Diskussionspapiere der DFG-Forschergruppe (Nr.: 3468269275):

Heterogene Arbeit: Positive und Normative Aspekte der Qualifikationsstruktur der Arbeit

Mathias Kifmann

Age-dependent taxation and the optimal retirement benefit formula

Dezember 2004
Age-dependent taxation and the optimal retirement benefit formula

Mathias Kifmann

Universität Konstanz
Fachbereich Wirtschaftswissenschaften
Fach D 136
78457 Konstanz
Germany

mail: mathias.kifmann@uni-konstanz.de
phone: +49-7531-88-2189
fax: +49-7531-88-4135

Zusammenfassung:

This paper presents a comprehensive view of life-time taxation including both explicit taxation through the general tax system and implicit taxation via the retirement benefit formula. Individuals are heterogeneous with respect to productivity. It is shown that the optimal structure of age-dependent taxation can be characterized by a generalized Ramsey formula. Furthermore, the paper derives the optimal retirement benefit formula in the presence of the general tax system.

JEL Klassifikation: H21, H55
Schlüsselwörter: optimal taxation, pay-as-you-go pension systems, implicit taxation, intra- and intergenerational equity
Download/Reference: http://www.wiwi.uni-konstanz.de/forschergruppewiwi/
Age-dependent taxation and the optimal retirement benefit formula

Mathias Kifmann∗†

December 2004

Abstract

This paper presents a comprehensive view of life-time taxation including both explicit taxation through the general tax system and implicit taxation via the retirement benefit formula. Individuals are heterogeneous with respect to productivity. It is shown that the optimal structure of age-dependent taxation can be characterized by a generalized Ramsey formula. Furthermore, the paper derives the optimal retirement benefit formula in the presence of the general tax system.

JEL-classification: H21, H55.

Keywords: optimal taxation, pay-as-you-go pension systems, implicit taxation, intra- and intergenerational equity.

∗I am grateful to Friedrich Breyer, Stefan Hupfeld and Dirk Schindler for helpful comments and suggestions. This research is part of the project “Lebensarbeitszeit, Humankapitalbildung und Transfersystem” funded by the Deutsche Forschungsgemeinschaft under grant BR 740/15-1.

†Universität Konstanz, Fach D 136, 78457 Konstanz, Germany; Email: Mathias.Kifmann@uni-konstanz.de.
1 Introduction

Pay-as-you-go pension schemes redistribute between generations if the economy is dynamically efficient. As Sinn (2000) has shown, transfers to previous generations are exactly equivalent to the implicit tax burden of the public pension system. In this context, the retirement benefit formula plays a key role. It establishes the link between contributions in each period and pension benefits and therefore determines how labor supply is implicitly taxed in each period.

To minimize the excess burden, age-dependent taxation through the retirement benefit formula should be optimally designed. Existing pension formulas, however, usually follow an actuarial logic and tax contributions when young higher than when old.¹ It is questionable whether such an implicit tax schedule is optimal from an incentive point of view. For example, Fenge, Übelmesser and Werding (2002) argue that income in early years should be taxed less because labor supply is more elastic in young age.

The framework to analyze the optimal design of the retirement benefit formula has so far been the standard Ramsey model with a representative agent. This approach does not consider that a key motivation for the use of distortionary taxes are unobservable differences in productivity of individuals. In this paper it is shown that the results have to be modified once heterogeneity of individuals is taken into account. In particular, the retirement benefit formula cannot only be based on age-dependent elasticities of labor supply. Furthermore, the analysis of the optimal retirement benefit formula has generally abstracted from the general tax system. This paper extends the analysis and provides a comprehensive view of lifetime taxation including both explicit taxation through the general tax system and implicit taxation via the retirement benefit formula.

The following Section 2 places the analysis in the context of the literature. In Section 3, the model is presented. Section 4 shows how the optimal tax problem can be solved in two steps. In Section 5, a generalized Ramsey rule for age-dependent taxation is derived. Section 6 discusses the implications for the design of the optimal retirement benefit formula. A final section concludes.

2 Literature review

Wrede (1999) was the first to analyze optimal age-dependent taxation through the retirement benefit formula and to show the equivalence to a standard Ramsey problem. His study was extended by Fenge et al. (2002) who derive the elasticity formula for optimal age-dependent taxation and estimate labor supply elasticities for different age groups. Assuming that the compensated elasticity of labor supply with respect to wages in other periods is zero, they find that the German pension formula is taxing income in early years too heavily. Instead, the optimal time structure of implicit taxes should follow an inversely “J-shaped” pattern, i.e. income in middle age should be taxed most heavily.

Another strand of the literature has examined age-dependent taxation without considering the link to the retirement benefit formula. Here the conclusion is that a tax system should generally be conditioned on age. In a standard life-cycle growth model, this result is obtained by Erosa and Gervais (2002) who solve the Ramsey problem of age-dependent taxation. They show that an optimizing government will generally use capital and labor income tax rates which vary with age. In a three period life-cycle model, Lozachmeur (2002) draws a similar conclusion. In his approach, a representative agent chooses the level of education when young and retirement age when old. The government sets a tax rate for labor income in each period. He even finds that the first best optimum can be reached by age-dependent taxation if the young can borrow. This solution, however, may not be possible if the young face borrowing constraints.

The papers by Erosa and Gervais (2002) and Lozachmeur (2002) assume a uniform tax rate for each age-period. A different approach is taken by Kremer (2001) who adapts the optimal tax model by Mirrlees (1971) to take into account differences in age. He shows that the government can improve the efficiency of the tax system by conditioning the marginal income tax rates on age. For example, he finds that young workers have a more elastic labor supply than prime-age workers and that income when young is relatively uncorrelated with income later in life. These age-dependent characteristics lead him to the conclusion that marginal tax rates should be lower for the young.
All papers mentioned look in isolation at either the retirement benefit formula or at the general tax system. However, Bütler (2002) shows that both ways of taxation can in principle be analyzed within one framework. Furthermore, all papers with the exception of Kremer (2001) assume that individuals are homogeneous or analyze an economy with a representative agent. Yet the second-best problem of income taxation arises mainly because individuals differ in productivity which cannot be observed by the government.

In these two aspects, the present paper differs from the existing literature. On the one hand, it characterizes the optimal retirement benefit formula in the presence of the general tax system. On the other hand, the paper considers a world in which individuals are heterogeneous with respect to productivity. Here the paper stays within the framework of Wrede (1999), Fenge et al. (2002) and Erosa and Gervais (2002) and assumes a uniform tax rate for each age-period. A generalized Ramsey formula is derived and contrasted with the findings of these papers.

3 The model

3.1 The economy

As Wrede (1999) and Fenge et al. (2002), I analyze a small open economy. The constant interest rate is $r^2$. The framework is an overlapping generations model in which individuals live for three periods. The population growth rate in period $t$ is $m_t$. Furthermore, I assume in the following that the interest rate is always higher than the growth rate of total wages. This rules out a Pareto-improvement by introducing a PAYG pension scheme or public debt.

3.2 The individual problem

Individuals work in the first two periods and are retired in the last period. Preferences of an individual born in period $t$ are described by the utility function

---

2The assumption of a constant interest rate is made to keep the exposition and notation as simple as possible and does not affect the main results of this paper. In footnotes 4 and 7 I point out possible changes if the interest rate is not constant.
With a perfect capital market, the problem an individual born in period \( t \) with productivity \( z \) is

\[
U (c_{1,t}, c_{2,t+1}, c_{3,t+2}, l_{1,t}, l_{2,t+1})
\]

where \( c_{a,t} \) and \( l_{a,t} \) are consumption and labor supply. The subscript \((a,t)\) denotes the individual’s “age” \( a = 1, 2, 3 \) in period \( t \). Labor income in period \( t \) is equal to \( w_t z l_{a,t} \) where \( w_t \) is the wage rate per efficiency unit in period \( t \) and \( z \) is the productivity of the individual. Individuals differ in productivity \( z \in \{z_l, z_r\} \) which is distributed according to the distribution function \( F_t(z) \) for individuals born in period \( t \). In the following, I assume that neither \( z \) nor labor supply \( l_{a,t} \) are observable.

The age-dependent tax rate on labor income is \( \tau_{a,t}^l \). Furthermore, each individual receives a tax-financed basic income \( B_{a,t} \) in each period of her life \( a = 1, 2, 3 \). Individuals must contribute to the public pension system when working. The contribution rate to the pension system in period \( t \) is \( b_t \). The pension benefit for individuals born in period \( t \) during retirement depends on contributions through the formula

\[
p_{t+2} = \alpha_{t+2} + \beta_1 b_tw_t z l_{1,t} + \beta_2 b_{t+1} b_{t+1} w_{t+1} z l_{2,t+1} \tag{1}
\]

where \( \alpha_{t+2} \) is the minimum pension and \( \beta_{a,t} \) is the return factor on contributions in period \( a \) of the individual’s life.

With a perfect capital market, the problem an individual born in period \( t \) with productivity \( z \) is

\[
\max_{c_{1,t}, c_{2,t+1}, c_{3,t+2}, l_{1,t}, l_{2,t+1}} U (c_{1,t}, c_{2,t+1}, c_{3,t+2}, l_{1,t}, l_{2,t+1}) \tag{2}
\]

subject to the intertemporal budget constraint

\[
c_{1,t} + \frac{c_{2,t+1}}{1 + r} + \frac{c_{3,t+2}}{(1 + r)^2} = (1 - b_t - \tau_{1,t}^l) w_{1,t} z l_{1,t} + B_{1,t}
\]
\[
+ \frac{(1 - b_{t+1} - \tau_{2,t+1}^l) w_{2,t+1} z l_{2,t+1} + B_{2,t+1}}{1 + r}
\]
\[
+ p_{t+2} + B_{3,t+2}
\]
\[
(1 + r)^2.
\]

Using the pension formula (1), this budget constraint can be simplified to

\[
c_{1,t} + \frac{c_{2,t+1}}{1 + r} + \frac{c_{3,t+2}}{(1 + r)^2} = (1 - \tau_{1,t}) w_{1,t} z l_{1,t} + \frac{(1 - \tau_{2,t+1}) w_{2,t+2} z l_{2,t+2}}{1 + r} + G_t. \tag{3}
\]
Here the total tax rate on labor income, \( \tau_{1,t} \), is defined by

\[
\tau_{1,t} \equiv \tau_{1,t}^l + \tau_{1,t}^b 
\]

and

\[
\tau_{2,t+1} \equiv \tau_{2,t+1}^l + \tau_{2,t+1}^b + \tau_{2,t+1}^b,
\]

where the indirect labor tax rates \( \tau_{a,t}^b \) through the pension system are given by

\[
\tau_{1,t}^b \equiv b_t \left( 1 - \frac{\beta_1}{(1+r)^2} \right) \quad \text{and} \quad \tau_{2,t+1}^b \equiv b_{t+1} \left( 1 - \frac{\beta_2}{1+r} \right).
\]

The present value of total transfers for an individual born in period \( t \) is given by

\[
G_t \equiv B_t + \frac{\alpha_{t+2}}{(1+r)^2},
\]

where \( B_t \) denotes the present value of tax-financed basic income

\[
B_t \equiv B_{1,t} + \frac{B_{2,t+1}}{1+r} + \frac{B_{3,t+2}}{(1+r)^2}.
\]

The budget constraint (3) shows that from the point of view of an individual only the total tax rates \( \tau_{a,t} \) and the total transfer \( G_t \) matter. Hence, utility maximization yields the indirect utility function

\[
V_t (\tau_{1,t}, \tau_{2,t+1}, G_t, w_t, w_{t+1}, z).
\]

### 3.3 Inter- and intragenerational welfare

Intragenerational welfare within generation \( t \) is captured by the social welfare function

\[
W_t = N_t \int \mathcal{Z} W (V_t (\tau_{1,t}, \tau_{2,t+1}, w_t, w_{t+1}, G_t, z)) dF_t(z).
\]

Furthermore, I assume that there is a concern for intergenerational welfare represented by the intergenerational welfare function

\[
\Psi = \sum_{s=0}^{\infty} \left( \frac{1}{1+\rho} \right)^s W_t
\]

where \( \rho > 0 \) is the intergenerational discount rate.

\[\text{See Büttler (2002, equation (4)) for a general formula for the total tax rate on labor income which allows for a arbitrary number of periods during which individuals work or are retired.}\]
3.4 Budget balance in the pension system

Pensions are organized on a pay-as-you-go basis. To facilitate the notation, I define
\[ Z_{a,t} = \int_{z} w_{a,t} z dF_{t}(z) \] \hspace{1cm} (11)
as the average income of individuals with age \( a \) in period \( t \). Then the condition for a balanced budget of the pension system in period \( t \) can be written as
\[ b_{t} (Z_{1,t}N_{t} + Z_{2,t}N_{t-1}) = \int_{z} p_{t} dF_{t-2}(z)N_{t-2} \] \hspace{1cm} (12)
\[ = (\alpha_{t} + \beta_{1,t-2}b_{t-2}Z_{1,t-2} + \beta_{2,t-1}b_{t-1}Z_{2,t-1})N_{t-2} \]
where \( N_{t} \) is the size of the generation entering the labor force in period \( t \). With \( N_{t} = (1 + m_{t-1})N_{t-1} \), the contribution rate is
\[ b_{t} = \frac{\alpha_{t} + \beta_{1,t-2}b_{t-2}Z_{1,t-2} + \beta_{2,t-1}b_{t-1}Z_{2,t-1}}{Z_{1,t}(1 + m_{t-1})(1 + m_{t-2}) + Z_{2,t}(1 + m_{t-1})} \] \hspace{1cm} (13)

3.5 Tax restrictions

The present value of labor income tax payments of generation \( t \) must finance tax-financed transfers \( B_{t} \) within their generation and transfers to other generations \( T_{t}^{d} \) in terms of a change in public debt (both in per capita terms):
\[ \tau_{1,t}^{b}Z_{1,t} + \frac{\tau_{2,t+1}^{b}Z_{2,t+1}}{1 + r} = B_{t} + T_{t}^{d} \] \hspace{1cm} (14)
In the pension system, individuals must finance the basic pension and implicit taxes \( T_{t}^{b} \) per person due to the gains of other generations, i.e.
\[ \tau_{1,t}^{b}Z_{1,t} + \frac{\tau_{2,t+1}^{b}Z_{2,t+1}}{1 + r} = T_{t}^{b} + \frac{\alpha_{t+2}}{(1 + r)^{2}} \] \hspace{1cm} (15)
The total tax per capita to be paid by generation \( t \), \( T_{t} \), is given by
\[ T_{t} = T_{t}^{b} + T_{t}^{d} \] \hspace{1cm} (16)
Combining (14), (15) and (16) and using the definition of $G_t$ (equation (6)) yields the intragenerational present value constraint for the maximization of intragenerational social welfare of generation $t$ in per capita terms:

$$\tau_{1,t} \bar{Z}_{1,t} + \tau_{2,t+1} \bar{Z}_{2,t+1} (1 + r)^{-1} = T_t + G_t.$$  \hspace{1cm} (17)

The constraint for intergenerational welfare maximization is \footnote{If the interest rate is not constant, this constraint becomes}

$$\sum_{t=0}^{\infty} (1 + r)^{-t} T_t N_t = D.$$  \hspace{1cm} (18)

Here $D$ is the explicit and implicit debt inherited from previous generations. Constraint (18) states that the present value of the taxes paid by all generations must be equal to $D$. This formulation rules out that the government runs a Ponzi game. An overview over the notation is given in Table 1 on page 20.

4 The optimal tax problem

The optimal tax problem is to find the optimal tax rates $\tau_{1,t}$, $\tau_{2,t+1}$, intragenerational transfers $G_t$ and intergenerational transfers $T_t$ for all generations $t$ which maximize intra- and intergenerational welfare. This problem can be solved in two steps:

1. **Maximization of intragenerational welfare subject to constraint (17)**

   For a given value of the total tax $T_t$ to be paid by generation $t$, the problem is

   $$\max_{\tau_{1,t}, \tau_{2,t+1}, G_t} W_t = N_t \int_{\bar{z}}^{z} W \left(V_t(\tau_{1,t}, \tau_{2,t+1}, G_t, w_t, w_{t+1}, z)\right) dF_t(z)$$  \hspace{1cm} (P1)

   s.t. $\tau_{1,t} \bar{Z}_{1,t} + \tau_{2,t+1} \bar{Z}_{2,t+1} (1 + r)^{-1} = T_t + G_t$.  \hspace{1cm} \footnote{If the interest rate is not constant, this constraint becomes}

   $$\sum_{t=0}^{\infty} \prod_{s=0}^{t} (1 + r_s)^{-1} T_s N_s = D.$$  \hspace{1cm} (P1)
This yields
\[ \hat{\tau}_{1,t}(T_t, w_t, w_{t+1}, F_t(z)), \hat{\tau}_{2,t+1}(T_t, w_t, w_{t+1}, F_t(z)), \hat{G}_t(T_t, w_t, w_{t+1}, F_t(z)) \]
and optimal intragenerational welfare \( \hat{W}_t(T_t, w_t, w_{t+1}, F_t(z)) \).

2. **Maximization of intergenerational welfare subject to constraint (P2)**

Here the problem corresponds to

\[
\max_{T_t} \Psi = \sum_{t=0}^{\infty} \left( \frac{1}{1+\rho} \right)^t \hat{W}_t(T_t, w_t, w_{t+1}, F_t(z)).
\]

s.t. \( \sum_{t=0}^{\infty} (1+r)^{-t} T_t N_t = D. \) (P2)

This yield the optimal intergenerational transfers \( T_t^* \) and therefore the optimal values

\[
\tau_{1,t}^*(w_t, w_{t+1}, F_t(z)), \tau_{2,t+1}^*(w_t, w_{t+1}, F_t(z)), G_t^*(w_t, w_{t+1}, F_t(z))
\]
for all periods \( t \).

From an incentive point of view, problem (P1) is the interesting economic problem. Since I assume that only labor income but neither \( z \) nor labor supply \( l_{a,t} \) are observable, it must be solved in a second-best analysis. It is similar to the optimum linear income tax problem (see Atkinson and Stiglitz (1980, p. 407)). However, the tax rate is allowed to be age-specific. It differs from the approach by Wrede (1999), Fenge et al. (2002) and the life-cycle taxation problem analyzed by Erosa and Gervais (2002) since these authors assume that all individuals of one generation are identical. In the following section 5 I solve problem (P1) and show that their results are modified when individuals are heterogenous.

The solution to problem (P2) depends decisively on the interest rate, the intergenerational discount rate and productivity changes over time. This can be seen by solving the Lagrangian

\[
L = \sum_{t=0}^{\infty} \left( \frac{1}{1+\rho} \right)^t \hat{W}_t(T_t, w_t, w_{t+1}, F_t(z)) + \mu \left( \sum_{t=0}^{\infty} (1+r)^{-t} T_t N_t - D \right).
\]
The first-order conditions are
\[
\frac{\partial L}{\partial T_t} = \left( \frac{1}{1+r} \right)^t \hat{W}_t + \mu (1+r)^{-t} N_t = 0 \quad \text{for all } t \tag{20a}
\]
\[
\frac{\partial L}{\partial \mu} = (1+r)^{-t} T_t N_t - D = 0 \tag{20b}
\]
Rewriting condition (20a) yields
\[
-\frac{\partial \hat{W}_t}{\partial T_t} N_t^{-1} = \mu \left( \frac{1+\rho}{1+r} \right)^t \quad \text{for all } t \tag{21}
\]
where \(-\partial \hat{W}_t^*/\partial T_t \times N_t^{-1}\) is marginal social welfare of intergenerational transfers per member of generation \(t\). With respect to the interest rate \(r\) and the intergenerational discount rate \(\rho\), a result similar to a standard life-cycle model for individual behavior follows: Marginal social welfare of intergenerational transfers per capita will fall over time if the interest rate rate exceeds the intergenerational discount rate and vice versa. If the intragenerational social welfare function \(\hat{W}_t(T_t, w_t, w_{t+1}, F_t(z))\) and wages were identical for all generations and the marginal social welfare of intergenerational transfers is decreasing in \(T_t\), \(\rho < r\) would therefore imply that later-born generations should bear a lower tax burden.

This result may not hold if the intragenerational social welfare function \(\hat{W}_t(T_t, w_t, w_{t+1}, F_t(z))\) is not identical for all generations. In particular, if productivity is rising over time, the marginal social welfare of intergenerational transfers will tend to fall as later-born generations become richer. In this case, it is possible that later-born generations should bear a higher tax burden even though \(\rho < r\).

5 The optimal structure of labor income taxation

In this section, I solve problem (P1) to find the optimal structure of labor income taxation. Since we look at one generation in isolation, the indices distinguishing generations are dropped. The Lagrangian is
\[
L = \int_{\tilde{z}} W (V(\tau_1, \tau_2, \tilde{G}; \tilde{z})) + \lambda (\tau_1 w_1 z_1 \tau_2 w_2 z_2 / (1+r) - T - \tilde{G}) \, dF(z) \tag{22}
\]
The first-order conditions are
\[
\frac{\partial L}{\partial \tau_1} = \int_z^\tau \left[ \frac{\partial W}{\partial V} \frac{\partial V}{\partial \tau_1} + \lambda \left( \tau_1 w_1 z \frac{\partial l_1}{\partial \tau_1} + \tau_2 w_2 z \frac{\partial l_2}{\partial \tau_1} \right) \right] dF(z) = 0 \tag{23a}
\]
\[
\frac{\partial L}{\partial \tau_2} = \int_z^\tau \left[ \frac{\partial W}{\partial V} \frac{\partial V}{\partial \tau_2} + \lambda \left( \tau_1 w_1 z \frac{\partial l_1}{\partial \tau_2} + \tau_2 w_2 z \frac{\partial l_2}{\partial \tau_2} \right) \right] dF(z) = 0 \tag{23b}
\]
\[
\frac{\partial L}{\partial G} = \int_z^\tau \left[ \frac{\partial W}{\partial V} \frac{\partial V}{\partial G} + \lambda \left( \tau_1 w_1 z \frac{\partial l_1}{\partial G} + \frac{\tau_2 w_2 z}{1 + r} \frac{\partial l_2}{\partial G} \right) - \frac{1}{1 + r} \left( \frac{\partial W}{\partial V} \frac{\partial V}{\partial \tau_1} \right) \right] dF(z) = 0 \tag{23c}
\]
\[
\frac{\partial L}{\partial \lambda} = \tau_1 w_1 z l_1 + \tau_2 w_2 z l_2 - T - G = 0. \tag{23d}
\]
I assume that the second-order condition is met. Applying Roy’s identity yields
\[
\frac{\partial V}{\partial \tau_1} = -\frac{\partial V}{\partial G} \frac{\partial l_1}{\partial \tau_1} = -\frac{\partial V}{\partial G} \frac{\partial z_1 l_1}{1 + r} \quad \text{and} \quad \frac{\partial V}{\partial \tau_2} = -\frac{\partial V}{\partial G} \frac{\partial z_2 l_2}{1 + r} \tag{24}
\]
where \( \partial V/\partial G \) is the private marginal utility of income. Furthermore, with the Slutsky relationship
\[
\frac{\partial l_k}{\partial \tau_t} = -w_t S_{kt} (1 + r)^{1-t} - w_t l_t (1 + r)^{1-t} \frac{\partial l_k}{\partial G}, \quad t, k = 1, 2, \tag{25}
\]
where \( S_{kt} \) is the derivative of the compensated demand curve, the first-order conditions (23a) and (23b) can be written as
\[
- \int_z^\tau \left[ \frac{\partial W}{\partial V} \frac{\partial V}{\partial G} \frac{\partial l_1}{\partial \tau_1} \right. \left. + \lambda \left( \tau_1 w_1 z \frac{S_{11}}{l_1} + \frac{\partial l_1}{\partial G} \right) + \tau_2 w_2 z \left( \frac{S_{21}}{l_1} + \frac{\partial l_2}{\partial G} \right) - 1 \right] w_1 z l_1 dF(z) = 0
\]
\[
- \int_z^\tau \left[ \frac{\partial W}{\partial V} \frac{\partial V}{\partial G} \frac{\partial l_2}{\partial \tau_2} \right. \left. + \lambda \left( \tau_1 w_1 z \frac{S_{12}}{l_2} + \frac{\partial l_1}{\partial G} \right) + \tau_2 w_2 z \left( \frac{S_{22}}{l_2} + \frac{\partial l_2}{\partial G} \right) - 1 \right] w_2 z l_2 dF(z) = 0.
\]
Defining the net social marginal valuation of income (compare Atkinson and Stiglitz (1980, p. 387)) by
\[
\nu(z) = \frac{\partial W}{\partial V} \frac{1}{k} + \tau_1 w_1 z \frac{\partial l_1}{\partial G} + \frac{\tau_2 w_2 z}{1 + r} \frac{\partial l_2}{\partial G}, \tag{26}
\]
the first-order conditions (23a) to (23c) can be transformed into

\[
\begin{align*}
- \int_{z}^{z} \left[ \nu(z) + \tau_1 w_1 z S_{11} \frac{l_1}{(1+r)l_1} + \tau_2 w_2 z S_{21} - 1 \right] w_1 z l_1 dF(z) &= 0 \quad (27a) \\
- \int_{z}^{z} \left[ \nu(z) + \tau_1 w_1 z S_{12} \frac{l_2}{(1+r)l_2} + \tau_2 w_2 z S_{22} - 1 \right] w_2 z l_2 dF(z) &= 0 \quad (27b) \\
\int_{z}^{z} [\nu(z) - 1] dF(z) &= 0 \quad (27c)
\end{align*}
\]

Condition (27c) implies \( \bar{\nu} = 1 \), i.e. the lump-sum element should be chosen such that the net social marginal valuation of income of the transfer on average equals its cost.

Now defining the compensated elasticities

\[
\varepsilon_{kt} \equiv \frac{(1 - \tau_t)(1 + r)^{1-t} w_t S_{kt}}{l_k} \Rightarrow \frac{\tau_t}{1 - \tau_t} \varepsilon_{kt} = \frac{\tau_t (1 + r)^{1-t} w_t S_{kt}}{l_k} \quad (28)
\]

and using the symmetry relationship \( S_{12} = S_{21} \) and \( \bar{\nu} = 1 \) allows to transform conditions (27a) and (27b) into

\[
\begin{align*}
\int_{z}^{z} \left[ \nu(z) - \bar{\nu} + \frac{\tau_1}{1 - \tau_1} \varepsilon_{11} + \frac{\tau_2}{1 - \tau_2} \varepsilon_{12} \right] w_1 z l_1 dF(z) &= 0 \quad (29a) \\
\int_{z}^{z} \left[ \nu(z) - \bar{\nu} + \frac{\tau_1}{1 - \tau_1} \varepsilon_{21} + \frac{\tau_2}{1 - \tau_2} \varepsilon_{22} \right] w_2 z l_2 \frac{1 + r}{1 + r} dF(z) &= 0 \quad (29b)
\end{align*}
\]

Rearranging and defining labor income by \( Z_t \equiv w_t z l_t \) yields\(^5\)

\[
\begin{align*}
\frac{\tau_1}{1 - \tau_1} \int_{z}^{z} \varepsilon_{11} Z_1 dF(z) + \frac{\tau_2}{1 - \tau_2} \int_{z}^{z} \varepsilon_{12} Z_1 dF(z) &= -\text{Cov}(\nu, Z_1) \\
\frac{\tau_1}{1 - \tau_1} \int_{z}^{z} \varepsilon_{21} Z_2 dF(z) + \frac{\tau_2}{1 - \tau_2} \int_{z}^{z} \varepsilon_{22} Z_2 dF(z) &= -\text{Cov}(\nu, Z_2)
\end{align*}
\]

since \( \text{Cov}(\nu, Z_t) = \int_{z}^{z} [\nu(z) - \bar{\nu}] Z_t dF(z) \). These covariances are marginal measures of inequality (Atkinson and Stiglitz (1980, p. 408)). Solving for \( \tau_t/(1 - \tau_t) \)

---

\(^5\)Note that the compensated elasticities will usually depend on productivity \( z \) and can therefore not be transferred to the other side of the integral sign.
results in

\[
\frac{\tau_1}{1 - \tau_1} = \frac{\text{Cov}(\nu, Z_2) \int_{z}^{\bar{z}} \varepsilon_{12} Z_1 \, dF(z) - \text{Cov}(\nu, Z_1) \int_{z}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z)}{\int_{z}^{\bar{z}} \varepsilon_{11} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z) - \int_{z}^{\bar{z}} \varepsilon_{12} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{21} Z_2 \, dF(z)}
\]

\[
\frac{\tau_2}{1 - \tau_2} = \frac{\text{Cov}(\nu, Z_1) \int_{z}^{\bar{z}} \varepsilon_{21} Z_1 \, dF(z) - \text{Cov}(\nu, Z_2) \int_{z}^{\bar{z}} \varepsilon_{11} Z_1 \, dF(z)}{\int_{z}^{\bar{z}} \varepsilon_{11} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z) - \int_{z}^{\bar{z}} \varepsilon_{12} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{21} Z_2 \, dF(z)}
\]

Combining (31a) and (31b) leads us to the following result for the optimal structure of age-dependent taxation:

**Proposition 1:** The optimal tax rates fulfil the generalized Ramsey rule

\[
\frac{\tau_1}{1 - \tau_1} = \frac{\text{Cov}(\nu, Z_2) \int_{z}^{\bar{z}} \varepsilon_{12} Z_1 \, dF(z) - \text{Cov}(\nu, Z_1) \int_{z}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z)}{\int_{z}^{\bar{z}} \varepsilon_{11} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z) - \int_{z}^{\bar{z}} \varepsilon_{12} Z_1 \, dF(z) \int_{z}^{\bar{z}} \varepsilon_{21} Z_2 \, dF(z)}
\]

(32)

In general, taxation will be age-specific as in Wrede (1999), Fenge et al. (2002) and Erosa and Gervais (2002). However, the analysis yields a different result because individuals are allowed to be heterogeneous. This can best be seen when condition (32) is compared to the standard rule (see Atkinson and Stiglitz (1980, p. 375) and Fenge et al. (2002, p. 17))

\[
\frac{\tau_1}{1 - \tau_1} = \frac{\varepsilon_{12} - \varepsilon_{22}}{\varepsilon_{21} - \varepsilon_{11}}
\]

(33)

which abstracts from income inequality and the possibility of lump-sum taxation.

Comparing conditions (32) and (33) shows two important differences:

6Erosa and Gervais (2002, p. 360) derive an alternative formulation of the Ramsey rule (33) for directly additive utility function (see Atkinson and Stiglitz (1980, p. 379)). They show that the relative age-dependent tax rates on labor income are inversely related to the relative income elasticities of age-specific labor supply.
1. In the general solution (32), elasticities are multiplied by income. This takes into account that individuals with high income pay higher taxes and must therefore be given more weight in the optimal tax rule.

2. In the optimal solution the covariance of net social marginal valuation of income and labor income is crucial. Assuming that labor income $Z_t$ is a non-decreasing function of productivity $z$, this covariance is non-positive if the net social marginal valuation of income is non-increasing with $z$. For instance, assume $\varepsilon_{12} = \varepsilon_{21} = 0$. Then

$$\frac{\tau_1}{1 - \tau_1} = \frac{\text{Cov}(\nu, Z_1) \int_{\underline{z}}^{\bar{z}} \varepsilon_{22} Z_2 \, dF(z)}{\text{Cov}(\nu, Z_2) \int_{\underline{z}}^{\bar{z}} \varepsilon_{11} Z_1 \, dF(z)},$$

i.e. the tax rate tends to be higher in the period in which the covariance of labor income and the net social marginal valuation of income is higher in absolute terms. Intuitively, income should be taxed more in the period with higher inequality. In particular, the wage rates influence $\text{Cov}(\nu, Z_t)$. Using $Z_t = w_t z l_t$ yields

$$\text{Cov}(\nu, Z_t) = w_t \int_{\underline{z}}^{\bar{z}} [\nu(z) - \bar{\nu}] z l_t \, dF(z).$$

Thus, if the wage rate rises over time, e.g. by more input of physical capital, the covariance will also tend to increase. This would be an argument for taxing individuals higher when old.

Note that if individuals are homogeneous, then $\text{Cov}(\nu, Z_1) = \text{Cov}(\nu, Z_2) = 0$ which implies $\tau_1 = \tau_2 = 0$ by (31a) and (31b). In this case, the budget constraint (17) implies

$$-G_t = T_t.$$  

Therefore, transfers to other generations and other public expenditure are financed by negative intragenerational transfers $-G_t$, i.e. by lump-sum taxation. In the models of Wrede (1999), Fenge et al. (2002) and Erosa and Gervais (2002) this possibility was excluded because they implicitly impose the restriction $G_t = 0$ which leads to the standard Ramsey rule (33).
Finally, it would be interesting to see whether the empirical results of Fenge et al. (2002) still hold if formula (32) is used instead of (33). They assume \( \varepsilon_{12} = \varepsilon_{21} = 0 \). In this case, equation (34) describes the optimal structure of taxation. A plausible hypothesis is that inequality increases with age, i.e. \( \text{Cov}(w_t, Z_2) > \text{Cov}(w_t, Z_1) \).

However, it may already be instructive to assume \( \text{Cov}(w_t, Z_1) = \text{Cov}(w_t, Z_2) \) and to see whether weighting elasticities by income instead of using average elasticities leads to different results.

## 6 The optimal retirement benefit formula

In this section, I take an optimal solution \( \mathcal{T}_t^*, \tau_{1,t}^*, \tau_{2,t+1}^*, G_t^* \) as given and examine the implications for the design of tax and pension systems for a particular generation \( t \), i.e. the parameters \( \tau_{1,t}^*, \tau_{2,t+1}^*, \alpha_{t+2}, \beta_{2,t+1}^* \). The following equations summarize the requirements for the parameters \( \tau_{1,t}^*, \tau_{2,t+1}^*, \alpha_{t+2}, \beta_{2,t+1}^* \) which can be chosen for each generation \( t \) (see equations (4) to (6) and (14) to (16)):

\[
\begin{align*}
\tau_{1,t}^* + \tau_{1,t}^* &= \tau_{1,t}^* + \tau_{1,t}^* \\
\tau_{2,t+1}^* + \tau_{2,t+1}^* &= \tau_{2,t+1}^* \\
\tau_{2,t+1}^* &= b_t \left(1 - \beta_{2,t+1}^*(1+r)^{-2}\right) \\
\tau_{2,t+1}^* &= b_t \left(1 - \beta_{2,t+1}^*(1+r)^{-1}\right) \\
G_t^* &= B_t^* + \frac{\alpha_{t+2}^*}{(1+r)^2} \\
\tau_{1,t}^*Z_{1,t}^* + \tau_{2,t+1}^*Z_{2,t+1}^* &= B_t^* + \mathcal{T}_t^d \\
\tau_{1,t}^*Z_{1,t}^* + \tau_{2,t+1}^*Z_{2,t+1}^* &= \mathcal{T}_t^b + \frac{\alpha_{t+2}^*}{(1+r)^2} \\
\mathcal{T}_t^b + \mathcal{T}_t^d &= \mathcal{T}_t^*.
\end{align*}
\]

Note that the contribution rates \( b_t \) and \( b_{t+1} \) (see equation (13))

\[
\begin{align*}
b_t &= \frac{\alpha_t + \beta_{1,t-2}b_{t-2}Z_{1,t-2} + \beta_{2,t-1}b_{t-1}Z_{2,t-1}}{Z_{1,t}(1+m_{t-1})(1+m_{t-2}) + Z_{2,t}(1+m_{t-1})} \\
b_{t+1} &= \frac{\alpha_{t+1} + \beta_{1,t-1}b_{t-1}Z_{1,t-1} + \beta_{2,t}b_{t}Z_{2,t}}{Z_{1,t+1}(1+m_{t})(1+m_{t-1}) + Z_{2,t+1}(1+m_{t})}
\end{align*}
\]

14
are determined by the retirement benefit formulas of past generations and independent of \( \alpha^*_t \), \( \beta^*_1, \beta^*_2 \).

Since the budget constraint
\[
\tau^*_t Z^*_{1,t} + \tau^*_t Z^*_{2,t+1} R^{-1} = G^*_t + \tau^*_t
\]
is met in the optimum, one of the equations (40) to (42) is redundant. Dropping equation (41) yields seven equations for ten unknowns. Three parameters can therefore be chosen with the restrictions

- either \( \alpha^*_t \) or \( B^*_t \) (condition (40)),
- either \( \tau^*_a, \tau^*_b \) or \( \beta^*_a, \beta^*_b \) (conditions (36) and (38))
- either \( \tau^*_a, \tau^*_b \) or \( \beta^*_a, \beta^*_b \) (conditions (37) and (39))

In particular, \( \alpha^*_t, \beta^*_1, \beta^*_2 \) can be chosen freely for each generation. This implies

**Proposition 2:** Without further restrictions, any retirement benefit formula can be optimal.

This result is in contrast to Wrede (1999) and Fenge et al. (2002) who did not consider the existence of a general tax system. It is explained by the fact that the public pension system and the general tax system are prefect substitutes with respect to labor income taxation. Note, however, that once a retirement benefit formula is chosen, all the remaining parameters \( \tau^*_a, \tau^*_b, \beta^*_a, \beta^*_b, B^*_t, T^*_d \) are determined, i.e. tax and debt policy must follow the design of the pension system.

Imposing restrictions on the general tax system, however, may yield an “optimal retirement benefit formula”. In particular, consider the case that direct taxes on labor income cannot be conditioned on age, i.e. \( \tau^*_a = \tau^*_b = \bar{\tau}^* \). Then the retirement benefit formula is essential in differentiating tax rates by age. Using (36) to (39) yields

\[
\beta^*_1 = \frac{b_t + \bar{\tau}^*_a}{b_t} (1 + r)^2 = \frac{b_t + \bar{\tau}^*_a - \tau^*_1}{b_t} (1 + r)^2
\]
\[
\beta^*_2 = \frac{b_t + \bar{\tau}^*_b}{b_t} (1 + r) = \frac{b_t + \bar{\tau}^*_b - \tau^*_2}{b_t} (1 + r).
\]
Proposition 3: If direct taxes on labor income cannot be conditioned on age, i.e. 
\[ \tau_{1,t}^{l*} = \tau_{2,t+1}^{l*} = \bar{\tau}^{l*}, \]
then the retirement benefit formula is essential in implementing age-dependent taxation.

This result is important in the context of the analysis by Erosa and Gervais (2002) who also consider the case that age-dependent tax rates are not possible. However, they only take into account explicit taxes. Through a pension system, it is easily possible to implement age-dependent tax rates. In fact, current retirement pension formulas usually imply age-dependent tax rates which fall through the contributor’s working life (Bütler (2002)). For example, this is the case in the German pension formula which may even lead to a negative implicit tax rate (Beckmann (2000)). Notional Defined Contributions Systems, in which the rate of return equals the growth rate of the aggregate wage sum, also imply an implicit tax schedule which is decreasing in age (See Lindbeck and Persson (2003, p. 85)). However, whereas these formulas rely on actuarial concepts, the analysis in section 5 has argued that age-dependent taxation should follow the generalized Ramsey rule (32). Substituting equations (36) and (37) into (32) and using 
\[ \tau_{1,t}^{l*} = \tau_{2,t+1}^{l*} = \bar{\tau}^{l*} \]
yields

\[
\frac{\tau_{1,t}^{b*} + \tau_{t+1}^{b*}}{1 - \tau_{1,t}^{b*} - \tau_{t+1}^{b*}} = \frac{\text{Cov}(\nu_t, Z_{2,t+1}) \int \varepsilon_1 Z_1 \, dF_t(z) - \text{Cov}(\nu_t, Z_{1,t}) \int \varepsilon_2 Z_{2,t+1} \, dF_t(z)}{\text{Cov}(\nu_t, Z_{1,t}) \int \varepsilon_2 Z_{2,t+1} \, dF_t(z) - \text{Cov}(\nu_t, Z_{1,t+1}) \int \varepsilon_1 Z_{1,t+1} \, dF_t(z)}
\]

and the corresponding values \( \beta_{1,t}^{*} \) and \( \beta_{2,t+1}^{*} \) from equations (46) and (47) as a rule for the implicit taxation through the retirement benefit formula. Only by chance will this rule correspond to existing retirement benefit formulas.

Note that the retirement benefit formula generally needs to be different for each generation if population growth or average income is not constant over time. This can be seen from equations (46) and (47) which show that the optimal retirement benefit formula is also determined by the contribution rates in periods \( t \) and \( t + 1 \). Equations (44) and (45) show that these depend on past population growth rates and average incomes of generations \( t - 1, t \) and \( t + 1 \) besides the retirement benefit.
formulas for generations $t - 2$ and $t - 1$. Thus, even if $\tau_{1,t}^{b_1}$ and $\tau_{2,t+1}^{b_1}$ are constant over time, $\beta_{1,t}^{*}$ and $\beta_{2,t+1}^{*}$ may vary from generation to generation.

Finally, I comment on the possibility of a constant contribution rate, a prominent objective in the political debate. Imposing the restriction that the contribution rate must be the same in period $t + 2$ as in period $t + 1$, the following condition must hold

$$b_{t+2} = \frac{\alpha_{t+2} + \beta_{1,t}^{*}b_t\tilde{Z}_{1,t+1} + \beta_{2,t+1}^{*}b_{t+1}\tilde{Z}_{2,t}}{\tilde{Z}_{1,t+1}(1 + m_{t+1})(1 + m_t) + \tilde{Z}_{2,t+1}(1 + m_{t+1})} = b_{t+1}. \quad (48)$$

If $\beta_{1,t}$ and $\beta_{2,t}$ are used to implement age-dependent taxation, then the minimum pension $\alpha_{t+2}$ can guarantee a constant contribution rate. However, this can imply a negative minimum pension which may not be enforceable if individuals with low productivity do not save for old age. In this case, there must either be an offsetting tax-financed transfer in old age to pay for the negative minimum pension. Alternatively, the pension system can be subsidized to guarantee $\alpha_{t+2} \geq 0$ and a constant contribution rate.

## 7 Conclusions

This paper has taken a comprehensive view of life-time taxation including both explicit taxation through the general tax system and implicit taxation via the retirement benefit formula. First, the analysis of Wrede (1999), Fenge et al. (2002) and Erosa and Gervais (2002) was extended by allowing individuals to differ in productivity. This led to a generalized Ramsey rule which requires elasticities to be weighted by income. Furthermore, income should be taxed more in periods with higher inequality. Second, the paper characterized the optimal retirement benefit formula in the presence of a general tax system. It was shown that the optimal retirement benefit formula cannot be defined independently of the general tax rates. However, a particular advantage of implicit taxation through the pension system is that age-dependent taxation can easily be implemented.

\[\beta_{1,t}^{*} = \frac{b_t + \tilde{q}_t^{*} - \tau_{1,t}^{*}}{b_t} (1 + r_t)(1 + r_{t+1}) \quad \text{and} \quad \beta_{2,t+1}^{*} = \frac{b_{t+1} + \tilde{q}_{t+1}^{*} - \tau_{2,t+1}^{*}}{b_{t+1}} (1 + r_{t+1}),\]

where $r_t$ is the interest rate in period $t$.\[\]
The analysis was carried out within an overlapping generation model in which individuals work for two periods and are retired in the last period. The results can be extended to an economy with more working and retirement periods. A limitation is that taxing capital income was not considered. As Erosa and Gervais (2002, p. 353) show, an optimal solution may include a tax rate on capital income different from zero.

The results of this paper may be useful for future empirical research. In particular, it would be interesting to examine the implications of the generalized Ramsey rule for optimal age-dependent taxation. Also the combined effect of income taxation and implicit taxation through the retirement benefit formula could be examined in an empirical study. Finally, it may be instructive to study how people perceive implicit taxation through the retirement benefit formula. In this paper, it was assumed that individuals can correctly infer the implicit age-dependent tax rates. However, the actuarial concepts which are necessary to calculate these tax rates may not be familiar to all citizens.
References


Table 1: Overview over the notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>interest rate</td>
</tr>
<tr>
<td>m_t</td>
<td>population growth rate in period t</td>
</tr>
<tr>
<td>t</td>
<td>time index and index for individuals born in period t (generation t)</td>
</tr>
<tr>
<td>a, t</td>
<td>subscript: in period t the individual has age a, i.e. is in period a of her life</td>
</tr>
<tr>
<td>c_{a, t}</td>
<td>consumption of an individual with age a in period t</td>
</tr>
<tr>
<td>l_{a, t}</td>
<td>labor supply of an individual with age a in period t</td>
</tr>
<tr>
<td>w_t</td>
<td>wage rate per efficiency unit of labor in period t</td>
</tr>
<tr>
<td>z</td>
<td>productivity of an individual, distributed in $z \in {z, \bar{z}}$ according to the distribution function $F_t(z)$ for generation t</td>
</tr>
<tr>
<td>Z_{a, t}</td>
<td>labor income $w_t z l_{a, t}$</td>
</tr>
<tr>
<td>\bar{Z}_{a, t}</td>
<td>average labor income of generation t</td>
</tr>
<tr>
<td>\tau_{a, t}</td>
<td>total tax rate on labor income for an individual with age a in period t</td>
</tr>
<tr>
<td>\tau_{a, t}^d</td>
<td>direct tax rate on labor income</td>
</tr>
<tr>
<td>\tau_{a, t}^b</td>
<td>indirect tax rate on labor income through the pension system</td>
</tr>
<tr>
<td>G_t</td>
<td>present value of total intragenerational transfers</td>
</tr>
<tr>
<td>B_{a, t}</td>
<td>tax-financed basic income for an individual with age a in period t</td>
</tr>
<tr>
<td>B_t</td>
<td>present value of tax-financed basic income</td>
</tr>
<tr>
<td>b_t</td>
<td>contribution rate to the public pension system in period t</td>
</tr>
<tr>
<td>p_t</td>
<td>pension benefit in period t</td>
</tr>
<tr>
<td>\alpha_t</td>
<td>minimum pension in period t</td>
</tr>
<tr>
<td>\beta_{a, t}</td>
<td>weight on contributions in period t if the individual has age a</td>
</tr>
<tr>
<td>T^{d}_t</td>
<td>transfers to other generations in terms of a change in public debt</td>
</tr>
<tr>
<td>T^{b}_t</td>
<td>implicit tax of the public pension system</td>
</tr>
<tr>
<td>T_t</td>
<td>total tax per capita to be paid by generation t</td>
</tr>
<tr>
<td>D</td>
<td>debt inherited from previous generations</td>
</tr>
<tr>
<td>V_t</td>
<td>indirect utility of a member of generation t</td>
</tr>
<tr>
<td>W_t</td>
<td>intragenerational welfare within generation t</td>
</tr>
<tr>
<td>\Psi</td>
<td>intergenerational welfare</td>
</tr>
<tr>
<td>\rho</td>
<td>intergenerational discount rate</td>
</tr>
</tbody>
</table>