

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

Andrés Erosa
The University of Western Ontario
aerosa@julian.uwo.ca

Martin Gervais
Federal Reserve Bank of Richmond
martin.gervais@rich.frb.org

February 6, 2001

Abstract

The behavioral implications of a tax system are determined by its marginal tax rates. 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. The question that arises is how close are the marginal tax rates derived from our life-cycle model to those faced by an average individual in the U.S. economy? We find that the age-profile of the labor income tax implied by the U.S. tax code is close to the optimal profile. The age-profile of the capital income tax implied by the U.S. tax code is, however, far from the optimal one, both in level and in shape. (JEL E62, H21)

*We would like to thank Daniel Feenberg from the NBER for calculating and providing us with the tax rates from the CPS data, and Huberto Ennis for helpful comments.

1 Introduction

The literature concerned with dynamic fiscal policy has evolved in two main directions over the last 20 years or so. On the one hand, there is a large literature on optimal taxation. In the context of a standard neoclassical growth model with infinitely-lived individuals, Chamley (1986) and Judd (1985) 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 should be taxed in the long run. On the other hand, most applied work concerned with the impact of dynamic fiscal policy uses the life-cycle framework (Auerbach, Kotlikoff and Skinner, 1983, Auerbach and Kotlikoff, 1987, and many others, surveyed in Kotlikoff, 1998). Unfortunately, the prescriptions that emanate from the former framework do not immediately generalize to the latter.

We show that in life-cycle economies, both capital and labor income taxes will generally be used by the 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 optimal tax rates vary over the life-time of individuals, that is, the optimal marginal tax rates are age-dependent. 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. A natural question to ask, then, is how close are the marginal tax rates derived from our life-cycle model to the ones faced by an average individual in the U.S. economy?

The first part of the paper is concerned with the characterization of optimal capital

and labor income taxes in life-cycle economies. Our characterization contrasts with previous prescriptions obtained in models with infinitely-lived agents. The classic result that capital income taxes should be zero in the long run essentially follows from the constancy of the consumption and work plan that defines the steady state of economies populated by infinitely-lived individuals. For the same reason the tax rate on labor is constant in the long run. This prescription has been shown to hold under a variety of utility functions and under many different types of model assumptions.¹ The fundamental difference between optimal taxation in life-cycle versus infinitely-lived agent economies is the life-cycle behavior inherent to the first type of models, which gives rise to potential roles for non-zero long run capital income taxes.

This paper focuses on one such role for capital income taxation. To illustrate this role, recall that a positive capital income tax is equivalent to taxing tomorrow's consumption at a higher rate than today's consumption. When individuals have infinite lives and the economy is in steady state, a positive tax on capital income implies that individuals essentially face an infinite tax on consumption goods in the distant future. In contrast, a capital income tax in life-cycle economies never implies that an individual will face infinite taxes in the future simply because agents die in finite time. When individuals' consumption and leisure profiles display life-cycle behavior, the government will in general want to tax consumption (or leisure) at different rates over the lifetime of individuals. One way to achieve this goal is to tax capital and labor income at rates that vary with age.

The optimal tax profiles are closely related to the leisure (and thus labor) profile chosen by individuals. Although the government cannot tax leisure directly, it can

¹For a review of this literature, see Chari and Kehoe (1999). A non-technical introduction can also be found in Atkeson, Chari and Kehoe (1999).

affect individuals' leisure decision over their lifetime by taxing more heavily goods that are complementary with leisure (Corlett and Hague, 1953). In life-cycle models, where consumption and leisure generally move together over time, consumption should be taxed more heavily when it is relatively high. Since consumption tends to increase over an individual's lifetime, optimal consumption taxes tend to increase as an individual ages, that is, capital income taxes tend to be positive, at least over a large portion of individuals' life. In contrast, consumption and leisure are constant in the steady state of infinitely-lived agent models. As a result the government in these models has no incentive to affect the relative price of leisure over time and a zero capital income tax is optimal.²

A sharp characterization of the optimal tax profiles is obtained under a Cobb-Douglas utility function. Under this utility function, the shape of the optimal tax profiles on capital and labor income is determined by the labor supply profile chosen by individuals.³ In particular, 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.

While the first part of the paper is concerned with the normative question of how tax rates should be set, the second part of the paper asks whether the U.S. tax

²A second role for capital income taxation arises when the fiscal authority cannot condition capital and labor income taxes on age. The government then uses capital income taxes to imitate tax rates that vary with age. For example, a positive tax on capital income, which implies taxing leisure tomorrow more than today, can be used to imitate a labor income tax rate that declines with age. See Alvarez, Burbidge, Farrell, and Palmer (1992) and Erosa and Gervais (2000).

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

code, which features progressive and thus age-dependent taxation, approximates the optimal tax rates generated by a parameterized version of the model. To do so, the NBER TAXSIM model is used to calculate the implied tax rates for individuals in the 1995 Current Population Survey.⁴ The shape of the marginal income tax profiles differ from the optimal ones in two important ways. First, capital income is taxed much too heavily, especially at young ages. Second, although the labor income tax profile shares the hump-shaped profile of the model, the peak occurs much later in the data than in the model. This 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 rest of the paper is organized as follows. In the next section we present the economic environment and formulate a Ramsey problem, named after Frank Ramsey's 1927 seminal paper. The main insight from that paper was to formally recognize the fact that individuals and firms react to changes in fiscal policy. The government thus needs to take individuals' reaction into account when choosing tax rates. We actually set up an alternative (and equivalent) formulation of the Ramsey problem in which the government chooses allocations rather than tax rates. We use this formulation to characterize optimal fiscal policies in section 3. In Section 4 we compare implied tax rates from the data to the optimal ones generated by a parameterized version of the model. Section 5 concludes the paper.

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

2 Economic Environment and Ramsey Problem

We consider an economy similar to that in Erosa and Gervais (2000). The 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 technology: factors are paid their marginal products. The payments received by individuals on their factors (capital and labor) are subject to proportional taxes which we assume 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 given a 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. We formally introduce this problem, the Ramsey problem, after presenting the basic economic environment.

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 we set $j_0(t) = 0$, 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 we let μ_j represent the share of age- j individuals in the population. The labor productivity level of an age- j individual is denoted z_j .

We 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. The problem faced by an individual born in period $t \geq -J$ is to maximize lifetime utility subject to a sequence of budget constraints:

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

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

Equation (1) expresses that individuals care about consumption and leisure in each period of their finite lives, and $\beta > 0$ is a factor used by individuals to discount utility in future periods to utility in the current period. In addition, individuals hold assets in order to transfer goods from one period to the next: $a_{t,j}$ 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 budget constraint (2) says 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.

These first order conditions will be used extensively in the rest of the paper. In particular, one could use these conditions 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 formulating a Ramsey problem. We will show below that these first order conditions can also be used to substitute prices out of the budget constraint (2). The resulting equation, which only involves quantities, can then be imposed as a constraint 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, $y_t = f(k_t, l_t)$, where y_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

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

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 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, we assume 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.⁵ 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

⁵In this framework, one can show that consumption taxes are redundant (see Erosa and Gervais, 2000).

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, 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

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

where $0 < \gamma < 1$ is the intergenerational discount factor and U^t denotes the indirect utility function of generation t as a function of the government tax policy. That is, U^t is given by the lifetime utility function (1), where consumption and labor supply at all ages are replaced by the demand functions that are implied by the first order conditions (3)–(5).

Note that given this utilitarian objective function, the government's desire to confiscate initial asset holdings is limited. To see this, consider the impact of confiscating the assets of a (possibly retired) individual who, at date 0, is in his last period of life.

Under this policy, this individual's consumption would be very low (may be zero) and so would his utility. Since this individual's utility has a positive and relatively large weight in the welfare function, the value of the government's objective would be driven down considerably by this front loading policy. This is not to say that the government would not tax initial assets at all, but rather that the extent to which the government will do so is limited, at least relative to what is optimal in infinitely-lived agent models.

Because initial asset holdings are given by history, taxing these assets perfectly imitates a (non-distortionary) lump-sum tax: agents cannot respond to the change in policy because their relevant decisions were made in the past. Since infinitely-lived agent models abstract from all intergenerational considerations, it is efficient to tax initial asset holdings at confiscatory rates. To circumvent this arguably trivial solution, exogenous bounds on feasible capital income taxes need to be imposed (Judd (1985), Chamley (1986), Jones et al. (1993), Chari et al. (1994)). These bounds, however, determine the magnitude of the welfare gains achieved by switching to the taxes prescribed by the Ramsey problem.⁶ The final steady state, and thus the optimal tax rate on labor income in the long run, is also determined by the magnitude of these bounds.

In life-cycle economies, as argued above, the government's desire to resort to a front loading policy is endogenously limited by intergenerational redistribution considerations: the individuals on whom the burden of the front loading policy falls are different from, and unrelated to, those who will later benefit from lower distortionary taxes in the future. Accordingly, there is no need in our framework to impose

⁶Indeed, with sufficiently high bounds a Pareto optimal equilibrium can be achieved if the return on the capital levied is sufficient to finance all future government outlays.

exogenous bounds on capital income tax rates.

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, we can formulate a Ramsey problem where the government chooses allocations rather than tax rates.⁷ 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) as well as the fact that factors are paid their marginal products to substitute out prices from consumers' 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 worth noting that implicit in this implementability constraint is the existence of age-dependent tax rates. Additional restrictions need to be imposed for an allocation to be implementable with age-independent taxes. In other words, the set of allocations that the government can pick depends crucially on the instruments available to the government.

Since these implementability constraints are constructed from the first order conditions of the consumers' problem, it can be shown that any competitive equilibrium

⁷This 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.

allocation satisfies (10). Conversely, one can show that if an allocation satisfies the implementability constraints (10) as well as the feasibility constraint (7), then it is a competitive equilibrium allocation.⁸ 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:

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

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: the first order conditions are included in the implementability constraint and need not be solved in order to define demand functions, as we previously did to set up the Ramsey problem in terms of prices (see equation (9)). If we let $\gamma^t \lambda_t$ be the Lagrange multiplier associated with generation t 's implementability constraint (10), then the function W_t is defined as

$$W_t = \sum_{j=j_0(t)}^J \beta^{j-j_0(t)} [U_{t,j} + \lambda_t (U_{c_{t,j}} c_{t,j} + U_{l_{t,j}} l_{t,j})] - \lambda_t A_{t,j_0(t)}, \quad (11)$$

where $U_{t,j} = U(c_{t,j}, 1 - l_{t,j})$. It should be noted that since government debt is unconstrained, the government budget constraint (8) does not effectively constrain the maximization problem and has therefore been omitted from the Ramsey problem. Once a solution is found, the government budget constraint can be used to back out the level of government debt.

⁸See Erosa and Gervais (2000) for details.

3 Optimal Fiscal Policy

In this section we show that the solution to the Ramsey problem generally features non-zero tax rates on labor and capital income and that these rates vary with age. In particular, and in contrast with infinitely-lived agent models, if the Ramsey allocation converges to a steady state solution, optimal capital income taxes will in general be different from zero even in that steady state. Although the main results of this section hold more generally, we will restrict attention to steady states for ease of exposition.

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:

$$1 - \delta + f_k = \frac{1 + n}{\gamma}, \quad (12)$$

$$(1 + \lambda)\beta^j U_{c_j} + \lambda\beta^j U_{c_j} H_j^c = \gamma^j \phi \mu_j, \quad j = 0, \dots, J, \quad (13)$$

$$(1 + \lambda)\beta^j U_{l_j} + \lambda\beta^j U_{l_j} H_j^l \leq -\gamma^j \phi \mu_j z_j f_l, \quad j = 0, \dots, J, \quad \text{with equality if } l_j > 0, \quad (14)$$

where

$$H_j^c = \frac{U_{c_j, c_j} c_j + U_{l_j, c_j} l_j}{U_{c_j}}, \quad (15)$$

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

as well as the feasibility and implementability constraints (7) and (10).

The first order condition with respect to capital, equation (12), 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 (13), 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, ϕ , 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 (13) 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, ϕ , appropriately weighted and discounted, times the marginal product of labor measured in efficiency units.

We now derive necessary conditions under which the Ramsey allocation features zero taxation of either labor or capital income. Since any optimal fiscal policy has to satisfy the consumer's optimality conditions, we derive these conditions by comparing the optimality conditions from the consumer's problem to those of the Ramsey problem.

Combining the consumer's first order conditions for consumption (3) and labor (4), for the non-trivial case of positive labor supply, we obtain

$$-\frac{U_{l_j}}{U_{c_j}} = z_j w_j = z_j \hat{w}(1 - \tau_j^w), \quad (17)$$

which corresponds to the usual optimality condition that the marginal rate of substitution between labor and consumption be equal to the relative price of labor faced by the consumer. We can then compare (17) to its analogue from the Ramsey problem. Combining the government's first order condition with respect to consumption (13) to that with respect to labor (14), again assuming a positive labor supply, implies

$$-\frac{(1 + \lambda)U_{l_j} + \lambda U_{l_j} H_j^l}{(1 + \lambda)U_{c_j} + \lambda U_{c_j} H_j^c} = z_j \hat{w}, \quad (18)$$

which equates the government's marginal rate of substitution between labor and consumption to the before-tax relative price of these two goods. Note that the government's marginal rate of substitution, unlike that of individuals, takes the implementability constraint into account. Also, the government cares about before-tax prices whereas individuals face after-tax prices.⁹

From equations (17) and (18), the tax rate on labor income for an age- j individual born in period t is given by

$$\tau_j^w = \frac{\lambda(H_j^l - H_j^c)}{1 + \lambda + \lambda H_j^l}. \quad (19)$$

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

The same logic applies to the tax rate on capital income. For this case, consider the consumer's first order condition for consumption (3) at age j and $j + 1$. Using the consumer's first order condition for asset holdings (5), we get

$$\begin{aligned} \frac{U_{c_j}}{\beta U_{c_{j+1}}} &= 1 + r_{j+1} \\ &= 1 + (1 - \tau_{j+1}^k)\hat{r}, \end{aligned} \quad (20)$$

which corresponds to the usual intertemporal condition that sets the marginal rate of substitution between consumption today and consumption tomorrow equal to the relative price of the same commodities, which is equal to the gross interest rate. The government's counterpart of (20) is obtained by using the ratio of the government's first order condition with respect to consumption (13) at age j and $j + 1$, as well as

⁹The reason why the government cares about before-tax prices is because production efficiency requires the pre-tax wage rate to be equal to the marginal rate of transformation of labor and consumption. In turn, the government wants to equate this marginal rate of transformation to its marginal rate of substitution, as equation (18) shows.

the first order condition with respect to capital (12):

$$\frac{(1 + \lambda)U_{c_j} + \lambda_t U_{c_j} H_j^c}{(1 + \lambda)\beta U_{c_{j+1}} + \lambda\beta U_{c_{j+1}} H_{j+1}^c} = 1 + \hat{r}. \quad (21)$$

Again, the resulting equation for the government (21) differs from its analogue from the consumer's side (20) in that the government takes the implementability constraint into account and cares about before-tax prices.

Dividing (20) by (21) we get

$$\frac{1 + \hat{r}}{1 + (1 - \tau_{j+1}^k)\hat{r}} = \frac{1 + \lambda + \lambda H_j^c}{1 + \lambda + \lambda H_{j+1}^c}, \quad (22)$$

which implies that the tax rate on capital income is different from zero unless $H_j^c = H_{j+1}^c$. These results are summarized in the following proposition.

Proposition 1 *At each date, (i) the optimal tax rate on labor income is different from zero unless $H_j^l = H_j^c$ and (ii) the optimal tax rate on capital income is different from zero unless $H_j^c = H_{j+1}^c$.*

Proposition 1 sheds some light on why the celebrated Chamley-Judd result on the optimality of not taxing capital income in the steady state of infinitely-lived agent models does not extend to life-cycle economies. Since consumption and leisure are constant in the steady state of infinitely-lived agent models, H_j^c is constant and, thus, zero capital income taxation is optimal regardless of the form of the utility function. In contrast, consumption and leisure are generally not constant over an individual's lifetime in life-cycle models, even in steady state. There is in fact no reason to expect $H_j^c = H_{j+1}^c$ and, consequently, capital income taxes will generally not be equal to zero in the long run. Obviously, if the economy is specified so that individuals' behavior

features no life-cycle elements, i.e. labor supply and consumption are independent of age, then optimal taxation works as in infinitely-lived agent models and capital income is not taxed.¹⁰

Although Proposition 1 demonstrates that labor and capital income taxes will in general not be equal to zero, it is silent on the shape of the tax profiles over the lifetime of individuals. In order to further characterize these tax rates, we consider a utility function of the form

$$U = \sum_{j=0}^J \beta^j \frac{c_j^{1-\sigma} (1-l_j)^\eta}{1-\sigma}, \quad (23)$$

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 θ measures the intensity of leisure in individuals' preferences. It is straightforward to show that optimal capital income taxes in this case are zero in the long run only under very restrictive conditions, as stated in footnote 10. 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 2 (Age-Profile of Optimal Taxes) *Assume that the utility function takes the form given by (23). 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 $l_{j+1} < l_j$.*

¹⁰To eliminate individual's life-cycle behavior, one has to set γ so that the interest rate is equal to the rate of time preferences of individuals and restrict the lifetime productivity profile to be flat—an empirically unappealing restriction. Another case in which optimal capital income taxes are zero is when preferences are such that uniform commodity taxation over individuals' lifetime is optimal. See Erosa and Gervais (2000) for details.

Proof. The proof follows directly from the definitions of H_j^c and H_j^l (equations (15) and (16)) under utility function (23). Since $H_j^c = -\sigma - \eta/(1 - l_j)$ we can re-write equation (22) as

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_j))}{1 + \lambda + \lambda(-\sigma - \eta/(1 - l_{j+1}))}. \quad (24)$$

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

$$\frac{1 + \hat{r}}{1 + r_{j+1}} = \frac{1 + \hat{r}}{1 + (1 - \tau_{j+1}^k)\hat{r}} > 1. \quad (25)$$

From equations (24) and (25), we obtain 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 (19) 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 $l_{j+1} < l_j$. ■

The government uses capital and labor income tax rates in order to imitate taxes on leisure and consumption. By taxing capital, the government makes consumption and leisure in the future more expensive than today. By taxing labor, the government lowers the price of leisure relative to consumption. Proposition 2 suggests the government wants to tax leisure relatively more when it is relatively high, and it does so through two mechanisms. First, when leisure is high tomorrow relative to today, the government taxes capital income in order to make leisure more expensive tomorrow relative to today. Second, when leisure is high tomorrow relative to today, the

government taxes labor income tomorrow less than today to make leisure relatively more expensive tomorrow.

An implication of the principle of optimal taxation developed in Proposition 2 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 2 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 2 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.

4 Age-Dependent vs Progressive Taxes

To compare the prescribed tax rates to those implied by the data, we need to parameterize the model and solve for the optimal tax rates. After computing the optimal tax rates implied by the Ramsey problem, we compare these rates to measured tax rates on labor and capital income from the Current Population Survey data.

4.1 Simulating Ramsey Taxes

We assume 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 ($z_0 = 1$).¹¹ The utility function is specified as in equation (23) 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 approximately 35 percent of total time ($\theta = 1.38$).

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 ($\delta = 0.065$).

We fix government spending such that it represents 19 percent of GDP. We then find the value of the intergenerational discount factor (γ) which makes government debt equal to zero in the final steady state. The value of γ which accomplishes this goal is 0.948. Equation (12) 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. The value of the intergenerational discount factor has a direct impact on the interest rate as well as the steady state value of government debt. Note that different values of γ influence the level of the labor and capital tax profiles rather than their shape: higher values of γ lead to lower government debt and lower tax rates in the long run.

¹¹We 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 $j = 0, \dots, 54$.

Figure 1 illustrates how taxes vary with age under our parameterization of the model. Following Proposition 2, capital income taxes are positive (negative) when the labor supply is decreasing (increasing), and labor income taxes follow the shape of the labor supply. As discussed previously, the government conditions both capital and labor income taxes on age in order to imitate a tax on leisure.

4.2 Comparing with the Data

In order to compute tax rates that individuals effectively face over their lifetime, average marginal tax rates were imputed using the NBER TAXSIM model on the Current Population Survey of the 1995 US Census.¹² The tax profiles generated by the data appear in Figure 2. As is the case for 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.

We now compare different tax rates generated by the data to those prescribed by Ramsey problem under our benchmark parameterization. Figure 3, which plots the total income (marginal) tax rates for the model and the data, 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 retired individuals, suggesting that older individuals are also over-taxed. This last point, however, may be an artifact of the representative agent model from which the Ramsey taxes are generated, which implies that all individuals retire at exactly the same age.

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

The difference between the marginal tax rates generated by the data and those generated by the Ramsey problem can best be explained by looking at the labor and capital income tax rate differentials. Figure 4 shows that the main difference between the two profiles comes from the taxation of capital income. The most striking feature of that Figure is that the U.S. tax code generally taxes capital income at very high rates relative to what the solution to the Ramsey problem prescribes. 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. Since the tax rate on capital income implied by the data is positive at young ages, individuals get a tax break while they have a negative net worth. This is reflected in Figure 2 which shows that the total income tax rate is below the labor income tax rate over the first few years.

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 Proposition 2 showed, 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.

5 Conclusion

This paper studies optimal taxation in a life-cycle economy similar to the one developed by Auerbach and Kotlikoff (1987) to study fiscal policies. We show that in general both capital and labor income taxes will be used by the government, even in the long run. Under a widely used utility function, we show that 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). We argue that though these principles, the government attempts to tax leisure relatively heavily when it is 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 much too heavily, especially at young ages. Second, although the labor income tax profile shares the hump-shaped profile of the model, the peak occurs much 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.

References

- [1] Alvarez, Yvette, Burbidge, John, Farrell, Ted and Palmer, Leigh. “Optimal Taxation in a Life Cycle Model.” *The Canadian Journal of Economics*, Vol. XXV (1), February 1992.
- [2] Auerbach, Alan J. and Kotlikoff, Laurence J. *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press, 1987.
- [3] Auerbach, Alan J. and Kotlikoff, Laurence J. and Skinner, Jonathan. “The Efficiency Gains from Dynamic Tax Reform.” *International Economic Review*, Vol. 24 (1), February 1983.
- [4] 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.
- [5] Atkinson, Anthony B. and Sandmo, Agnar. “Welfare Implications of the Taxation of Savings.” *The Economic Journal*, Vol. 90, September 1980.
- [6] Chamley, Christophe. “Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives.” *Econometrica*, Vol. 54 (3), May 1986.
- [7] Chari, V.V., Christiano, Lawrence J. and Kehoe, Patrick J. “Optimal Fiscal Policy in a Business Cycle Model.” *Journal of Political Economy*, Vol. 102 (4), August 1994.
- [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] Corlett, W.J. and Hague, D.C. “Complementarity and the Excess Burden of Taxation.” *The Review of Economic Studies*, Vol. 21 (1), 1953.
- [10] Diamond, Peter A. “National Debt in a Neoclassical Model.” *American Economic Review*, Vol. LV, (no 5, part 1), 1965.
- [11] Erosa, Andrés and Gervais, Martin. “Optimal Taxation in Life-Cycle Economies.” Federal Reserve Bank of Richmond WP-02, 2000.
- [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] 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.
- [15] Jones, Larry E., Manuelli, Rodolfo E. and Rossi, Peter E. “Optimal Taxation in Models of Endogenous Growth.” *Journal of Political Economy*, Vol. 101 (3), June 1993.
- [16] Judd, Kenneth L. “Redistributive Taxation in a Simple Perfect Foresight Model.” *Journal of Public Economics*, Vol. 28, 1985.
- [17] Klein, Paul and Ríos-Rull, José Víctor. “Time-Consistent Optimal Fiscal Policy.” Manuscript, 1999.
- [18] Kotlikoff, Laurence J. “The A-K Model: It’s Past, Present, and Future.” NBER Working Paper 6684, August 1998.

- [19] Krusell, Per, Quadrini, Vincenzo and Ríos-Rull, José Víctor. “Are Consumption Taxes Really Better Than Income Taxes?” *Journal of Monetary Economics*, Vol. 37 (3), June 1996.
- [20] Kydland, Finn E. and Prescott, Edward C. “Rules Rather than Discretion: The Inconsistency of Optimal Plans.” *Journal of Political Economy*, Vol. 85, 1977.
- [21] Ramsey, Frank P. “A Contribution to the Theory of Taxation.” *Economic Journal*, Vol. 37, 1927.
- [22] Samuelson, Paul A. “The Two-Part Golden Rule Deduced as the Asymptotic Turnpike of Catenary Motions.” *Western Economic Journal*, Vol. 6, 1968.

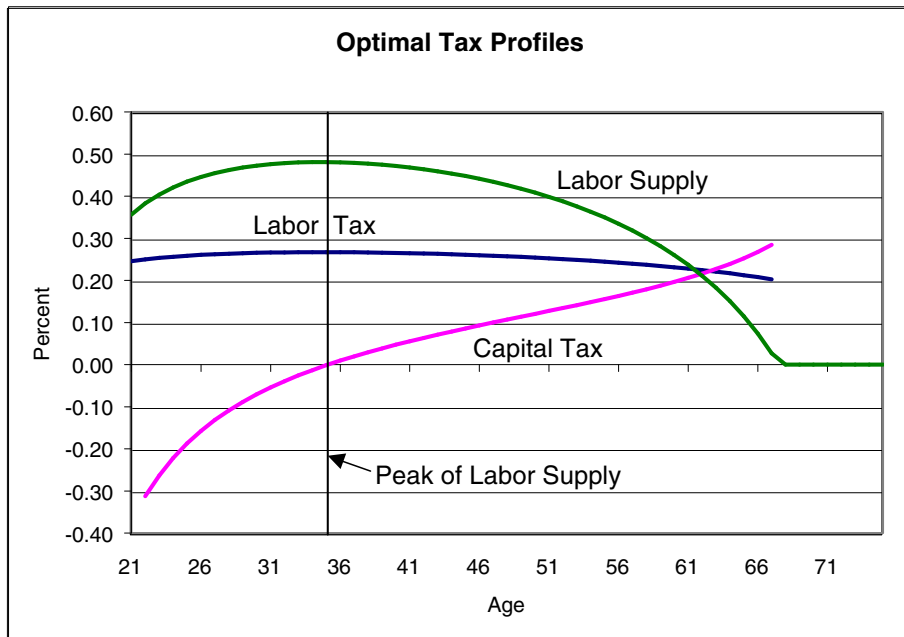


Figure 1: Optimal tax rates over the lifetime of individuals

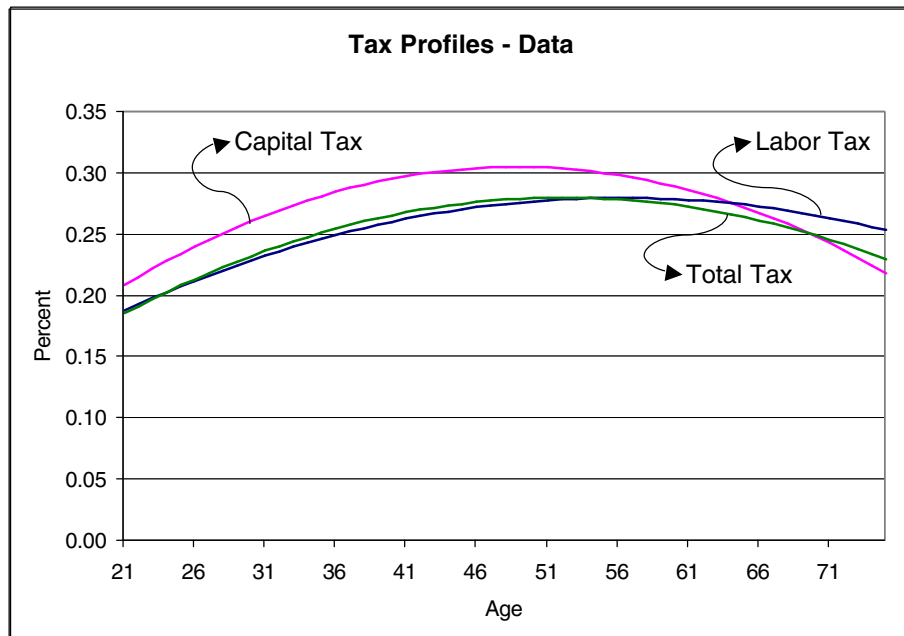


Figure 2: Tax rates over the lifetime of individuals implied by the data

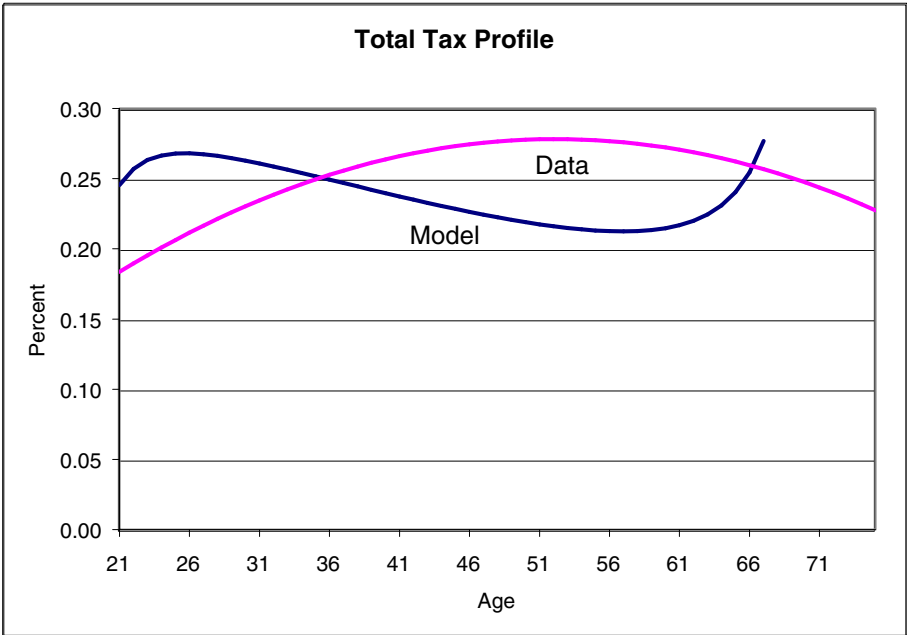


Figure 3: Total income tax rates over the lifetime of individuals

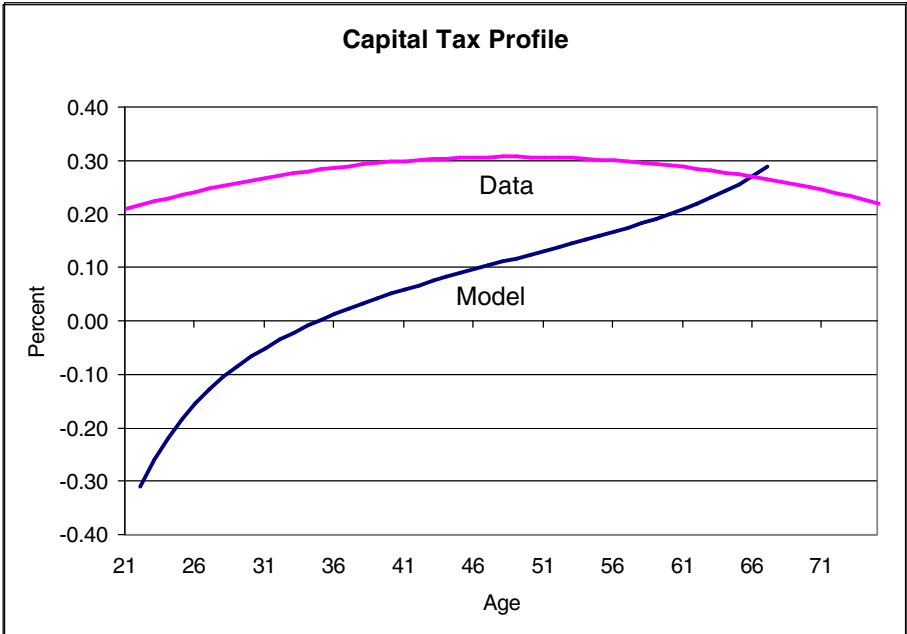


Figure 4: Capital income tax rates over the lifetime of individuals

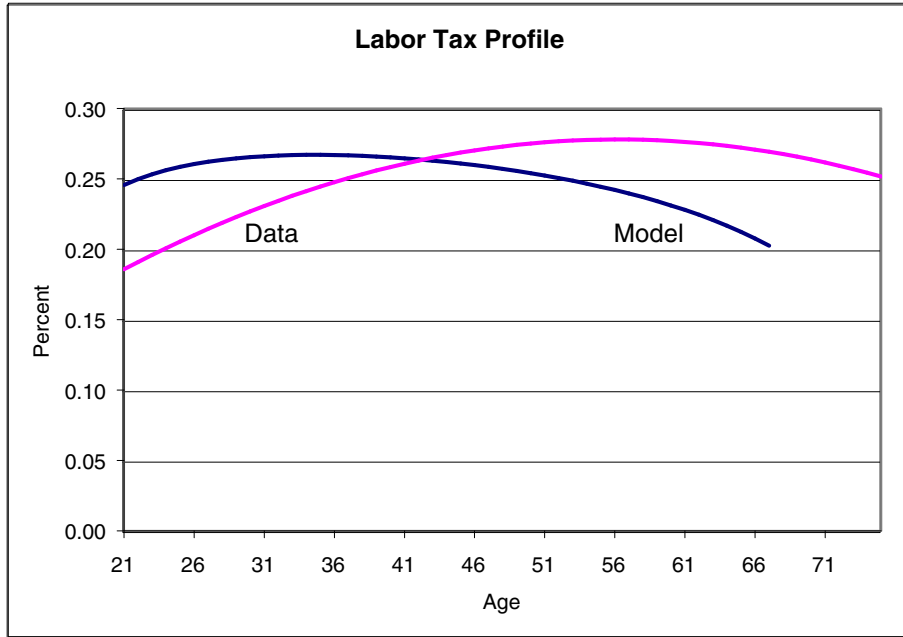


Figure 5: Labor income tax rates over the lifetime of individuals

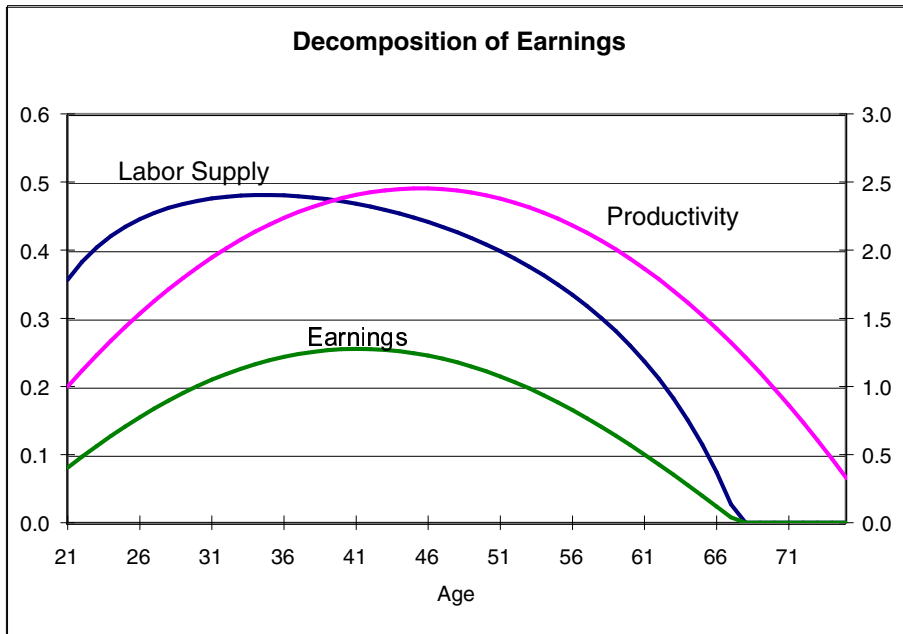


Figure 6: Labor earnings decomposition over the lifetime of individuals