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

Martin Gervais

University of Southampton and IFS
gervais@soton.ac.uk

December 11, 2008

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 of 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: Progressive Taxation, Optimal 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, Andrés Erosa, Andreas Hornstein, Igor Livshits, Miguel Molico, and seminar participants at various conferences and universities 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 and Keen (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 (2005). In this paper, I argue that progressive taxation may have a role purely on efficiency grounds, without relying on any re-distribution 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.

In life-cycle economies both capital and labor income taxes are generally used by an optimizing government, even in the long run. Erosa and Gervais (2002) show that when the government has access to a full set of proportional, age-conditioned, tax rates on capital and labor income, the 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. 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. Likewise, 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.

1 Note, however, that the labor supply profile that individuals choose is itself a function of the tax rates.
Given optimal tax rates generated by a parameterized version of a standard lifecycle 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. 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.

These results suggest, at least as far as labor income taxes are concerned, that progressive taxation may provide a way for the government to imitate optimal age-dependent taxes without conditioning tax rates on individuals’ age. As is well known, however, progressive taxation introduces a wedge between marginal and average tax rates that is not present under age-dependent proportional taxes. I evaluate the cost of this wedge by computing equilibria of the life-cycle model under different specifications of progressive tax systems. I compare the allocations obtained under these tax systems to a benchmark allocation taken to be the solution to a Ramsey problem in which the government is forced to pick age-independent (proportional) tax rates and is precluded from issuing debt—a feasible alternative when the government cannot condition tax rates on age.

The desirability of progressive taxation naturally depends on the level and shape of the lifetime profiles of consumption and leisure. Under a progressive tax system in which the tax base is total income, the smoother profiles do not make up for the level effect. The capital stock under the progressive tax system is 11% below its benchmark level, yet individuals would be indifferent between progressive and flat taxes if their consumption were increased by less than 0.5% in every period of their life. Surprisingly, this result does not hinge on capital income taxes being too high: when capital income is free from taxation and only labor income is taxed at progressive rates, the consumption compensation increases to 2.7% per year, even though the stock of capital is slightly higher than its benchmark level. However, with

\(^2\)For more information about the TAXSIM model, see Feenberg and Coutts (1993).

\(^3\)Stockman (2001) studies optimal taxation without government debt in a stochastic neoclassical growth model populated a representative infinitely-lived individual. He shows that the celebrated result from Chamley (1986) and Judd (1985) on the optimality of not taxing capital income in the long run holds without government debt.
a capital income tax rate of 40%, the lower after-tax interest rate induces individuals
to choose relatively flat consumption and leisure profiles. These flat profiles more
than offset the lower level of aggregate consumption, as individuals prefer to live in
the world with progressive labor income taxes than under optimal age-independent
taxes (the consumption compensation is equal to −1.5% per year). An important
lesson from these results is that the merits of any tax system need to be evaluated as
a whole.

The rest of the paper is organized as follows. The next section presents the eco-
nomic environment and formulates a Ramsey problem. Section 3 compares implied
tax rates from the data to optimal age-dependent taxes 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 sec-
tion 4. Section 6 concludes the paper.

2 Economic Environment and Ramsey Taxes

Consider an economy populated by overlapping generations of individuals with finite
lives. Individuals make consumption and labor/leisure choices in each period so as
to maximize their lifetime utility. Firms operate a neoclassical production technol-
gy: 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 ab-
sorbs 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.

---

4This section draws from Erosa and Gervais (2002).
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, many generations are already alive. To account for these initial generations, born in periods \(t = -J, -J + 1, \ldots, 0\), it is convenient to denote by \(j_0(t)\) the age of these individuals at date 0. For all other generations, \(j_0(t) = 0\), 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 (time-invariant) 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 occur 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:

\[
U_t(\pi) \equiv \max_{j = j_0(t)} \beta^{j - j_0(t)} U(c_{t,j}, 1 - l_{t,j}),
\]

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

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 twice continuously differentiable strictly increasing in both arguments, strictly concave, and to satisfy standard Inada conditions. In 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 net of taxes, 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

\[ \beta^{j - j_0(t)} U_{c_{t,j}} - p_{t,j} = 0, \]  
\[ \beta^{j - j_0(t)} U_{l_{t,j}} + p_{t,j} w_{t,j} z_j \leq 0, \quad \text{with equality if } l_{t,j} > 0, \]  
\[ -p_{t,j} + p_{t,j+1} (1 + r_{t,j+1}) = 0, \]  
\[ a_{t,j+1} = 0, \] 

\[ j = j_0(t), \ldots, J, \] 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.\(^5\)

**Technology and Feasibility** The production technology is represented by a 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

\[ c_t + (1 + n)k_{t+1} - (1 - \delta)k_t + g_t \leq q_t, \]  

where \( c_t \) and \( g_t \) respectively denote aggregate (per capita) private and government consumption at date \( t \). Note that tomorrow’s per capita stock of capital 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 (effective labor supply) of all individuals alive at date \( t \), where the weights are given by the fraction of the population that each individual represents:

\[ c_t = \sum_{j=0}^{J} \mu_j c_{t-j,j}, \]
\[ l_t = \sum_{j=0}^{J} \mu_j z_j l_{t-j,j}. \]

\(^5\)The Inada conditions guarantee that consumption and leisure will be strictly positive in each period.
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 $\tau^k_{t-j,j}$ and $\tau^w_{t-j,j}$, respectively. In per capita terms, the government budget constraint at date $t \geq 0$ is given by

$$(1 + \hat{r}_t)b_t + g_t = (1 + n)b_{t+1} + \sum_{j=0}^{J}(\hat{r}_t - r_{t-j,j})\mu_j a_{t-j,j} + \sum_{j=0}^{J}(\hat{w}_t - w_{t-j,j})\mu_j z_{t-j,j},$$

where $w_{t,j} \equiv (1 - \tau^w_{t,j})\hat{w}_{t+j}$, $r_{t,j} \equiv (1 - \tau^k_{t,j})\hat{r}_{t+j}$, and $b_t$ denotes government debt issued at date $t$.\(^6\) Equation (10) 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 welfare (Samuelson (1968) and Atkinson and Sandmo (1980)). In other words, the government chooses a sequence of tax rates in order to maximize

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

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 policy $\pi$.

\(^6\)Note that in overlapping generations economies, the present value of taxes collected need not equal the present value of expenditures. However, since debt per capita is bounded, there is no need to impose a limit on government debt.
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.\footnote{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.} 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

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

(12)

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. As we will see in section 4, 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 allocation satisfies (12). Conversely, one can show that if an allocation satisfies the implementability constraints (12) as well as the feasibility constraint (7), then there exists a fiscal policy for which the allocation is a competitive equilibrium.\footnote{For details, see Atkeson et al. (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 which imposes age-independent tax rates.} This
equivalence allows one to set up a Ramsey problem in which the government chooses quantities rather than tax rates.

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,

$$
\max_{\{c_{t,j},l_{t,j}\}_{j=0(t)}^{J}} \sum_{t=-J}^{\infty} \gamma^{t}W_{t}
$$

subject to feasibility (7) for $t = 0,1,\ldots$. 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 (12), the function $W_{t}$ is defined as

$$
W_{t} = \sum_{j=0(t)}^{J} \beta^{j-0(t)} \left[ U(1 - l_{t,j}) + \lambda_{t}(U_{c_{t,j}}c_{t,j} + U_{l_{t,j}}l_{t,j}) \right] - \lambda_{t}A_{t,0(t)}.
$$

It should be noted that since government debt is unconstrained, the government budget constraint (10) 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.

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:

$$1 - \delta + f_{k} = \frac{1 + n}{\gamma},$$

$$\gamma^{j}\phi_{j} = (1 + \lambda)\beta^{j}U_{c_{j}} + \lambda\beta^{j}U_{c_{j}}H_{c_{j}}^{j} = \gamma^{j}\phi_{j}, \quad j = 0,\ldots,J,$$

$$\gamma^{j}\phi_{j} = (1 + \lambda)\beta^{j}U_{l_{j}} + \lambda\beta^{j}U_{l_{j}}H_{l_{j}}^{j} \leq -\gamma^{j}\phi_{j}z_{j}f_{t}, \quad j = 0,\ldots,J, \quad \text{with equality if } l_{j} > 0,$$

where

$$H_{c_{j}}^{j} = \frac{U_{c_{j},c_{j}}c_{j} + U_{l_{j},c_{j}}l_{j}}{U_{c_{j}}},$$

$$H_{l_{j}}^{j} = \frac{U_{c_{j},l_{j}}c_{j} + U_{l_{j},l_{j}}l_{j}}{U_{l_{j}}}.$$
as well as the feasibility and implementability constraints (7) and (12). Atkeson et al. (1999) refer to the terms $H^c$ and $H^l$ as \textit{general equilibrium elasticities} since they capture the relevant distortions for setting the capital and labor income tax rates in general equilibrium. 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 individuals’ life, as the general equilibrium elasticities $H^l_j$ and $H^c_j$ change as individuals age.

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 ($z_0 = 1$).\footnote{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 $j = 0, \ldots, 54$, which is then normalized so that $z_0 = 1$.} The utility function is specified as follows:

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

(20)

where $\eta = \theta(1 - \sigma)$. Here, $1/\sigma$ is the intertemporal elasticity of substitution and $\theta$ reflects the intensity of leisure in individuals’ preferences. I set the intertemporal
elasticity of substitution equal to 0.5 ($\sigma = 2$) and the discount rate 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 steady state output. Simultaneously, the value of the intergenerational discount factor ($\gamma$) is chosen to make government debt equal to zero in the final steady state. The value of $\gamma$ which accomplishes this goal is 0.948. Equation (15) 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 this parametrization of the model. It shows that capital income taxes are positive (negative) when the labor supply is decreasing (increasing), and labor income taxes follow the shape of the labor supply. By taxing or subsidizing capital, the government makes consumption and leisure in the future more or less expensive than today. The government thus uses capital income taxes to smooth individuals’ leisure and consumption profiles over their lifetime. When leisure is high tomorrow relative to today, the government taxes 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. Since leisure is constant during retirement—which is endogenous—capital income is not taxed while individuals are retired. Notice, however, that consumption and leisure during retirement are taxed indirectly by taxing the return on savings prior to retirement.

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. The fact that the tax on labor income peaks at the same age as the labor supply follows from the specification of the utility function, which assumes a unitary intratemporal elasticity of substitution between consumption and leisure. When the elasticity of substitution between consumption and leisure is below (above) one, the labor income tax profile peaks before (after) the labor supply profile peaks.

It is generally known that taxing capital income is equivalent to taxing consumption tomorrow more than today. It is less often pointed out that the same holds true for leisure: taxing capital is also equivalent to taxing leisure tomorrow more than today.
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.\footnote{For more information about the TAXSIM model, see Feenberg and Coutts (1993).} 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.\footnote{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. On the other hand, the information available in the CPS corresponds closely to the model, which also abstract from housing, the main component of itemized deductions.} 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 reminder 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.\footnote{The total income tax rates from the model are actually average tax rates, computed as the ratio of tax liability to 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.} 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.

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 the most 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. Accordingly, many different specifications of tax systems, in particular with respect to capital income taxation, will be investigated in section 5.

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, the optimal labor income tax peaks at the same time as individuals’ working time (or labor supply). In contrast, the tax profile from a progressive tax on labor income will by definition peak at the same age as earnings. Figure 6 shows that the age-profile of earnings peaks later that hours worked in the model, suggesting that the tax profile that emerges from a progressive tax system will naturally peak later than the optimal tax profile.

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, which is the subject of the next section.

4 Steady State Ramsey 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. The idea is that since we want to compare steady states where government debt is zero, we need a way of

\[\text{\footnotesize \cite{cite16}}\]

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 as $\gamma \to 1$, as long as the government can issue debt. Of course, the objective function of the full Ramsey problem (13) is not well-defined when the government does not discount the utility of future generations. As will become clear below, the restriction on government debt substantially complicates the problem.
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:

\[
\max \sum_{j=0}^J \beta^j U(c_j, l_j) \tag{21}
\]
\[
s.t. \quad c_j + a_{j+1} = w_j z_j l_j + (1 + r_j) a_j, \quad j = 0, 1, \ldots, J, \tag{22}
\]

where \( a_0 = 0 \);

2. Factor prices are competitive:

\[
\hat{r} = f_k - \delta, \tag{23}
\]
\[
\hat{w} = f_l; \tag{24}
\]

3. The government budget constraint is satisfied:

\[
g = \sum_{j=0}^J (\hat{w} - w_j) \mu_j z_j l_j + \sum_{j=0}^J (\hat{r} - r_j) \mu_j a_j; \tag{25}
\]

4. The allocation is feasible:

\[
\sum_{j=0}^J \mu_j c_j + (n + \delta) k + g = f(k, l). \tag{26}
\]

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 Proposition 1, the set of fiscal instruments is not restricted other than by the restriction in Definition 1 that the government balances its budget on a period by period basis. Proposition 2 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 1** An allocation \( \{ (c_j, l_j) \}_{j=0}^{J}, k \) is implementable if and only if it is feasible (26) and satisfies the following implementability constraints:

\[
\sum_{j=0}^{J} \beta^j \left( U_{c_j} c_j + U_{l_j} l_j \right) = 0, \tag{27}
\]

\[
\sum_{j=0}^{J-1} \mu_{j+1} \sum_{i=j+1}^{J} \beta^{i-j} \left[ \frac{U_{c_i} c_i + U_{l_i} l_i}{U_{c_j}} \right] = k. \tag{28}
\]

**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 equals the aggregate stock of capital, thereby forcing government debt to be zero.

**Proposition 2** An allocation \( \{ (c_j, l_j) \}_{j=0}^{J}, k \) is implementable with age-independent taxes if and only if it is feasible (26), satisfies the implementability constraints (27)–(28), as well as the following additional implementability constraints:

\[
l_j \left[ \frac{U_{l_j}}{z_j U_{c_j}} - \frac{U_{l_0}}{z_0 U_{c_0}} \right] = 0 \quad j = 1, \ldots, J, \tag{29}
\]

\[
\frac{U_{c_j}}{U_{c_j+1}} - \frac{U_{c_1}}{U_{c_1}} = 0 \quad j = 1, \ldots, J - 1. \tag{30}
\]

**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.\(^{17}\) Only then can an allocation be decentralized with age-independent tax rates. The first constraint is multiplied by \( l_j \) since the term inside the brackets need not equal zero when individuals are retired \( (l_j = 0) \).

\(^{17}\)Note that constraint (30) forces the government to subsidize borrowing when individuals choose to hold a negative amount of assets.
The Ramsey problem than consists of maximizing steady state utility (21) subject to the feasibility constraint (26) as well as the appropriate implementability constraints: equations (27)–(28) for age-dependent taxes; and equations (27)–(30) 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 (negative debt) 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 in that steady state with government debt (savings).

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 age-independent 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. The average and marginal tax functions implied by the data are not consistent with each other. Since the average tax rates are likely more reliable than marginal tax rates, I use the average tax data to estimate an average
tax function from which the marginal tax function is derived. Figure 7 depicts the average tax rates as a function of income. Although many different functional forms could be used to approximate the progressivity of the U.S. tax code, the simplest one is perhaps a log-linear function, also shown on Figure 7.\textsuperscript{18}

The fact that the marginal and average tax data are not consistent with each other can be illustrated as follows. Let the average and marginal tax functions be respectively given by

\[
\bar{\tau}(y) = \pi_0 + \pi_1 \log y,
\]

\[
\tau(y) = (\pi_0 + \rho \pi_1) + \pi_1 \log y.
\]

Of course, these tax functions are only consistent with each other if the parameter \(\rho\) is equal to one. Nevertheless, the value of this parameter which best fits the data on marginal tax rates is approximately \(\rho = 2.5\), which suggests that the tax code is in some sense ‘more progressive’ than the log-linear tax functions imply. This should be kept in mind in interpreting the results of the simulations, which use consistent tax functions (\(\rho = 1\)).

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 2, that is, maximizing steady state utility (21) subject to the feasibility constraint (26) as well as implementability constraints (27)–(30). I do so using the same parameter values as in

\textsuperscript{18}The tax function estimated by Gouveia and Strauss (1994)—which has been used by Sarte (1997), Conesa and Krueger (2005) and Castañeda et al. (2003)—has the same shape as the log-linear tax function over the relevant range of income. Although the log-linear tax function is much easier to use, it does not have some of the nice properties that the Gouveia-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.
section 3.1, except for per capita 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 of Total Income** Under a progressive tax system where the tax base is total income, the problem that individuals face is the maximization of

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

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

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:

$$
-U_{c_j} + \beta U_{c_{j+1}}[1 + (1 - \tau(y_{j+1}))\hat{r}] = 0,
$$

$$
U_{l_j} + U_{c_j}(1 - \tau(y_{j+1}))\hat{w}z_j \leq 0, \quad \text{with equality if } l_j > 0.
$$

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 appear in Table 1. The second column presents the results from the optimal age-independent tax code, and the third column presents results under a progressive tax system in which the tax base is total income. The last row of Table 1 gives the percentage by which consumption would need to be increased in each period of one’s 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 age-independent tax code and country B operates under some progressive tax system. Then $c_{comp}$ 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 being born in either country.

Table 1 indicates that capital and labor are respectively 11 and 2 percent lower under the progressive tax system with income as the tax base relative to their levels under the optimal age-independent tax code. These lower levels of inputs translate into a 5.4 percent reduction in output. Figure 8 shows that the consumption profile is lower under progressive taxation than under proportional tax rates while the leisure profile is higher. It is also important to note that the consumption and leisure profiles are much flatter under progressive taxation 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. The benefits of progressive taxation (flatter profiles,
and higher leisure), however, do not outweigh its costs (lower consumption) in this case, as individuals would require consumption to be 0.42 percent higher in every period to be indifferent between the two tax systems under consideration.\footnote{19}{As one would expect, the desirability of progressive taxation is very sensitive to the degree of progressivity. In particular, for slightly lower progressivity, progressive taxation is preferred to optimal age-independent taxes.}

We saw in section 3 that the capital income tax profile from the Ramsey problem was very different from those implied by the data. Accordingly, it will prove interesting to identify the extent to which the above results are affected by the tax treatment of capital income.

**Progressive Taxation of 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 rate is more hump-shaped than the one shown in Figure 1.

The budget constraint under this tax system is given by

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

where \(y_j^l = \hat{w}z_j l_j\). 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:

\[
-U_{c_j} + \beta U_{c_{j+1}}[1 + (1 - \tau^k)\hat{r}] = 0,
\]

\[
U_{l_j} + U_{c_j}[1 - \tau(y_j^l)]\hat{w}z_j \leq 0, \quad \text{with equality if } l_j > 0.
\]

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 vary the tax rate on capital income to measure its impact on the economy, adjusting the parameter \(\pi_0\) of the labor 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 \(\tau^k = 9.92\%\),
Table 2: Progressive Taxation vs Optimal Proportional Taxation

<table>
<thead>
<tr>
<th></th>
<th>Age-Independent</th>
<th>Progressive Taxation of Labor Income</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Taxes</td>
<td>( \tau^k = 9.92% )</td>
<td>( \tau^k = 0.0% )</td>
<td>( \tau^k = 40.0% )</td>
<td></td>
</tr>
<tr>
<td>( q )</td>
<td>1.000</td>
<td>0.973</td>
<td>0.974</td>
<td>0.948</td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>0.607</td>
<td>0.585</td>
<td>0.580</td>
<td>0.587</td>
<td></td>
</tr>
<tr>
<td>( i )</td>
<td>0.203</td>
<td>0.198</td>
<td>0.204</td>
<td>0.171</td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>2.711</td>
<td>2.643</td>
<td>2.721</td>
<td>2.277</td>
<td></td>
</tr>
<tr>
<td>( l )</td>
<td>0.571</td>
<td>0.554</td>
<td>0.546</td>
<td>0.579</td>
<td></td>
</tr>
<tr>
<td>( c\text{comp} ) (%)</td>
<td>0.000</td>
<td>1.072</td>
<td>2.711</td>
<td>(-1.502)</td>
<td></td>
</tr>
</tbody>
</table>

capital is only 2.5 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 slightly lower. Nevertheless, the shape of the consumption and leisure profiles, shown in Figure 9 for \( \tau^k = 9.92\% \), is such that individuals prefer the age-independent tax system.

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 the fourth column of Table 2 indicates. The results in that column were obtained under progressive labor income taxes and a proportional tax on capital income equal to zero. While the capital stock is now higher than under age-independent taxes, the labor supply is much lower. Since the interest rate is higher than under age-independent taxes, the consumption profile, shown in Figure 10, is not as flat as it is under age-independent taxes. The fact that leisure is higher during the first and last few years of life does not make up for the lower consumption profile, and age-independent taxes are still preferred to progressive taxation.

The value of \( \tau^k \) which achieves the highest level of utility—adjusting \( \pi_0 \) to keep tax revenues constant—is around 40.0\%. Results from this experiment appear in the last column of Table 2. Even though capital is 16 percent below its benchmark level, consumption is only 3 percent lower. As Figure 11 shows (under \( \tau^k = 40.0\% \)), the higher tax rate on capital income lowers the after-tax interest rate and induces individuals to choose flatter consumption/leisure profiles. These flat profiles more than compensate for the lower aggregate consumption and leisure levels, as the progressive
tax system is now preferred to the optimal age-independent tax system.

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.

Finally, allocations under different progressive tax systems are compared to a benchmark allocation taken to be the solution to a Ramsey problem which imposes tax rates to be age-independent. Results indicate that even if the progressivity of the tax code introduces additional distortions, a tax system in which labor is taxed at progressive rates with a relatively high tax on capital income is preferred to optimal flat taxes.
Appendix A  Proofs

Proof of Proposition 1. By construction, implementable allocations are feasible and satisfy the implementability constraints (27) and (28). The converse needs to be shown.

Suppose that \( \{c_j, l_j\}^J_{j=0} \) is a feasible allocation (satisfies equation (7)) which satisfies the implementability constraints (27) and (28). Define before-tax prices as \( \hat{r} \equiv f_k(k, l) - \delta \) and \( \hat{w} \equiv f_l(k, l) \), where \( l \) is per capita effective labor (see equation (9)). Define the sequence of after-tax prices \( \{w_j, r_{j+1}\}^J_{j=0} \) as follows:

\[
\begin{align*}
    w_j &\equiv -\frac{U_l}{z_j U_{c_j}}, \\
    r_{j+1} &\equiv \frac{U_{c_j}}{\beta U_{c_{j+1}}} - 1,
\end{align*}
\]

and let \( p_j = \beta^i U_{c_j} > 0 \) for \( j = 0, \ldots, J \). Then, by construction, \( \{c_j, l_j\}^J_{j=0} \) satisfies the steady state version of the consumer’s first order conditions (3)–(5).

To show that the budget constraints (2) and the transversality condition (6) are satisfied, from \( a_0 = 0 \) define recursively for \( j = 0, \ldots, J \)

\[
a_{j+1} = w_j z_j l_j + (1 + r_j) a_j - c_j,
\]

and note that, given the definition of after-tax prices, the implementability constraint implies that \( a_{J+1} = 0 \).

I still need to show that the government budget constraint is satisfied, but first let me show that the sum of individuals’ assets exactly matches the stock of capital. To do so, notice that given prices as defined above and the fact that \( a_{J+1} = 0, a_{j+1} \) from equation (31) can be written as

\[
a_{j+1} = \frac{1}{p_j} \sum_{i=j+1}^J \beta^i \left( U_{c_i} c_i + U_{l_i} l_i \right)
= \sum_{i=j+1}^J \beta^{i-j} \left[ \frac{U_{c_i} c_i + U_{l_i} l_i}{U_{c_j}} \right].
\]

Multiplying \( a_{j+1} \) by \( \mu_{j+1} \) and summing over \( j \), we get

\[
\sum_{j=0}^{J-1} \mu_{j+1} a_{j+1} = \sum_{j=0}^{J-1} \mu_{j+1} \sum_{i=j+1}^J \beta^{i-j} \left[ \frac{U_{c_i} c_i + U_{l_i} l_i}{U_{c_j}} \right],
\]
which, by equation (28), equals $k$. Since the sum of all assets held by individuals is equal to the capital stock, government debt must be equal to zero. Finally, by multiplying the age-$j$ budget constraint by $\mu_j$ and adding them, we have

$$\sum_{j=0}^{J} \mu_j(c_j + a_{j+1}) = \sum_{j=0}^{J} \mu_j[w_j z_j l_j + (1 + r_j) a_j]$$

which can be re-written as

$$c + (1 + n)k = k + \sum_{j=0}^{J} \mu_j[w_j z_j l_j + r_j a_j]. \quad (32)$$

Since the production function is homogeneous of degree one, we can write the feasibility constraint as

$$c + (1 + n)k - (1 - \delta)k + g = (\hat{r} + \delta)k + \hat{w} \sum_{j=0}^{J} \mu_j z_j l_j. \quad (33)$$

Combining equations (32) and (33), we have

$$g - \hat{r}k - \hat{w} \sum_{j=0}^{J} \mu_j z_j l_j = \sum_{j=0}^{J} \mu_j z_j l_j + \sum_{j=0}^{J} \mu_j r_j a_j.$$

Using the fact that $k = \sum_{j=0}^{J} \mu_j a_j$ and rearranging the previous expression:

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

which completes the proof.

**Proof of Proposition 2.** This proof follows immediately from the previous proof, as equations (29) and (30) together with prices as defined above imply that after-tax prices are age-independent. The rest of the proof is identical to the previous one.

It is interesting to note that instead of imposing the standard implementability constraint (27), we could instead impose the following constraint

$$\beta U_{c_1}[c + (1 + n)k] = U_{c_0} k - \beta U_{l_1} / z_1,$$

which is similar to the implementability constraint used in Stockman (2001) in an infinitely-lived agent economy.
References


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

![Labor Tax Profile](image)

Figure 6: Labor earnings decomposition over the lifetime of individuals

![Decomposition of Earnings](image)
Figure 7: Average Tax Rates from the Data

![Average Tax Rates and Function](image)

Figure 8: Age-Independent vs Progressive Taxation on Total Income

![Consumption and Leisure Profiles](image)
Figure 9: Age-Independent vs Progressive Taxation on Labor Income: $\tau^k = 9.92\%$

![Figure 9](image)

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

![Figure 10](image)
Figure 11: Age-Independent vs Progressive Taxation on Labor Income: $\tau^k = 40\%$