing capital, the government makes consumption and leisure in the future more or less expensive than today. Proposition 1 suggests the government uses capital income taxes to smooth individuals' leisure and consumption profiles over their lifetime. Under Cobb-Douglas utility, the share of consumption in an individual's total expenditures is constant, so that consumption and leisure always move together over time. If consumption and leisure are high tomorrow relative to today, then the government will tax the return on today's savings at a positive rate

tomorrow. Doing so, the government gives individuals an incentive to consume more and to save less today, and thus to consume less tomorrow.

An implication of the principle of optimal taxation developed in Proposition 1 is that capital income should not be taxed during retirement. This follows directly from the fact that labor supply is constant during retirement. Notice, however, that leisure time during retirement is taxed indirectly by taxing the return on savings prior to retirement.

Proposition 1 puts a lot of structure on the profiles of optimal capital and labor income tax rates. Furthermore, the shape of these profiles is entirely determined by the profile of the labor supply. For instance, imagine that the labor supply profile is hump-shaped. Then the tax rate on capital income should be negative until the labor supply peaks, become positive until retirement, and be zero thereafter. Proposition 1 also tells us that under a hump-shaped labor supply profile the tax rate on labor income should be hump-shaped and peak at the same age as the labor supply.

### 3 Age-Dependent and Progressive Tax Profiles

To compare the prescribed tax rates to those implied by the data, the model is parameterized and optimal tax rates are computed numerically. After computing the optimal tax rates implied by the Ramsey problem, these rates are compared to implied tax rates on labor and capital income from the 1995 Current Population Survey data.

#### 3.1 Simulating Ramsey Taxes

It is assumed that individuals live for 55 years ( $J = 54$ ) and the population grows at one percent per annum ( $n = 0.01$ ). In this setting, one can think of individuals as beginning their economic life at age 21, which corresponds to model age 0, and living until real age 75. The labor productivity profile is taken from Hansen (1993) and normalized so that labor productivity is equal to one in the first year ( $z_0 = 1$ ).<sup>10</sup>

---

<sup>10</sup>I actually use a smoothed version of the profile. The equation generating the productivity profile is  $z_j = 0.4817 + 0.0679(j + 1) - 0.0013(j + 1)^2$  for  $j = 0, \dots, 54$ , which is then normalized so that

The utility function is specified as in equation (18) with intertemporal elasticity of substitution equal to 0.5 ( $\sigma = 2$ ), and discount rate equal to 1.5 percent per year ( $\beta = (1 + 0.015)^{-1}$ ). The parameter determining the intensity of leisure is set such that aggregate working time represents a third of total time ( $\theta = 1.47$ ).

The production function is given by  $f(k, l) = k^\alpha l^{1-\alpha}$ . The capital share of output is set to 36 percent ( $\alpha = 0.36$ ) and capital depreciates at a rate of 6.5 percent per year ( $\delta = 0.065$ ).

The level of (per capita) government spending is set so that it represents 19 percent of output. Simultaneously, the value of the intergenerational discount factor ( $\gamma$ ) is chosen to make government debt equal to zero in the final steady state.<sup>11</sup> The value of  $\gamma$  which accomplishes this goal is 0.948. Equation (13) then implies that the pre-tax interest rate is equal to 6.5 percent, and the steady state capital-output ratio is equal to 2.76.

Figure 1 illustrates how taxes vary with age under our parameterization of the model. Following Proposition 1, capital income taxes are positive (negative) when the labor supply is decreasing (increasing), and labor income taxes follow the shape of the labor supply.

## 3.2 Comparison with the Data

Because of the need to compute tax rates that individuals face at each age of their life, which is not available from the IRS, tax rates were imputed using the NBER TAXSIM model on data from the Current Population Survey (CPS) of the 1995 US Census.<sup>12</sup> Given the information available in the CPS data, the TAXSIM model calculates the tax liability that each individual in the sample faced under the 1995 U.S. tax code.<sup>13</sup>

---

$z_0 = 1$ .

<sup>11</sup>Note that different values of  $\gamma$  influence the level of the labor and capital tax profiles more than their shape: higher values of  $\gamma$  lead to lower government debt and lower tax rates in the long run.

<sup>12</sup>The data is available at <http://www.nber.org/~taxsim/byage/>. For more information about the TAXSIM model, see Feenberg and Coutts, (1993) or <http://www.nber.org/~taxsim>.

<sup>13</sup>It should be noted that the CPS is not an ideal source of property income data. In particular, it is assumed that all individuals are given only the standard deduction. This assumption may make the tax code appear more progressive than it is in reality and may paint a misleading picture for

For each individual, the marginal tax rate is calculated as the marginal rate from a one percent change in all income items, and the average tax rate is equal to the ratio of tax liability to adjusted gross income. The data used in the remainder of this paper consist of the mean marginal and average tax rates for individuals at each age.

Figure 2 shows the marginal tax profiles generated by the CPS data. Much like in the model, the total income tax rate follows the labor income tax rate until the last few years where it tracks the capital income tax.

The marginal tax rates generated by the data are now compared to those prescribed by the Ramsey problem under our benchmark parametrization. Figure 3 plots the total income tax rates from the model and the total (marginal) income tax rates from the data.<sup>14</sup> This Figure suggests that the U.S. tax code under-taxes young individuals and taxes middle-aged individuals too much. Although the total tax rate declines for older individuals, Ramsey taxes are zero for retirees, suggesting that older individuals are also over-taxed. This last point, however, may be an artifact of the representative agent model from which Ramsey taxes are generated, which implies that all individuals retire at exactly the same age.

The difference between the marginal tax rates generated by the data and those generated by the Ramsey problem can best be explained by analyzing labor and capital income taxes separately. Figure 4 suggests that capital income taxation is an important source of discrepancy between the two profiles. The most striking feature of that Figure is that the U.S. tax code generally taxes capital income at very high rates relative to the rates prescribed by the solution to the Ramsey problem. Note that the negative tax rate on capital income during the first few years is actually a tax on borrowing, as individuals have a negative net worth for the first 9 years. The fact that the tax rate on capital income implied by the data is positive at young ages suggests that individuals get a tax break while they have a negative net worth. Again, it should be emphasized that the CPS data is far from ideal for computing

---

capital income tax rates. I am currently computing tax rates using 2001 PSID data, for which the capital income tax rates can be much more reliably calculated.

<sup>14</sup>The total income tax rates from the model are actually average tax rates, computed as the ratio of tax liability and total income. They thus represent the tax rate on an extra dollar of income assuming that the distribution of income between capital and labor is unaffected by the extra income.

empirical capital income tax rates.

Unlike the capital income tax, the labor income tax rates implied by the data are fairly close to those prescribed by the Ramsey problem. Figure 5 shows that both profiles are hump-shaped at approximately the same level. The profile implied by the data, however, peaks much later than the Ramsey profile. The origin of this discrepancy is that the model taxes labor income according to a different principle from the U.S. tax code. In the model, as shown in Proposition 1, the optimal labor income tax is based on individuals' working time (or labor supply). In contrast, the labor income tax in the U.S. is a function of earnings, where earnings are equal to the wage rate times productivity times working time. Since actual marginal tax rates are an increasing function of earnings—by definition since the tax code is progressive—the tax profile in the data will peak at the same time as earnings. Figure 6 shows that the age-profile of productivity and hours worked do not peak at the same time, as the age-profile for productivity peaks later in life. That means that the tax profile from the data will naturally peak later than that of the model since in the model, the tax rate peaks at the same time as the labor supply.

Before exploring the desirability of using progressive taxation, we first need to establish a simple benchmark that the government can achieve when it cannot age-condition tax rates nor issue debt. In order to establish this benchmark in which capital and labor income are taxed at fixed rates over the lifetime of individuals, we need to formulate a Ramsey problem that can be used to compute optimal tax rates. This is subject of the next section.

## 4 Steady State Problem Without Debt

In this section I study a Ramsey problem in which the objective of the government or Planner is to maximize the steady state utility of a representative individual, under the restriction that the government cannot issue debt.<sup>15</sup> The idea is that since we

---

<sup>15</sup>Strictly speaking this problem is very different from the Ramsey problem discussed in section 2. Loosely speaking, however, the first order conditions of this steady state problem correspond to the steady state first order conditions of the full Ramsey problem when  $\gamma = 1$ . Of course, the objective

want to compare steady states where government debt is zero, we need a way of finding the best the government can do under a proportional tax system, which is a feasible alternative to progressive taxation when the government is not allowed to base tax rates on age.

**Definition 1** *A steady state allocation  $\{\{c_j, l_j\}_{j=0}^J, k\}$  is implementable if there exists a fiscal policy  $\pi \in \Pi$  and a sequence of asset holdings  $\{a_{j+1}\}_{j=0}^J$  such that:*

1. *Given prices implied by the fiscal policy  $\pi$ , the allocation  $\{c_j, l_j, a_{j+1}\}_{j=0}^J$  solves the following consumer problem:*

$$\max \sum_{j=0}^J \beta^j U(c_j, l_j) \quad (19)$$

$$\text{s.t. } c_j + a_{j+1} = w_j z_j l_j + (1 + r_j) a_j, \quad j = 0, 1, \dots, J, \quad (20)$$

where  $a_0 = 0$ ;

2. *Factor prices are competitive:*

$$\hat{r} = f_k - \delta, \quad (21)$$

$$\hat{w} = f_l; \quad (22)$$

3. *The government budget constraint is satisfied:*

$$g = \sum_{j=0}^J (\hat{w} - w_j \mu_j z_j l_j) + \sum_{j=0}^J (\hat{r} - r_j \mu_j a_j); \quad (23)$$

4. *The allocation is feasible:*

$$\sum_{j=0}^J \mu_j c_j + (n + \delta)k + g = f(k, l). \quad (24)$$

As the next two Propositions show, the set of implementable allocations depends crucially on the set of fiscal policies  $\Pi$  from which the government can choose. In function of the full Ramsey problem (11) is not well-defined when the government does not discount the future.

Proposition 2, the set of fiscal instruments is not restricted other than by the restriction in Definition 1 that the government balance its budget on a period by period basis. Proposition 3 restricts the set of feasible fiscal policies to only two proportional tax rates:  $\Pi = \{\tau^k, \tau^w\}$ , thus forcing the government to choose age-independent tax rates in addition to balancing the budget.

**Proposition 2** *An allocation  $\{\{c_j, l_j\}_{j=0}^J, k\}$  is implementable if and only if it is feasible (24) and satisfies the following implementability constraints:*

$$\sum_{j=0}^J \beta^j (U_{c_j} c_j + U_{l_j} l_j) = 0, \quad (25)$$

$$\sum_{m=1}^J -\frac{\mu_m}{\beta^m U_{c_m}} \sum_{j=0}^m \beta^j (U_{c_j} c_j + U_{l_j} l_j) = k. \quad (26)$$

**Proof.** See Appendix.

The first implementability constraint insures that the allocation satisfies the consumer's optimality conditions. The second constraint imposes that the weighted sum of all assets held by individuals at any point in time equals the aggregate stock of capital, thereby imposing government debt to equal zero.

**Proposition 3** *An allocation  $\{\{c_j, l_j\}_{j=0}^J, k\}$  is implementable with age-independent taxes if and only if it is feasible (24), satisfies the implementability constraints (25)–(26), as well as the following additional implementability constraints:*

$$l_j \left( \frac{U_{l_j}}{z_j U_{c_j}} - \frac{U_{l_0}}{z_0 U_{c_0}} \right) = 0 \quad j = 1, \dots, J, \quad (27)$$

$$\left( \frac{U_{c_j}}{U_{c_{j+1}}} - \frac{U_{c_0}}{U_{c_1}} \right) = 0 \quad j = 1, \dots, J - 1. \quad (28)$$

**Proof.** See Appendix.

The interpretation of the additional implementability constraints is straightforward. Given any tax rates, individuals will set marginal rates of substitution equal to after-tax prices. If we want to restrict the government to choose age-independent

tax rates, we must restrict the set of allocations it can choose to those that feature equal marginal rates of substitution across individuals of all ages. Only then can an allocation be decentralized with age-independent tax rates. The first constraint is multiplied by  $l_j$  since the constraint need not hold when individuals are retired ( $l_j = 0$ ).

The Ramsey problem then consists of maximizing steady state utility (19) subject to the feasibility constraint (24) as well as the appropriate implementability constraints: equations (25)–(26) for age-dependent taxes; and equations (25)–(28) for age-independent taxes. It is interesting to note that the optimality condition for the choice of capital implies that the allocation no longer has the modified golden rule property. When the government has access to government debt, it can issue debt or hold capital (negative debt) to make up the difference between the aggregate amount of assets individuals wish to hold and the golden rule amount. Without government debt, the government loses its ability to do so. Under the parameter values of the previous section, with  $\gamma$  close to unity, the government holds a lot of capital in order to achieve the golden rule amount of aggregate capital. It follows that the level of capital will be lower under budget balance than it was in that steady state.

## 5 Simulating Progressive Taxes

The discussion in section 3 suggests that, at least as far as the labor income tax is concerned, progressive taxation can generate marginal tax profiles that resemble optimal Ramsey tax profiles. Instead of comparing implied tax rates, this section compares the behavior of individuals under different specifications of progressive tax systems to their behavior under optimal Ramsey taxes. As is well known, even if progressive taxation could perfectly imitate optimal age-dependent marginal tax rates, the progressivity of the tax code implies additional distortions as it introduces a wedge between the average and marginal tax rates that is not present under an age-dependent tax system.

## 5.1 Average and Marginal Tax Functions

To specify a functional form for the tax function, I use the tax rates implied by the CPS data discussed above. Figure 7 depicts the average and marginal tax rates as a function of income.<sup>16</sup> Although many different functional forms could be used to approximate the progressivity of the U.S. tax code, the fact that the marginal tax rates appear to sit right on top of the average tax rates suggests that a log-linear function is an appropriate functional form to use. These functions are also shown on Figure 7.<sup>17</sup>

The fact that the intercepts of the marginal and average tax functions are very similar suggests that the tax code does not become more or less progressive as income increases, and that the distance between the two tax functions is a good proxy for the degree of progressivity in the tax code. However, the functional form used for the average and marginal tax functions are just approximations of the true functions, which means they need not be consistent with each other. More precisely, if we respectively let the average and marginal tax functions be given by

$$\begin{aligned}\bar{\tau}(y) &= \pi_0 + \pi_1 \log y, \\ \tau(y) &= (\pi_0 + \rho\pi_1) + \pi_1 \log y,\end{aligned}$$

then for these tax functions to be consistent with each other requires setting  $\rho$  equal to one. The data suggests that the tax code is more progressive than the log-linear tax function, that is,  $\rho > 1$ . Values of  $\rho \in (0, 1)$  imply a tax code that is less progressive than the log-linear case, and setting  $\rho$  equal to zero brings the tax code closer to the age-dependent tax code, which is not progressive.

---

<sup>16</sup>At this time, I only have average tax rates for total income, so I will use the tax functions from total income throughout the rest of the paper.

<sup>17</sup>Other functional forms, such as that estimated by Gouveria and Strauss (1994)—which has been used by Sarte (1997) and Castañeda, Días-Giménez and Ríos-Rull (2000)—or a quadratic form, do not fit the data as well. Note, however, that the log-linear tax function does not have some of the nice properties that the Gouveria and Strauss tax function has. In particular, the tax rate at zero income is not zero, and the distance between the average and marginal tax functions does not converge to zero as income gets large.

## 5.2 Simulations

I can now solve for optimal proportional tax rates and compare the allocation implied by this tax code to those obtained under different specifications of progressive tax systems.

**Age-Independent Optimal Taxation** In order to simulate optimal age-independent taxes, we need to solve the problem implied by Proposition 3, that is, maximizing steady state utility (19) subject to the feasibility constraint (24) as well as implementability constraints (25)–(28). I do so using the same parameter values as in section 3.1, except for government spending and the leisure preference parameter ( $\theta$ ). I set these parameters such that the resulting equilibrium under optimal tax rates has the ratio of government spending to output equal to 19 percent, and such that aggregate working time represents a third of total available time ( $\theta = 1.427$ ). The resulting optimal tax rates on labor and capital income are, respectively, 26.8 percent and 9.9 percent. The interest rate is equal to 6.78 percent and the capital to output ratio is equal to 2.71.

**Progressive Taxation on Total Income** Under a progressive tax system where the tax base is total income, the problem that individuals face is the maximization of

$$\sum_{j=0}^J \beta^j \frac{c_j^{1-\sigma} (1-l_j)^\eta}{1-\sigma},$$

$$\text{s.t. } c_j + a_{j+1} = (1 - \bar{\tau}(y_j))y_j + a_j, \quad j = 0, \dots, J,$$

where

$$y_j = \hat{w}z_j l_j + \hat{r}a_j.$$

The optimality conditions, although standard, are key to understand the results of this section:

$$\begin{aligned} -U_{c_j} + \beta U_{c_{j+1}} [1 + (1 - \tau(y_{j+1}))\hat{r}] &= 0, \\ U_{l_j} + U_{c_j} (1 - \tau(y_{j+1}))\hat{w}z_j &\leq 0, \quad \text{with equality if } l_j > 0. \end{aligned}$$

The distance between the tax function  $\bar{\tau}()$  and  $\tau()$  thus determines how costly it is for the government to collect a certain amount of tax revenues, that is, it determines the wedge between the average and the marginal tax rates: whereas  $\bar{\tau}()$  in the budget constraints measures the resources lost by the individual to the tax authority,  $\tau()$  in the optimality conditions measures how costly it is to collect these taxes at the margin.

To compare the allocations under progressive taxation to that under age-independent taxes, the economies need to be parameterized so that they are indeed comparable. In particular, per capita government spending is maintained at the same level across all economies, and government debt is always equal to zero. I adjust the parameter  $\pi_0$  from the tax functions in order for the latter requirement to hold in all economies. All other parameters of the model are kept at their initial values.

The results for different values of  $\rho$ —which controls the degree of progressivity by setting the wedge between the average and marginal tax functions—are presented in Table 1. The second column presents the results from the optimal age-independent tax code, and the next five present results under tax systems of various degrees of progressivity. With  $\rho = 2.471$ , the distance between the average and marginal tax functions matches that of the estimated tax functions on Figure 2. A value of unity for the parameter  $\rho$  implies that the tax functions are consistent with each other.

The last row of table 1 gives the percentage by which consumption would need to be increased in each period of ones' life for that individual to be indifferent between a given allocation and the allocation obtained under the optimal age-independent tax code. In other words, suppose that country A operates under the optimal tax code and country B operates under some progressive tax system. Then *ccomp* gives the percentage by which consumption at each age for an individual living in country B would need to be increased in order for that individual to be indifferent between living in either country.

Table 1 indicates that when  $\rho = 2.471$ , capital and labor are respectively 24 and 8 percent lower under the progressive tax system with income as the tax base relative to its level under the optimal age-independent tax code. These lower levels of inputs translate into a 14 percent reduction in output. Figure 8 shows that the

Table 1: Progressive Taxation on Total Income

	Age-Ind	Progressive Taxation on Total Income				
$\rho$	Taxes	2.471	1.500	1.000	0.667	0.333
$q$	1.000	0.859	0.917	0.946	0.964	0.981
$c$	0.607	0.515	0.556	0.576	0.588	0.601
$i$	0.203	0.155	0.172	0.180	0.185	0.191
$k$	2.711	2.062	2.287	2.398	2.469	2.540
$l$	0.571	0.525	0.549	0.560	0.568	0.575
$ccomp$ (%)	0.000	6.249	2.097	0.423	-0.557	-1.442

consumption profile is lower under progressive taxation (pink or light lines) than under proportional tax rates (blue or dark lines) while the leisure profile is higher. The higher leisure profile does not make up for lower consumption, however, as individuals would require consumption to be 6.25 percent higher in every period to be indifferent between the two tax systems under consideration. It is important to note that the consumption and leisure profiles are much flatter under the progressive tax code than under proportional taxes, as effective after-tax interest rates are generally lower—except for the last two periods of life—under the progressive tax system.

Table 1 also suggests that the welfare cost of progressive taxation is very sensitive to the degree of progressivity of the tax code. When the distance between the marginal and average tax function is such that the functions are consistent with each other, the welfare cost decreases to less than half a percent of consumption per period. As the degrees of progressivity decreases further, the allocation under the progressive tax system becomes preferable to that under the optimal age-independent tax system.

Given how different the capital income tax profiles from the Ramsey problem are from those implied by the data, it will prove interesting to identify the extent to which the treatment of capital income taxation affects the above results.

**Progressive Taxation on Labor Income** I now consider a tax system where labor income is taxed at progressive rates while capital income is taxed at a fixed proportional rate. It is interesting to note that the optimal labor income tax profile under the restriction that capital income be taxed at a flat proportional rate is more

Table 2: Progressive Taxation on Labor Income ( $\tau^k = 9.92\%$ )

$\rho$	Age-Ind	Progressive Taxation on Labor Income				
	Taxes	2.471	1.500	1.000	0.667	0.333
$q$	1.000	0.918	0.955	0.973	0.984	0.995
$c$	0.607	0.540	0.570	0.585	0.594	0.603
$i$	0.203	0.188	0.195	0.198	0.200	0.202
$k$	2.711	2.504	2.599	2.643	2.671	2.698
$l$	0.571	0.522	0.544	0.554	0.561	0.568
$ccomp$ (%)	0.000	4.176	1.963	1.072	0.553	0.088

hump-shaped than the one shown in Figure 1.

The budget constraint under this tax system is given by

$$c_j + a_{j+1} = [1 - \bar{\tau}(y_j^l)]y_j^l + [1 + (1 - \tau^k)\hat{r}]a_j, \quad j = 0, \dots, J,$$

where  $y_j^l = \hat{w}z_jl_j$ . The optimality conditions under this tax code, which show that the asset accumulation decision is no longer directly affected by progressive taxation, are as follows:

$$\begin{aligned} -U_{c_j} + \beta U_{c_{j+1}}[1 + (1 - \tau^k)\hat{r}] &= 0, \\ U_{l_j} + U_{c_j}[1 - \tau(y_j^l)]\hat{w}z_j &\leq 0, \quad \text{with equality if } l_j > 0. \end{aligned}$$

As a starting point, the tax rate on capital is set at  $\tau^k = 9.92\%$ , its value under the optimal age-independent tax system. I then simulate the economy under different values for  $\rho$ , adjusting the parameter  $\pi_0$  from the tax functions to keep government debt equal to zero under a constant level of per-capita government spending. Results appear in Table 2.

The impact on the capital stock of taxing labor income at progressive rates is much less pronounced than with total income as a tax base. With  $\rho = 2.471$ , capital is only 7.6 percent lower than under the optimal age-independent tax code. This translates into higher output and consumption than under progressive taxation on total income, even though labor is lower. Nevertheless, the shape of the consumption and leisure profiles, shown in Figure 9 under  $\rho = 0.333$ , is such that even for relatively low degrees of progressivity, individuals prefer the age-independent tax system.

Table 3: Progressive Taxation on Labor Income ( $\tau^k = 0.0\%$ )

$\rho$	Age-Ind	Progressive Taxation on Labor Income				
	Taxes	2.471	1.500	1.000	0.667	0.333
$q$	1.000	0.915	0.955	0.974	0.984	0.998
$c$	0.607	0.533	0.564	0.580	0.594	0.599
$i$	0.203	0.193	0.200	0.204	0.200	0.209
$k$	2.711	2.568	2.673	2.721	2.671	2.782
$l$	0.571	0.512	0.535	0.546	0.561	0.560
$ccomp$ (%)	0.000	6.234	3.714	2.711	2.131	1.611

Interestingly, the fact that individuals are not better off under a flat capital income tax is not because the level of that tax is too high, as Table 3 illustrates. The results presented in that table were obtained under a progressive labor income tax and a proportional tax on capital income equal to zero. The reason again lies in the shape of the consumption and leisure profiles. Figure 10 shows these profiles when  $\rho = 0.333$ . Notice that under that tax code aggregate consumption is close to its level under age-independent taxes, and that the labor supply is higher, suggesting that leisure is higher than under age-independent taxes. However, consumption is lower than under age-independent taxes except for the last 15 years of life, which are discounted much more heavily than the early years, during which consumption is lower than under age-independent taxes. The fact that leisure is higher during all but the first and last few years of life does not make up for lower consumption early in life.

Indeed, the value of  $\tau^k$  which achieves the highest level of utility—adjusting  $\pi_2$  to keep tax revenues constant—is around 40.0%. Results from this experiment appear in Table 4, which reveals that even for fairly large degrees of progressivity, progressive taxation is preferred to the optimal age-independent tax system. As Figure 11 shows (under  $\rho = 1.0$ ), the higher tax rate on capital income, which lowers the after-tax interest rate, induces individuals to choose a flatter consumption/leisure profile.

Table 4: Progressive Taxation on Labor Income ( $\tau^k = 40.0\%$ )

	Age-Ind	Progressive Taxation on Labor Income				
$\rho$	Taxes	2.471	1.500	1.000	0.667	0.333
$q$	1.000	0.903	0.933	0.948	0.957	0.966
$c$	0.607	0.549	0.575	0.587	0.595	0.603
$i$	0.203	0.163	0.168	0.171	0.172	0.174
$k$	2.711	2.179	2.245	2.277	2.297	2.317
$l$	0.571	0.550	0.569	0.579	0.585	0.591
$ccomp$ (%)	0.000	0.735	-0.849	-1.502	-1.886	-2.234

## 6 Conclusion

This paper studies optimal and progressive taxation in a life-cycle economy similar to the one developed by Auerbach and Kotlikoff (1987). In this economy, the government generally uses both capital and labor income taxes, even in the long run. Under a widely used utility function, the optimal tax rate on capital and labor income vary with age and are a function of the labor supply: when the labor supply increases (decreases), the tax rate on capital income is negative (positive) and the tax rate on labor income is increasing (decreasing). Through these principles, the government essentially attempts to tax consumption and leisure relatively heavily when they are relatively high.

The marginal tax rates that individuals face in the U.S. also depend on age. This follows from the progressivity of the U.S. tax code as well as the fact that earnings vary over the lifetime of individuals. A comparison of the shape of the optimal income tax profiles to the implied profiles from the data reveals that these profiles differ in two important ways. First, capital income is taxed too heavily, especially at young ages and during retirement. Second, although the labor income tax profile shares the hump shape of the Ramsey tax profile, the peak occurs later in the data than in the model. This discrepancy can be explained by the fact that the optimal labor tax scheme depends on hours worked (or the labor supply), and the tax rates in a progressive income tax system are an increasing function of earnings, which peak later in life than hours worked.

Finally, allocations under different progressive tax systems are compared to the allocation obtained under taxes obtained from a Ramsey problem which imposes a no debt restriction and restricts tax rates to be age-independent. Results indicate that even if the progressivity of the tax code introduces additional distortions, a mild degree of progressivity is preferred to flat progressive taxes. An important questions left for future research has to do with the optimal degree of progressivity.

## Appendix A Proofs

**Proof of Proposition 1.** Combining the consumer's first order conditions for consumption (3) and labor (4), for the non-trivial case of positive labor supply, implies

$$-\frac{U_{l_j}}{U_{c_j}} = z_j w_j = z_j \hat{w}(1 - \tau_j^w). \quad (29)$$

This expression can then be compared to its analogue from the Ramsey problem. Combining the government's first order condition with respect to consumption (14) to that with respect to labor (15), again assuming a positive labor supply, implies

$$-\frac{(1 + \lambda)U_{l_j} + \lambda U_{l_j} H_j^l}{(1 + \lambda)U_{c_j} + \lambda U_{c_j} H_j^c} = z_j \hat{w}. \quad (30)$$

Combining equations (29) and (30), the tax rate on labor income for an age- $j$  individual born in period  $t$  is given by

$$\tau_j^w = \frac{\lambda(H_j^l - H_j^c)}{1 + \lambda + \lambda H_j^l}. \quad (31)$$

Since  $\lambda$  is in general different from zero, this tax rate on labor income will be equal to zero only if  $H_j^l = H_j^c$ .

Similarly, consider the consumer's first order condition for consumption (3) at age  $j$  and  $j + 1$ . Together with the consumer's first order condition for asset holdings (5) implies

$$\begin{aligned} \frac{U_{c_j}}{\beta U_{c_{j+1}}} &= 1 + r_{j+1} \\ &= 1 + (1 - \tau_{j+1}^k) \hat{r}. \end{aligned} \quad (32)$$

The government's counterpart of (32) is obtained by using the ratio of the government's first order condition with respect to consumption (14) at age  $j$  and  $j + 1$ , as well as the first order condition with respect to capital (13):

$$\frac{(1 + \lambda)U_{c_j} + \lambda U_{c_j} H_j^c}{(1 + \lambda)\beta U_{c_{j+1}} + \lambda \beta U_{c_{j+1}} H_{j+1}^c} = 1 + \hat{r}. \quad (33)$$

Dividing (32) by (33) gives

$$\frac{1 + \hat{r}}{1 + (1 - \tau_{j+1}^k)\hat{r}} = \frac{1 + \lambda + \lambda H_j^c}{1 + \lambda + \lambda H_{j+1}^c}, \quad (34)$$

which implies that the optimal tax rate on capital income is different from zero unless  $H_j^c = H_{j+1}^c$ .

The proof of the Proposition then follows directly from the definitions of  $H_j^c$  and  $H_j^l$  (equations (16) and (17)) under utility function (18). Since  $H_j^c = -\sigma - \eta/(1 - l_j)$  equation (34) can be re-written as

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_j))}{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_{j+1}))}. \quad (35)$$

Notice that  $\tau_{j+1}^k > 0$  if and only if

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \hat{r}}{1 + (1 - \tau_{j+1}^k)\hat{r}} > 1. \quad (36)$$

From equations (35) and (36) follows that  $\tau_{j+1}^k > 0$  if and only if  $l_{j+1} < l_j$ , which proves the first part of the Proposition.

For the second part of the Proposition, notice that  $H_j^l - H_j^c = 1/(1 - l_j)$ . Equation (31) then implies that

$$\begin{aligned} \tau_j^w/(1 - \tau_j^w) &= \frac{\lambda(H_j^l - H_j^c)}{1 + \lambda + \lambda H_j^c} \\ &= \frac{\lambda}{1 + \lambda - l_j(1 + \lambda(1 - \sigma)(1 + \theta)) - \lambda\sigma}. \end{aligned}$$

It follows that the ratio  $[\tau_j^w/(1 - \tau_j^w)]/[\tau_{j+1}^w/(1 - \tau_{j+1}^w)]$  is bigger than one if and only if  $l_{j+1} < l_j$ . ■

**Proof of Proposition 2.** TBW ■

**Proof of Proposition 3.** TBW ■

## References

- [1] Ambler, Steeve. 2000. "Optimal Time Consistent Fiscal Policy with Overlapping Generations." Manuscript.
- [2] Auerbach, Alan J. and Kotlikoff, Laurence J. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press, 1987.
- [3] Atkeson, Andrew, Chari, V.V. and Kehoe, Patrick J. "Taxing Capital income: A Bad Idea." Federal Reserve Bank of Minneapolis Quarterly Review, Vol. 23 (3), Summer 1999.
- [4] Atkinson, Anthony B. and Sandmo, Agnar. "Welfare Implications of the Taxation of Savings." *The Economic Journal*, Vol. 90, September 1980.
- [5] Boadway, Robin W. and Keen, Michael. 2000. "Redistribution," in Anthony B. Atkinson and Francois Bourguignon (eds.), *Handbook of Income Distribution Volume 1* (Amsterdam: North-Holland), pp. 677-789.
- [6] Castañeda, Ana, Días-Giménez, Javier and Ríos-Rull, José-Víctor. "Accounting for Earnings and Wealth Inequality." Manuscript, 2000.
- [7] Chamley, Christophe. "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica*, Vol. 54 (3), May 1986.
- [8] Chari, V.V. and Kehoe, Patrick J. "Optimal Fiscal and Monetary Policy," in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics, Vol. 1*. Amsterdam: North-Holland, 1999.
- [9] Conesa, Juan Carlos and Krueger, Dirk. 2002. "On the Optimal Progressivity of the Income Tax Code." Manuscript.
- [10] Erosa, Andrés and Gervais, Martin. "Optimal Taxation in Infinitely-Lived Agent and Overlapping Generations Models: A Review." *Federal Reserve Bank of Richmond Economic Quarterly*, Vol. 87 (2), Spring 2001.

- [11] Erosa, Andrés and Gervais, Martin. “Optimal Taxation in Life-Cycle Economies.” *Journal of Economic Theory*, Vol. 105 (2), 2002.
- [12] Escolano, Julio. “Optimal Fiscal Policy in Overlapping Generations Models.” Manuscript, The University of Minnesota, 1991.
- [13] Feenberg, Daniel and Coutts, Elisabeth. “An Introduction to the TAXSIM Model.” *Journal of Policy Analysis and Management*, Vol. 12 (1), Winter 1993.
- [14] Gouveia, Miguel and Strauss, Robert P. “Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis.” *National Tax Journal*, Vol. 47 (2), June 1994.
- [15] Hansen, Gary D. “The Cyclical and Secular Behavior of the Labor Input: Comparing Efficiency Units and Hours Worked.” *Journal of Applied Economics*, Vol. 8, 1993.
- [16] Judd, Kenneth L. “Redistributive Taxation in a Simple Perfect Foresight Model.” *Journal of Public Economics*, Vol. 28, 1985.
- [17] Mirrlees, James A. 1971. “An exploration in the Theory of Optimum Income Taxation,” *Review of Economic Studies*, Vol. 38, pp. 175–208.
- [18] Ramsey, Frank P. “A Contribution to the Theory of Taxation.” *Economic Journal*, Vol. 37, 1927.
- [19] Samuelson, Paul A. “The Two-Part Golden Rule Deduced as the Asymptotic Turnpike of Catenary Motions.” *Western Economic Journal*, Vol. 6, 1968.
- [20] Sarte, Pierre-Daniel G.. “Progressive Taxation and Income Inequality in Dynamic Competitive Equilibrium.” *Journal of Public Economics*, Vol. 66 (1), 1997.
- [21] Stockman, David R. 2001. “Balanced-Budget Rules: Welfare Loss and Optimal Policies.” *Review of Economic Dynamics*, Vol. 4, pp. 438–459.

Figure 1: Optimal tax rates over the lifetime of individuals

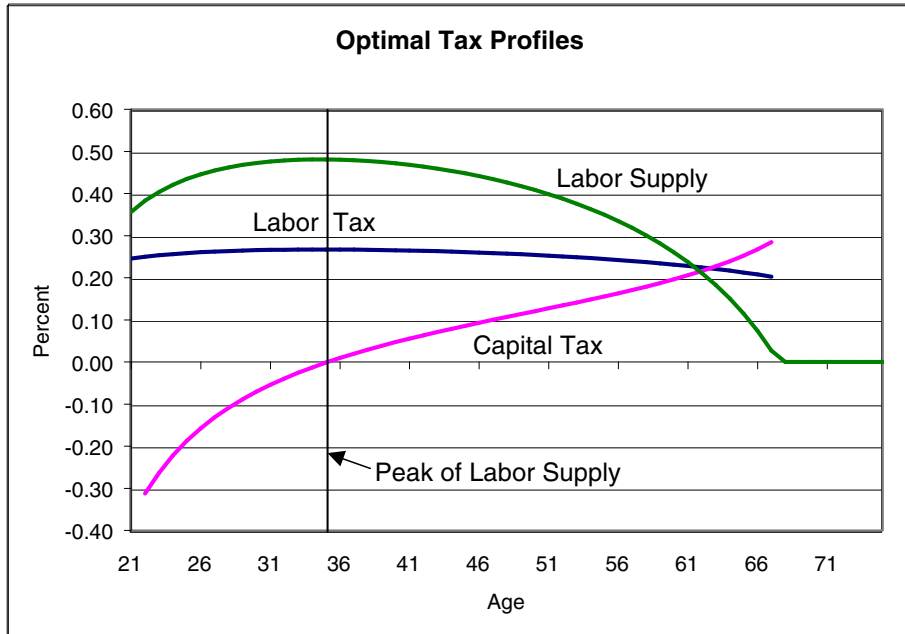


Figure 2: Tax rates over the lifetime of individuals implied by the data

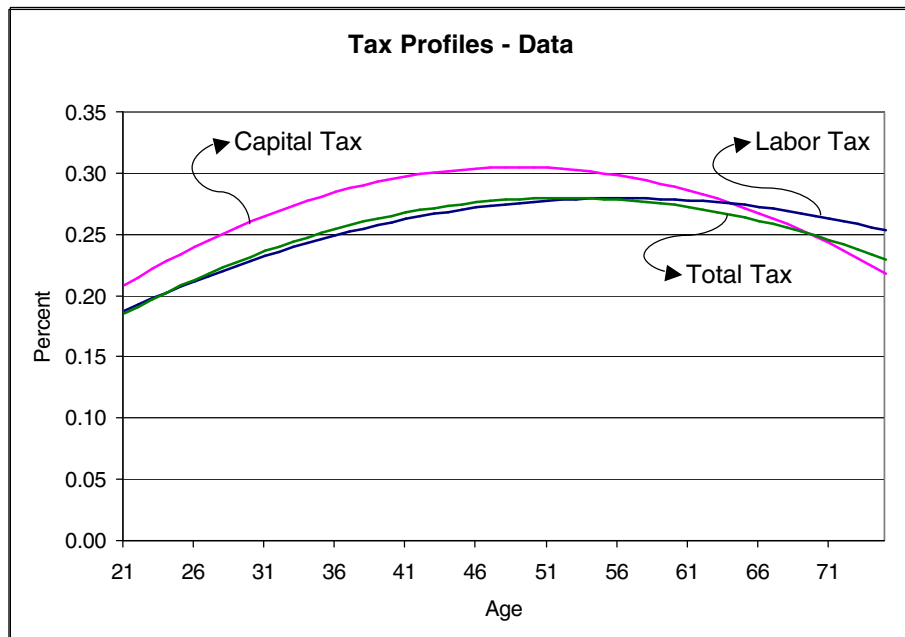


Figure 3: Total income tax rates over the lifetime of individuals

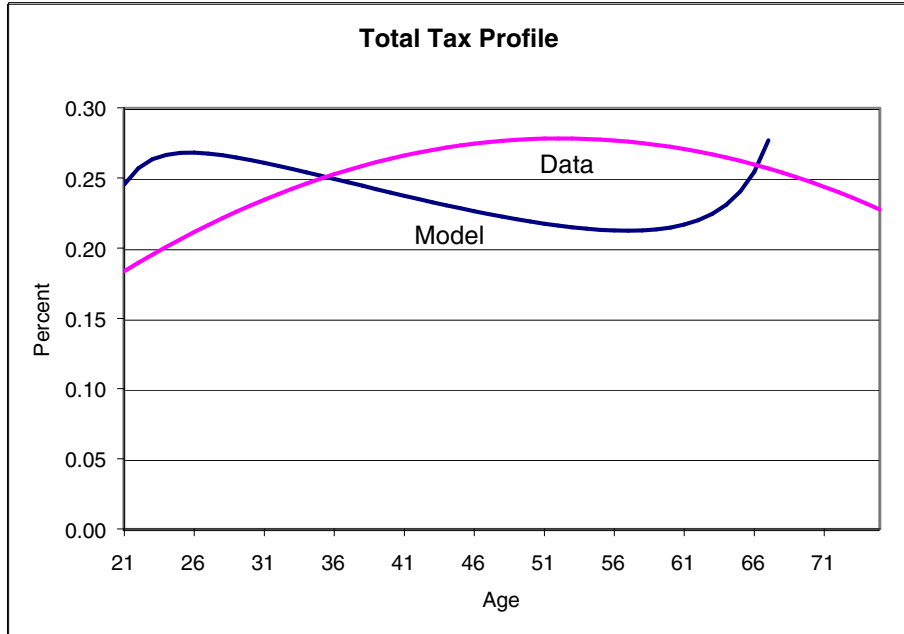


Figure 4: Capital income tax rates over the lifetime of individuals

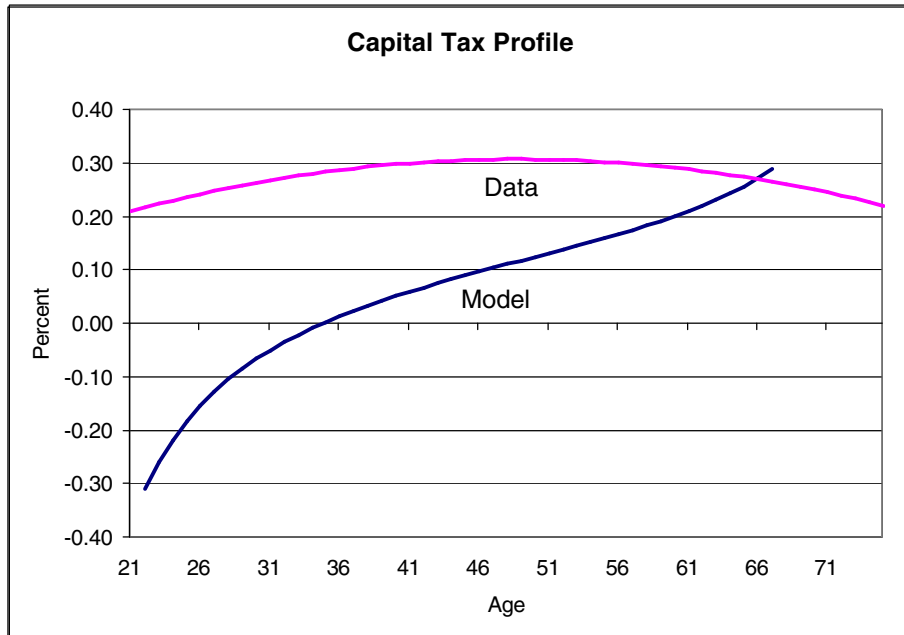


Figure 5: Labor income tax rates over the lifetime of individuals

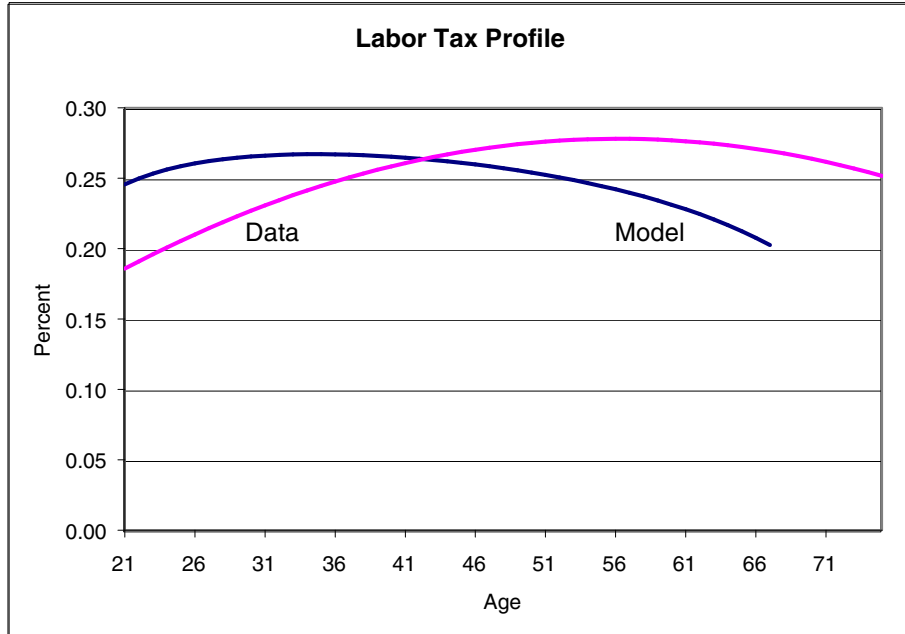


Figure 6: Labor earnings decomposition over the lifetime of individuals

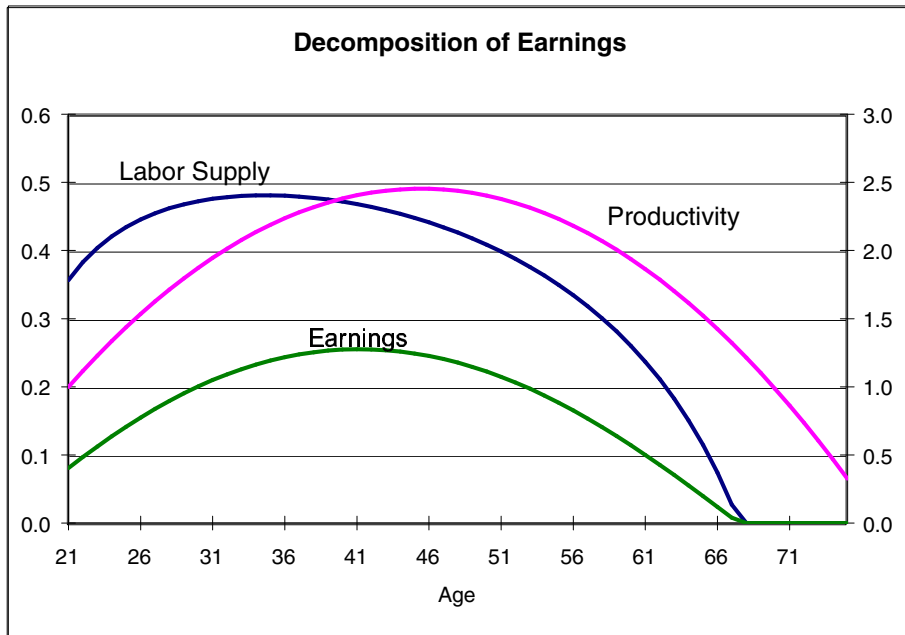


Figure 7: Average and Marginal Tax Rates from the Data

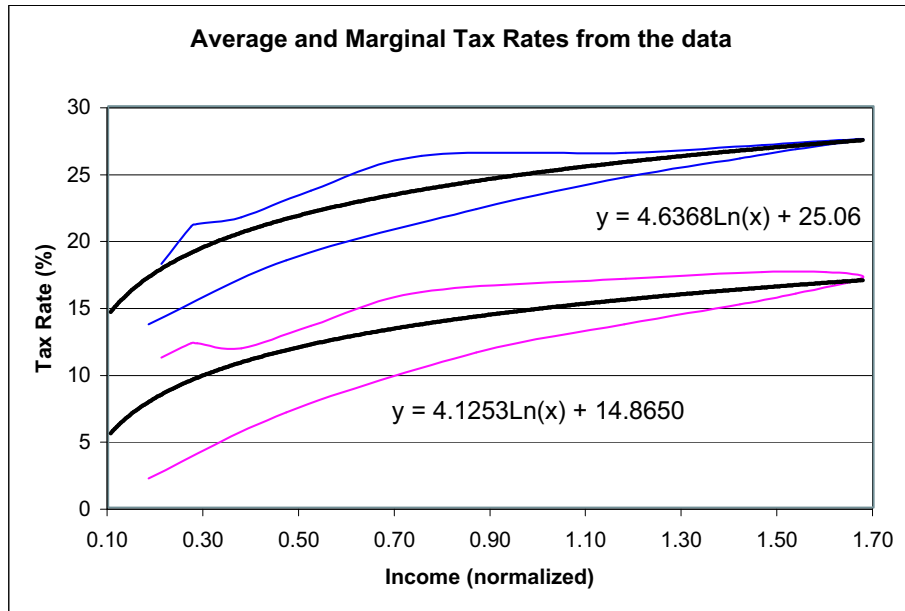


Figure 8: Age-Independent vs Progressive Taxation on Total Income:  $\rho = 2.471$

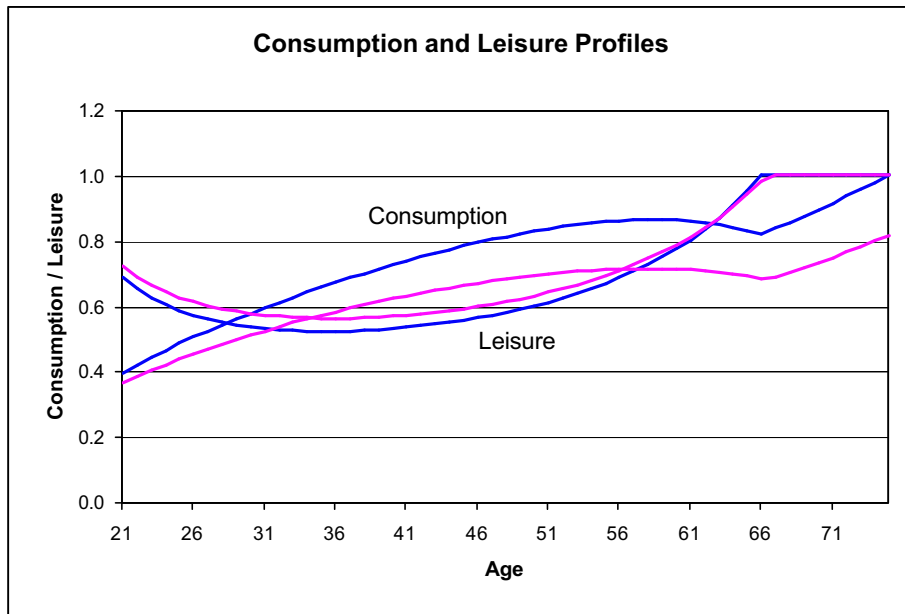


Figure 9: Age-Independent vs Progressive Taxation on Labor Income:  $\tau^k = 9.92\%$ ,  $\rho = 0.333$

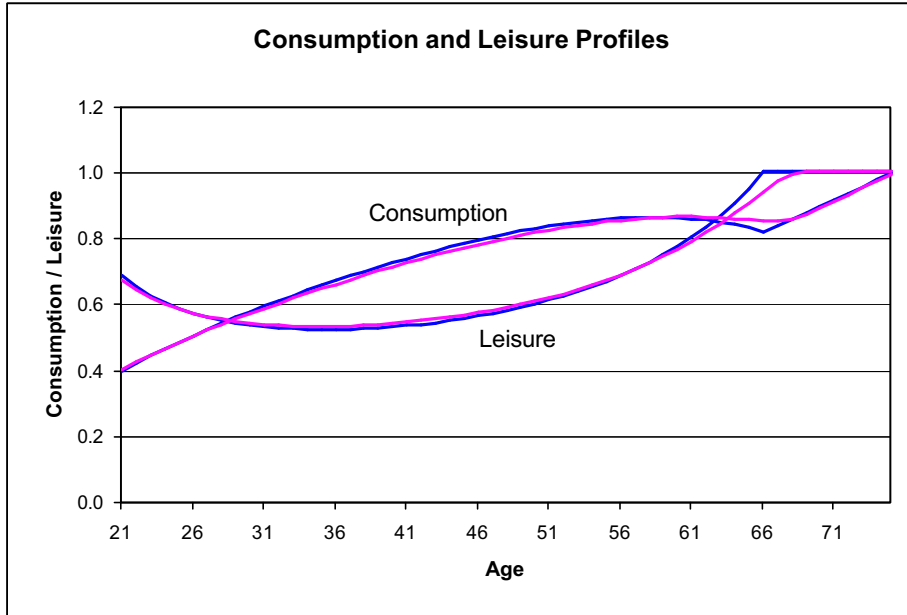


Figure 10: Age-Independent vs Progressive Taxation on Labor Income:  $\tau^k = 0\%$ ,  $\rho = 0.333$

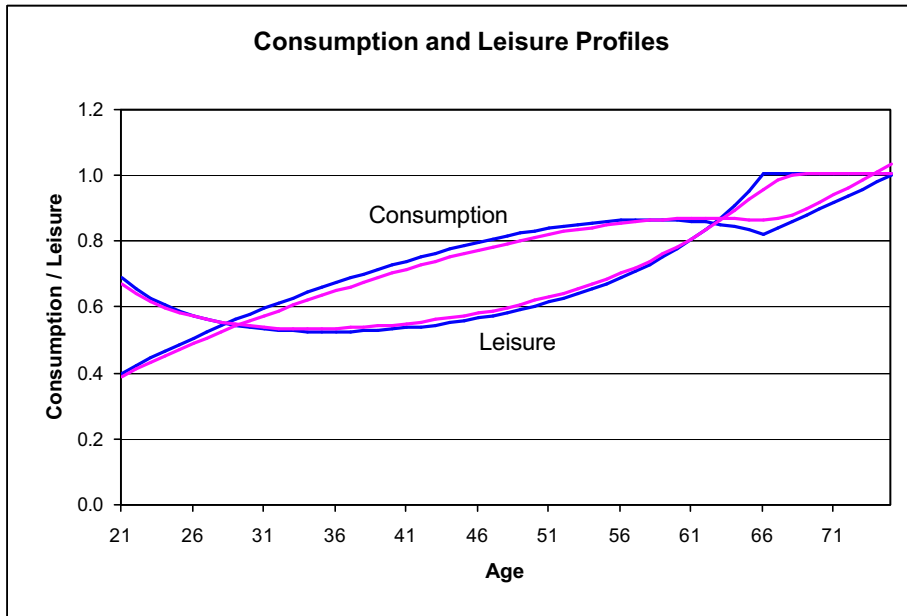


Figure 11: Age-Independent vs Progressive Taxation on Labor Income:  $\tau^k = 40\%$ ,  $\rho = 1.0$

