Optimal Age Specific Income Taxation

Jean-Marie Lozachmeur[1]

October, 2005

[1]GREMAQ, University of Toulouse. Mailing: Gremaq, 21 Allée de Brienne F-31000 Toulouse. E-mail: jean-marie.lozachmeur@univ-tlse1.fr. I whish to thank four anonymous referees and the associate editor for their very helpful suggestions. Comments from Alain Jousten, Maurice Marchand and Pierre Pestieau are gratefully acknowledged.

Abstract

This paper studies optimal earnings taxation in a three period life-cycle model where taxes can be differentiated according to age. Agents choose their level of education when young and their retirement age when old. I study the problem both without and with borrowing constraints. It is shown that, without borrowing constraints, a first best optimum can be decentralized by setting a zero tax rate in the third period and a first period tax lower than the second period one. With borrowing constraints, the first best can no longer be achieved. The gap between the first and second period tax rates is larger, while the third period tax rate is generally different from zero.

JEL Classification Numbers: D91, H21

Keywords: Optimal Income Taxation, Life-cycle, Borrowing Constraints.

1 Introduction

Should tax rates on labor earnings vary over the life-cycle? In reality, statutory tax rates are usually independent of age. However, there are several reasons why the *effective* tax rate on labor earnings is not constant over the life-cycle. Apart from exemptions for young taxpayers, tax rates vary with age due to the progressiveness of the tax system combined with the increasing earning profile (see Gervais 2001). If the tax system is progressive, 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 workers also vary with age. Another reason comes from the tax/social security scheme targeted toward the old (see Gruber and Wise 1999). Above a certain age, workers face a double loss when deciding to continue their activity, namely the additional contribution and the foregone benefit stemming from the social security scheme.

A natural question concerns the optimality of such a tax system. A first rationale for age specific earning tax rates is related to the optimum tax theory initiated by Mirrlees (1971) and the use of categorical transfers (see Akerlof 1978, Immonen et al. 1998 and Viard 2001). When the government cannot observe ability but can observe an exogenous characteristic which is correlated with ability, it can improve the distributional effects of the tax system using lump sum grants. But it can also improve the efficiency of the tax system by conditioning the marginal income tax rates on the observable characteristic. This last criterion is analyzed in Kremer (1997) in the context of age specific income taxation. He concludes that young workers should optimally face lower marginal tax rates than older workers. Another rationale concerns the optimal Ramsey taxation within a standard life-cycle growth model in which agents' tastes for consumption and leisure change with age. This feature implies a non constant consumption-leisure choice for the representative agent, even in a steady-state economy. In this context, age specific income and consumption taxes are always optimal (Erosa and Gervais 2002).[1]

The aim of the present paper is also to study Ramsey-optimal age-dependent income tax rates. The analysis thus abstracts from any (intra-cohort) redistributive aspects of the tax system. The originality of the analysis comes from the introduction of some important life-cycle decisions. In Erosa and Gervais (2002), individuals'only choice variables are savings and per period labor supply. The present study incorporates important life-cycle decisions such as education and the retirement age. It also allows for incorporation of borrowing constraints. This gives rise to age-dependent tastes for consumption and leisure that are explained by life-cycle decisions.

The framework is a simple representative agent model. The individual lives for three periods of equal length. He chooses the amount of savings in the first and second periods, and the fraction of time devoted to education in

[1] The tax rates paths depend on the earnings profile and preferences (see Erosa and Gervais 2002).

the first period. He chooses when to retire during the third period, whereas he inelastically supplies one unit of labor in the second period. Consequently, this model assumes an endogenous *extensive* labor supply (i.e. the duration of a career) but a non-elastic *intensive* labor supply (i.e. weekly or yearly labor supply). The government chooses the age-specific flat income tax rates so as to finance an exogenous amount of public expenditure by maximizing the life-cycle utility of the representative agent.

The approach allows one to address different efficiency issues that are often discussed in the literature. Two important efficiency issues that arise in life-cycle models are related to the decisions to enter and to leave the labor market. When education is taken into account, the decision to enter the labor market is subject to an opportunity cost, namely the trade-off between higher future income due to the returns of education and the loss of income due to the postponement of entering the labor market. Additionaly, above a certain age, the individual faces a double loss when working an additional year, namely income taxes and foregone retirement benefits. This double taxation creates an implicit tax on continued activity which is negatively correlated with the labor force participation of the elderly (see Gruber and Wise 1999). This positive result is often used to advocate reforms tending to remove the bias in the benefit formulas. This raises the question of whether a bias in the benefit formula in favor of early retirement is necessarily the sign of a bad policy.[2] Another issue is the existence of borrowing constraints. If the capital market is imperfect, there may exist borrowing constraints especially when workers are young, and the tax system is likely to affect the strength of this constraint, which will in turn affect the individual's decisions. An age-specific tax rate could at least partially eliminate this imperfection by imposing some transfers between generations.

I start with the benchmark case of no borrowing constraints and show that the availability of age-dependent income tax rates leads to a first best outcome. The first period tax rate should be positive but lower than the second period one. The third period tax rate should be zero, which is compatible with a marginally fair social security system. With borrowing constraints, the gap between the first and the second period tax rates is higher and the third period tax is different from 0. This last property comes from the introduction of both endogenous retirement age and education. If these two variables are positively related, then the third period tax should be negative. Otherwise, it should be positive.

2 The model

The model is a simple partial equilibrium, three period life-cycle model. I Assume zero interest and discount rates and a zero population growth

[2] See Cremer et al. (2004) who deal with this issue in an optimal redistributive income taxation framework.

rate.[3] In period 1, the individual supplies a fraction $1-e$ units of labor and spends a fraction e in education while saving s_1. In period 2, he inelastically supplies one unit of labor and saves s_2. In the third period, the individual chooses his retirement date l, i.e, the fraction of this period during which he remains at work. His consumption is financed by his earnings and the return from his savings. The real wage w is an increasing and concave function of education:

$$w = w(e), w'(e) > 0, w''(e) < 0.$$

Letting c_i denote period i consumption, the life-cycle utility function $U(c_1, c_2, c_3, l)$ is assumed to be separable between periods and between consumption and leisure:

$$U(c_1, c_2, c_3, l) = u(c_1) + u(c_2) + u(c_3) + v(1-l),$$

where $u(.)$ and $v(.)$ satisfy the usual properties: $u'(.), v'(.) > 0, u''(.), v^{''}(.) < 0$ where $u'(.)$ and $u''.)$ denote the first and second derivatives of $u(.)$ respectively. The same holds for $v'(.)$ and $v^{''}(.)$. Finally, $\sigma(c_i)$ denotes the elasticity of the marginal utility that is, $\sigma(c_i) = -c_i u''(c_i)/u'(c_i)$, which is also the inverse of the elasticity of intertemporal substitution.

[3]Implicitely, we assume a stationnary equilibrium within an overlapping generation model. With this assumption, taxing an individual over his life cycle or taxing a society at a given point in time are identical problems.

This model implies that labor is endogenous in the first and the third period. The assumption of inelastic labor in period 2 may seem somewhat restrictive, but labor elasticity should be thought of in terms of “time spent in the labor force” rather than “work hours per week” supplied. I come back to this in the conclusion.

The following section studies the individual’s problem for given levels of tax rates. Section 3.1 starts by studying the problem without liquidity constraint and identifies the conditions under which the individual chooses a positive amount of savings in the first period. Section 3.2 presents the constrained problem and some resulting comparative statics which are useful for the rest of the analysis. Section 4 presents the problem facing the government.

3 The problem of the individual

The individual determines the level of education e, first and second period savings, s_1 and s_2, and the retirement age, l, maximizing his life-cycle utility under the constraint that first period savings must be non negative. I assume throughout the paper that s_2 is always positive, i.e that the second period borrowing constraint is not binding. Letting t_i denote the period i proportional income tax rate, the problem facing the individual can be written as follows:

$$\begin{aligned}
\underset{e,l,s_1,s_2}{\text{Max}} \quad & u(c_1) + u(c_2) + u(c_3) + v(1-l), \\
\text{s.to: } c_1 &= (1-e)(1-t_1)w(e) - s_1, \\
c_2 &= (1-t_2)w(e) - s_2 + s_1, \\
c_3 &= l(1-t_3)w(e) + s_2, \\
s_1 &\geq 0.
\end{aligned}$$

Letting μ denote the Lagrangian multiplier associated with the borrowing constraint on s_1, the first order conditions with respect to e, l, s_1 and s_2 for an interior solution are:

$$[w'(e)(1-e) - w(e)](1-t_1)u^{'}(c_1) +$$

$$w'(e)(1-t_2)u^{'}(c_2) + w'(e)l(1-t_3)u^{'}(c_3) = 0, \tag{1}$$

$$w(e)(1-t_3)u^{'}(c_3) - v^{'}(1-l) = 0, \tag{2}$$

$$u'(c_2) - u^{'}(c_1) + \mu = 0, \tag{3}$$

$$u^{'}(c_2) - u^{'}(c_3) = 0, \tag{4}$$

$$\mu \geq 0.$$

Equation (1) states that the marginal utility loss due to the foregone income in the first period must equal the marginal utility gain from the higher wage in the future. Equation (2) implies the usual equality between the marginal rate of substitution between consumption and leisure and net labor productivity. Because s_2 is assumed to be positive, the individual always sets $c_2 = c_3$.

3.1 The unconstrained problem

Let me first consider the case $\mu = 0$ in the individual's problem, that is, either the capital market is perfect, or the individual chooses a non negative s_1. I first present the first-order conditions determining the optimal choices of e and l and then infer the conditions for which these choices lead to non negative savings in the first period. From (1) to (2), the optimal choices of e and l are given by the two following equations:

$$\begin{aligned} & w'(e)\left[2-e+l\right]-w(e) \\ = \; & t_1\left[(1-e)w'(e)-w(e)\right]+w'(e)\left[t_2+lt_3\right], \quad (5) \\ & w(e)(1-t_3)u^{'}(c)-v^{'}(1-l)=0, \quad (6) \end{aligned}$$

where $c = c_i \equiv 1/3[lw(e)(1-t_3)+w(e)(1-t_2)+(1-e)w(e)(1-t_1)]$. Given that interest and discount rates are zero, the individual chooses to equate consumption $c_i = c$ across the three periods. (5) states that the individual chooses the amount of education that maximizes his *life-cycle income.*

It is easy to show that the optimal first period savings are given by:

$$s_1 = \frac{2}{3}(1-e)(1-t_1)w(e) - \frac{1}{3}(1-t_2)w(e) - \frac{1}{3}lw(e)(1-t_3).$$

Using the first order condition (5), this can be rewritten as:

$$s_1 = \frac{w(e)(1-t_1)}{3w'(e)}\left[3w'(e)(1-e) - w(e)\right].$$

Defining e_c as the level of education such that $3w'(e_c)(1-e_c) - w(e_c) = 0$, s_1 is negative if $e > e_c$. When the individual chooses $e > e_c$, the first period income is lower than future income since labor supply is small relative to that in the second and third period. This makes the individual borrow from future period income.

3.2 The constrained problem

Assume now that $e > e_c$ in the unconstrained problem. The first order condition for e becomes:

$$\begin{aligned} U_e &= (1-t_1)[w'(e)(1-e) - w(e)]u^{'}(c_1) + \\ & \quad w'(e)\left[(1-t_2) + l(1-t_3)\right]u^{'}(c) = 0, \end{aligned} \tag{7}$$

with

$$c_1 = (1-e)w(e)(1-t_1),$$
$$c = c_2 = c_3 = \frac{1}{2}w(e)\left[(1-t_2)+l(1-t_3)\right],$$

where U_e denotes the first-order derivative of life-cycle utility U with respect to e. Note that since the constraint $s_1 \geqslant 0$ is binding, $c_1 < c$. The main difference of this setup is that the individual is not able to smooth consumption between the first period and the later ones. The system formed by (7) and (6) defines the individual decision variables as a function of the tax rates: $e = e^*(t_1, t_2, t_3)$ and $l = l^*(t_1, t_2, t_3)$.

Define the optimal level of education given in (7) as a function of the $t_i's$,and *for a given* retirement age $l : \overline{e}(t_1, t_2, t_3, l)$. In the government's problem, the following comparative statics will be useful:

$$\frac{d\overline{e}}{dl} = -\frac{w'(e)(1-t_3)u'(c)(1-\sigma(c))}{U_{ee}} \gtreqless 0 \text{ if } \sigma(c) \lesseqgtr 1,$$
$$\frac{d\overline{e}}{dt_1} = \frac{[w'(e)(1-e)-w(e)]\left[1-\sigma(c_1)\right]u'(c_1)}{U_{ee}} \gtreqless 0 \text{ if } \sigma(c_1) \lesseqgtr 1,$$
$$\frac{d\overline{e}}{dt_2} = \frac{w'(e)u'(c)\left[1-\sigma(c)\right]}{U_{ee}} \gtreqless 0 \text{ if } \sigma(c) \gtreqless 1,$$

where U_{ee} is strictly negative.

The sign of these expressions depends upon $\sigma(c)$ or $\sigma(c_1)$. To understand this, observe in equation (7) that increasing l has a positive substitution effect on e since it increases the marginal gain of income in period 3 by $w'(e)(1-t_3)$. It also has a negative income effect since it increases disposable income in period 3 by $w(e)(1-t_3)$. Recall that $\sigma(c)$ is a measure of the utility cost of variability in consumption. When $\sigma(c) > 1$, the income effect dominates since this utility cost is relatively high so that e reacts negatively to an increase in l. Otherwise, the substitution effect dominates and e reacts positively to an increase in l.

Similarly increasing t_1 decreases the marginal loss of income in period 1 by $-(w'(e)(1-e) - w(e))$. It also has a negative income effect since it decreases period 1 disposable income by $w(e)(1-e)$. The positive (resp. negative) effect dominates if $\sigma(c_1)$ is lower (resp. higher) than 1. The effect of t_2 on e also includes a substitution and an income effect whose signs are, however, opposite to those of the effects of a rise in t_1. An increase in t_2 has a negative substitution effect since it decreases the marginal gain of income in period 2 by $w'(e)$. It has also a positive income effect since it decreases period 2 disposable income by $w(e)$. Again, the income effect will dominate if $\sigma(c) > 1$.

4 The problem of the government

The government needs to finance an exogenous per capita amount of public expenditure R_0 with flat-rate income taxes and maximizes the representative individual's indirect utility function. The budget constraint of the government is:

$$R_0 \equiv R^c = w(e)\left[(1-e)t_1 + t_2 + lt_3\right], \tag{8}$$

where R^c defines the total tax revenue. Letting $W(t_1, t_2, t_3)$ denote the indirect utility function of the individual, the government solves:

$$\underset{t_1,t_2,t_3}{\text{Max}}\, W(t_1, t_2, t_3) + \lambda\left[R^c - R_0\right].$$

In a first part, I describe the solution of the government's problem when the individual is not constrained. I turn to the constrained problem of the government in section 4.2.

4.1 The non constrained problem of the government

When the borrowing constraint is not binding, the availability of age-specific tax rates allows the government to implement the first best allocation. To see this, note that such an allocation can be decentralized by setting (t_1, t_2, t_3) such that:

$$w'(e)\left[t_1(1-e) + t_2 + lt_3\right] - w(e)t_1 = 0, \tag{9}$$

$$t_3 = 0, \tag{10}$$

$$R^c = R_0.$$

where $e \leq e_c$, so that the individual is not constrained. As can be inferred by (5), equation (9) is the condition for education to be not distorted. Clearly, the decision to invest in education is not distorted as long as the tax bill on additional earnings associated with education (first term) is equal to the subsidies to this investment (second term).

Note also for future reference that (9) and (10) imply that the individual's decision variables have no marginal effect on tax revenues, that is

$$\frac{\partial R^c}{\partial e} = w'(e)\left[t_1(1-e)+t_2+lt_3\right]-w(e)t_1=0, \tag{11}$$

$$\frac{\partial R^c}{\partial l} = w\left(e\right)t_3=0. \tag{12}$$

Combining equation (9) and the fact that the first best education is described by $w'(e)\left[2+l-e\right]=w(e)$, it is straightforward to show that:

$$t_1(1+l)=t_2+lt_3.$$

This equation means that the first period tax must be equal to the average income tax in the second and the third period. A uniform proportional tax fulfills this condition but since the third period income should not be taxed, the tax rates will be differentiated according to the periods. Using the budget constraint of the government (8), one is now able to characterize the optimum:

Proposition 1 *When there is no (binding) borrowing constraint, the use of age dependent income tax rates leads to a first best optimum with:*

$$t_1 = \frac{R_0}{w(e)(2+l-e)} = \frac{t_2}{(1+l)},$$
$$t_3 = 0.$$

where $e \leq e_c$*.*

The first period tax is then the ratio of the per capita public expenditure to the gross working-life wage income. The second period tax will never be lower than the first period tax. The special case of a uniform taxation between the first and the second period arises when the fraction of time devoted to work in the third period is 0. The fact that the first period tax is lower than the second one comes from the introduction of an endogenous retirement age. Because it is optimal not to tax the third period income since it would distort the choice of l, income that is not taxed in this period will be taxed in the second period in order not to distort education.

4.2 The constrained problem of the government

Assume from now on that $e > e_c$, where e is the optimal level of education chosen by the individual as a function of the tax rates given in proposition 1. These tax rates imply that the individual is constrained so that they are not optimal. In the appendix, it is shown that the constrained optimum yields the two following equations:

$$w(e)(1-e)\left(\frac{u'(c_1)}{u'(c)}-1\right) = -\min\left[0, \frac{\partial R^c}{\partial e}\left(\frac{(1-e)u'(c_1)}{u'(c)}\frac{d\overline{e}}{dt_2}-\frac{d\overline{e}}{dt_1}\right)\right] \quad (13)$$

$$\frac{\partial R^c}{\partial e}\frac{d\overline{e}}{dl}+\frac{\partial R^c}{\partial l}=0, \quad (14)$$

where $\partial R^c/\partial e > 0$ and $\partial R^c/\partial l$ are given respectively by equations (11) and (12).

To understand the intuition behind equation (13), assume one starts from the optimal tax rates as given in proposition 1, where $\partial R^c/\partial e = 0$ as shown in (11). From this point, an increase in t_2 combined with a welfare balanced decrease in t_1 increases tax revenue by the LHS of (13). In other words, reducing the gap between c_1 and c by a decrease of t_1 compensated by an increase in t_2 increases tax revenues. Not surprisingly, this implies that the gap between the first and the second period tax rates is larger with borrowing constraints. The LHS is thus the direct tax revenue gain from an increase in t_2 combined with a welfare balanced decrease of t_1.

The second term in brackets of the RHS of (13) represents the tax revenue variation induced by the same (compensated) change *via the effect on education.* If this change represents a loss of resources (in which case this term is negative), one has a simple trade-off between this loss and the marginal gain represented by the LHS. Otherwise, if this change represents a gain (in which case the term is positive), it is always optimal to increase t_2

and decrease t_1 until the intertemporal equality of consumptions is restored.

Therefore, a necessary condition for having $c_1 = c$ at the optimum is that the second term in the RHS of (13) is positive. This inequality would imply that an increase in t_2 combined with *a welfare balanced* decrease of t_1 has a positive effect on tax revenue *via* the effect on education. Since optimality necessarily requires that $\partial R^c/\partial e > 0$,[4] this regime occurs if a decrease in t_1 combined with a *welfare balanced* increase in t_2 has a positive effect on the level of education:

Proposition 2 *With borrowing constraints, the optimum involves $c_1 = c$ if and only if:*

$$(1-e)\frac{d\overline{e}}{dt_2} - \frac{u'(c)}{u'(c_1)}\frac{d\overline{e}}{dt_1} \geq 0,$$

which is equivalent to

$$w'(e)(1-e)\left(1-\sigma(c)\right) - \big(w'(e)(1-e) - w(e)\big)\left(1-\sigma(c_1)\right) \geq 0. \qquad (15)$$

Before going further, consider the simple case where the utility function is CES. In this case, $\sigma(c_i) = \sigma$ so that inequality (15) is equivalent to $\sigma < 1$. In this case, increasing t_2 and decreasing t_1 decrease the level of education, which in turn decrease tax revenues. Consequently, there is a

[4] Assume instead that $\partial R^c/\partial e < 0$ with $c_1 < c$. Then a marginal decrease in e would create a budget surplus and relax the borrowing constraint. This contradicts the fact that it is an optimum.

tradeoff between reducing the gap between c_1 and c (which increases tax revenue) and the tax revenue loss implied by the decrease in the level of education. However, if $\sigma \geq 1$, reducing the gap between c_1 and c by means of a compensated change in t_1 and t_2 does not decrease the level of education, which in turn does not decrease tax revenues. In this case, intertemporal equality of consumption occurs at the optimum.

In the more general case of a non CES utility function, one can make the same reasoning when both $\sigma(c_1)$ and $\sigma(c)$ are lower (or higher) than 1. When $\sigma(c_1) > 1$ and $\sigma(c) < 1$,[5] matters get more complicated since a rise in t_2 has a negative effect on education (the substitution effect dominates) and a decrease in t_1 has a positive effect on education (the income effect dominates). The total effect of a welfare balanced change of t_1 and t_2 on education then depends upon the difference between $\sigma(c_1)$ and $\sigma(c)$ but also upon the relative returns to education in period 1: if the relative returns to education in period 1 as measured by $\left[\left(w'(e)(1-e) - w(e)\right)/w'(e)(1-e)\right]$ are high in absolute value (i.e. an increase in education implies a high loss in period 1 income), the positive income effect on education due to the decrease in t_1 may dominate the negative substitution effect of t_2. In this case, reducing the gap between c_1 and c increases the level of education and condition (15) is fulfilled. A symmetric argument applies to the case where $\sigma(c_1) < 1$ and $\sigma(c) > 1$.

[5] Note that a necessary condition for this is that the elasticity of marginal utility is a decreasing function of consumption.

Equation (14) states that the retirement age has no marginal effect on tax revenue, taking into account its indirect effect on education. Since t_3 is the only tax with a distortive effect on l (t_1 and t_2 only have an income effect), the optimum involves the compensated effect of t_3 on l having no (marginal) effect on tax revenues. Remember that this inequality is fulfilled in the first best optimum since $\partial R^c/\partial e = \partial R^c/\partial l = 0$, as shown in (11) and (12). Now, these equalities are not sustainable since the individual is constrained.

It turns out that the third-period tax is generally different from zero. To see this, suppose one starts from a zero third-period tax rate. In this case, one has $\partial R^c/\partial l = w(e)\, t_3 = 0$. Increasing t_3 thus makes tax revenues increase if $(\partial R^c/\partial e)\,(-d\overline{e}/dl) > 0$, that is, if the decrease in the compensated retirement age makes education vary such that the tax revenues increase. Since one has $\partial R^c/\partial e > 0$, an increase in t_3 has a positive impact on tax revenues if $d\overline{e}/dl < 0$, that is, if education and retirement are negatively related. In the opposite case where education and the retirement age are positively related, it is optimal to decrease t_3 below zero.

The sign of t_3 then depends upon the sign of $d\overline{e}/dl$. Using the comparative statics, one has:

Proposition 3 *With borrowing constraints, the optimum leads to sign* $(t_3) =$ *sign* $[\sigma(c) - 1]$.

When $\sigma(c) = 1$, e does not directly depends upon l. In this case, there is no reason to deviate from a zero third period tax rate. When $\sigma(c) < 1$, consumptions are relatively high substitutes so that e and l are positively correlated. This means that a compensated decrease in l (due to an increase in t_3) indirectly decreases education and hence tax revenues. Starting from a zero third period tax rate, it is thus optimal to subsidize labor by decreasing the third period tax. This increases the compensated retirement age so that education increases ,which in turn increase tax revenues. However, when $\sigma(c) > 1$, an increase in t_3 reduces the compensated retirement age and positively affects the level of education which in turn increases tax revenues. In this case, a positive third period tax rate is desirable.

To sum up, the constrained optimum still leads to a first period tax that is lower than the second period one. The gap between these two tax rates is larger than without borrowing constraints. The third period tax is no longer equal to 0. This tax rate may be positive or negative following the negative or positive relation between education and the retirement age.

5 Conclusion

In this paper, I have considered a Ramsey tax problem in which it is possible to differentiate the flat earnings tax rates according to age. A three period model is built in which education when young and retirement when old are endogenous. Without borrowing constraints when young, the optimal tax rates involve a lower tax for the young than for the middle aged workers.

Labor earnings when old are not taxed. With borrowing constraints, the gap between the tax rate for the young and the middle aged workers is higher. Depending upon the values of the elasticities of marginal utilities and the relative return of education, it may be optimal to restore intertemporal equality of consumptions. The labor tax rate on the old may be either positive or negative depending upon the relation between retirement age and education. If both are positively related (which is the case if the elasticity of marginal utility is lower than one), the tax rate is negative. In the reverse case where they are negatively correlated (which is the case if the elasticity of marginal utility is higher than one), the tax rate is positive.

In order to tackle the problem without too many difficulties, the model used some simplifying assumptions. An important one is that of exogenous *intensive* labor supply. Relaxing this assumption would imply that a first best outcome would not be attainable with the age dependent tax rates. However, the qualitative properties of the model would remain unchanged. Since intensive labor supply would be endogenous in each period, each age-dependent tax rates would have an equivalent effect on each (per period) intensive labor supply. Therefore, the problem would just be rescaled. Another assumption is the one of separable utility functions. If the utility functions are not separable over time, the individual would choose to equalize marginal utilities of consumptions instead of equalizing consumptions. This would complicate the analysis without changing the main mechanisms of the model. In particular, the first best outcome would remain unchanged. With borrowing constraints, the important issue would remain the elasticity of substitution between consumptions.

This model could be extended in many ways. First, tax instruments are limited to wage tax rates. Extending the present analysis in a general equilibrium Ramsey problem with consumption taxes is on the research agenda. One may also introduce education expenses instead of treating the education decision as a mere trade off between foregone time and higher future earnings. Finally, the model could be extended to the case where one allows redistribution between agents of the same cohort. The tax rate on old individuals'labor earnings would of course be different from zero given the desirability to redistribute income within generations.

References

[1] Akerlof, G. A. (1978), The Economics Of Tagging As Applied To The Optimal Income Tax, Welfare Programs, And Manpower Planning, *American Economic Review 68, 8-19.*

[2] Cremer, H., J.M. Lozachmeur and P. Pestieau (2004), Social Security, Retirement Age and Optimal Taxation, *Journal of Public Economics 88, 11, 2259-2281.*

[3] Erosa, A. and M. Gervais (2002), Optimal Taxation In Life-Cycle Economies, *Journal of Economic Theory* 105, 338-369.

[4] Gervais, M. (2001), On the Optimality of Age-Dependent Taxes and the Progressive U.S. Tax System, Manuscript.

[5] Gruber, G. and D. Wise (1999), *Social Security and Retirement around the World*, The Chicago University Press, Chicago.

[6] Immonen, R., R. Kanbur, M. Keen, and M. Tuomala (1998), Tagging And Taxing: The Use Of Categorical And Income Information In Designing Tax/Transfer Schemes, *Economica 65, 179-192.*

[7] Kremer, M. (1997), Should taxes be independent of age?, NBER Working Paper 6304.

[8] Mirrlees, J.A. (1971), An exploration in the theory of optimum income taxation, *Review of Economic Studies 38*, 175-208.

[9] Viard, A.D. (2001), Optimal Categorical Transfer Payments: The Welfare Economics of Limited Lump-Sum Redistribution, *Journal of Public Economic Theory* 3, 483-500.

Appendix

A Comparative statics

In this particular problem, the individual chooses two variables. In analyzing the effect of t_i on these two decision variables, it is useful for the following to distinguish their direct effects from the indirect effect via the change in the other decision variable. Formally, $e^* (t_1, t_2, t_3)$ and $l^* (t_1, t_2, t_3)$ are given by the system (7) and (6). Define $\overline{e}(t_1, t_2, t_3, l)$ and $\bar{l}(t_1, t_2, t_3, e)$ as the level of education and the retirement age (for a given level of retirement age and education respectively) as given implicitly by (7) and (6). de^*/dt_i and dl^*/dt_i can be decomposed as follows:

$$\begin{aligned}
\frac{de^*}{dt_i} &= C\left[\frac{d\overline{e}}{dt_i} + \frac{d\overline{e}}{dl}\frac{d\bar{l}}{dt_i}\right], \\
\frac{dl^*}{dt_i} &= C\left[\frac{d\bar{l}}{dt_i} + \frac{d\bar{l}}{de}\frac{d\overline{e}}{dt_i}\right],
\end{aligned}$$

where $C > 0$[6] and $d\overline{e}/dt_i$ and $d\overline{l}/dt_i$ are direct uncorrected effects of t_i on e and l and obtained respectively by differentiation of (7) and (6). This gives:

$$\begin{aligned}
\frac{d\overline{e}}{dt_1} &= \frac{\left[w'(e)(1-e) - w(e)\right]\left[1-\sigma\left(c_1\right)\right]u^{'}(c_1)}{U_{ee}} \gtreqless 0 \text{ if } \sigma(c_1) \lesseqgtr 1, \\
\frac{d\overline{l}}{dt_1} &= 0, \\
\frac{d\overline{e}}{dt_2} &= \frac{w'(e)u^{'}(c)\left[1-\sigma(c)\right]}{U_{ee}} \gtreqless 0 \text{ if } \sigma(c) \gtreqless 1, \\
\frac{d\overline{l}}{dt_2} &= \frac{\left[\frac{1}{2}\left(w(e)\right)^2(1-t_3)u^{''}(c)\right]}{U_{ll}} > 0, \\
\frac{d\overline{e}}{dt_3} &= l\frac{d\overline{e}}{dt_2}, \\
\frac{d\overline{l}}{dt_3} &= \frac{d\overline{l}}{dt_2} + \frac{d\widetilde{\overline{l}}}{dt_3},
\end{aligned}$$

where $U_{ll} < 0$ is the second order derivative of the LHS of (6) with respect to l. $d\widetilde{\overline{l}}/dt_3 = w(e)u^{'}(c)/U_{ll} < 0$ is the compensated effect of t_3 on l.

[6] The corrective term $C = \left[1 - \frac{de}{dl}' \frac{dl}{de}'\right]^{-1} = \frac{\left[U_{ee}U_{ll} - U_{el}^2\right]}{U_{ee}U_{ll}}$ is greater than zero given that the second order conditions of the individuals problem holds.

B Derivation of expressions (14) and (13)

The government program leads to first order conditions with respect to t_1 t_2 and t_3 given by:

$$
\begin{aligned}
&-(1-e)w(e)u'(c_1)+\lambda\left[w(e)(1-e)+\frac{\partial R^c}{\partial e}\frac{de^*}{dt_1}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_1}\right]=0,\\
&-w(e)u'(c)+\lambda\left[w(e)+\frac{\partial R^c}{\partial e}\frac{de^*}{dt_2}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_2}\right]=0,\\
&-lw(e)u'(c)+\lambda\left[lw(e)+\frac{\partial R^c}{\partial e}\frac{de^*}{dt_3}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_3}\right]=0,
\end{aligned}
$$

where $\partial R^c/\partial e$ and $\partial R^c/\partial l$ are given by the left hand sides of (11) and (12). Combining first order conditions with respect to t_2 and t_3, one obtains:

$$\frac{\partial R^c}{\partial e}\left(l\frac{de^*}{dt_2}-\frac{de^*}{dt_3}\right)+\frac{\partial R^c}{\partial l}\left(l\frac{dl^*}{dt_2}-\frac{dl^*}{dt_3}\right)=0.$$

Decomposing direct and indirect effects, one has:

$$\frac{\partial R^c}{\partial e}\left(\begin{array}{c} l\frac{d\overline{e}}{dt_2}-\frac{d\overline{e}}{dt_3}\\ +\frac{d\overline{e}}{dl}\left(l\frac{d\overline{l}}{dt_2}-\frac{d\overline{l}}{dt_3}\right)\end{array}\right)+\frac{\partial R^c}{\partial l}\left(\begin{array}{c} l\frac{d\overline{l}}{dt_2}-\frac{d\overline{l}}{dt_3}\\ +\frac{d\overline{l}}{de}\left(l\frac{d\overline{e}}{dt_2}-\frac{d\overline{e}}{dt_3}\right)\end{array}\right)=0. \tag{A1}$$

By the comparative statics, one has:

$$
\begin{aligned}
l\frac{d\overline{e}}{dt_2} &= \frac{d\overline{e}}{dt_3},\\
\text{and } l\frac{d\overline{l}}{dt_2}-\frac{d\overline{l}}{dt_3} &= -\frac{d\widetilde{\overline{l}}}{dt_3}.
\end{aligned}
$$

Thus, $(A1)$ simplifies to:

$$\frac{\partial R^c}{\partial e}\left(\frac{d\bar{e}}{dl}\right)\frac{\delta\widetilde{\bar{l}}}{\delta t_3}+\frac{\partial R^c}{\partial l}\frac{d\widetilde{\bar{l}}}{dt_3}=0.$$

Factorizing by $\frac{d\widetilde{\bar{l}}}{dt_3}$ leads to equation (14). Similarly, combining the optimality conditions of t_1 and t_2, one has:

$$\begin{aligned}\frac{u'(c_1)(1-e)}{u'(c)}&\left(\frac{\partial R^c}{\partial e}\frac{de^*}{dt_2}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_2}+w(e)\right)\\ &-\left(\frac{\partial R^c}{\partial e}\frac{de^*}{dt_1}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_1}+(1-e)w(e)\right) = 0.\end{aligned}$$

Next, decompose the effect of the taxes on tax revenue via the decision variables. One has:

$$\begin{aligned}&\frac{\partial R^c}{\partial e}\frac{de^*}{dt_i}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_i}\\ =\ & C\frac{\partial R^c}{\partial e}\left(\frac{d\bar{e}}{dt_i}+\frac{d\bar{e}}{dl}\frac{d\bar{l}}{dt_i}\right)+C\frac{\partial R^c}{\partial l}\left(\frac{d\bar{l}}{dt_i}+\frac{d\bar{l}}{de}\frac{d\bar{e}}{dt_i}\right)\\ =\ & C\frac{d\bar{e}}{dt_i}\left(\frac{\partial R^c}{\partial e}+\frac{\partial R^c}{\partial l}\frac{d\bar{l}}{de}\right)+C\frac{d\bar{l}}{dt_i}\left(\frac{\partial R^c}{\partial e}\frac{d\bar{e}}{dl}+\frac{\partial R^c}{\partial l}\right)\\ =\ & C\frac{d\bar{e}}{dt_i}\left(\frac{\partial R^c}{\partial e}+\frac{\partial R^c}{\partial l}\frac{d\bar{l}}{de}\right),\end{aligned}$$

where (14) was used . Again using (14) for the expression in brackets, one can rewrite:

$$\frac{\partial R^c}{\partial e}\frac{de^*}{dt_i}+\frac{\partial R^c}{\partial l}\frac{dl^*}{dt_i}=C\frac{d\bar{e}}{dt_i}\frac{\partial R^c}{\partial e}\left(1-\frac{d\bar{e}}{dl}\frac{d\bar{l}}{de}\right).$$

Given that $C = \frac{1}{1-\frac{d\overline{e}}{dl}\frac{d\overline{l}}{de}}$, one finally obtains:

$$\frac{\partial R^c}{\partial e}\frac{de^*}{dt_i} + \frac{\partial R^c}{\partial l}\frac{dl^*}{dt_i} = \frac{\partial R^c}{\partial e}\frac{d\overline{e}}{dt_i}. \tag{A2}$$

Substituting (A_2) in:

$$\begin{aligned} \frac{u'(c_1)(1-e)}{u'(c)}\left(\frac{\partial R^c}{\partial e}\frac{de^*}{dt_2} + \frac{\partial R^c}{\partial l}\frac{dl^*}{dt_2} + w(e)\right) & \\ -\left(\frac{\partial R^c}{\partial e}\frac{de^*}{dt_1} + \frac{\partial R^c}{\partial l}\frac{dl^*}{dt_1} + (1-e)w(e)\right) &= 0. \end{aligned}$$

gives equation (13) , provided that $c_1 \leq c$.