Optimal capital income taxation and redistribution

Ulrike Ludden*
University of Mannheim

January 2000

Abstract

This paper studies the effects of agent heterogeneity on optimal capital income tax rates. In a two period model with arbitrarily many heterogeneous agents, we explicitly derive the welfare effects of taxation depending on the distribution of the agents' characteristics. In particular, we show that the sign of the optimal capital income tax rate depends not on the extent of inequality in goods endowments and productivities each by itself, but on a measure of inequality in their joint distribution.

Keywords: Optimal taxation; Capital income; Heterogeneous agents; Redistribution.

JEL classification: H21, H24

1 Introduction

The study of optimal tax systems in a dynamic framework has mainly focused on efficiency aspects.[1] In the present paper we choose a different approach, focusing on the impacts of agent heterogeneity on optimal tax rates, where taxes are collected for redistributional purposes. Under the assumption that the government maximizes a social welfare function, we ask which forms of taxation are optimal for different sources of inequality? How do correlations between labor income and wealth affect optimal tax rates?

*Graduiertenkolleg "Allokation auf Finanz- und Gütermärkten," 68131 Mannheim, Germany. Phone: -49-621-181-3425. Email: *uludden@econ.uni-mannheim.de*. I am indebted to Martin Hellwig and Marco Bassetto for many helpful discussions. Further thanks go to Larry Jones and Michele Tertilt for their comments on earlier drafts. Part of this work was done during my visit at Northwestern University, whose hospitality is gratefully acknowledged. All remaining errors are my own.

[1]See, for example, Chamley (1986), Chari, Christiano, and Kehoe (1994), Milesi-Ferretti and Roubini (1995), and Jones, Manuelli, and Rossi (1997).

To answer these questions we develop a two period model, in which households make labor-leisure choices and decide how much to consume and how much to save. The government uses linear tax rates on labor and capital income to maximize a social welfare function. Households are heterogeneous with respect to their endowments and abilities.

Remaining tractable analytically, the present model allows us to analyze the effects of different sources of heterogeneity among households on optimal tax rates. In particular, we show that the optimal capital income tax rate generally is non-zero. It depends crucially on the joint distribution of initial wealth and productivities and on the flexibility of labor income taxes over time.

How do our main results relate to earlier findings on optimal capital income taxation? Infinite horizon representative agent models in the tradition of Chamley (1986) show that it is not optimal to finance an exogenous stream of government expenditures through capital income taxes in the long run. However, there are a number of initial periods (the number depending on a possible upper bound on tax rates) in which the optimal capital income tax is strictly positive, declining to zero afterwards. The main effect of initially high levels of capital income taxes is to extract the endowments from the consumers. The government builds a large surplus in the initial periods from which it finances part of its expenditures thereafter.

A limitation of this approach is its reliance on representative agents with an infinite horizon. If there is only one agent, intragenerational distribution is not an issue. Furthermore, maximization in infinite horizon models with linear discounting often implies high tax rates in earlier periods combined with lower or zero taxes in later periods. If the infinitely living agent is interpreted as a succession of generations, this implies a high burden on earlier generations while later generations benefit. Since the different period's utilities are summed up over time and only the sum is maximized, intergenerational distribution is irrelevant for the optimal policy. One further limitation of most infinite horizon models is the assumption that the government is allowed to build up a substantial surplus in the early periods, which is often limited only by the assumption that taxes ought to be no higher than 100%. Tax rates of this magnitude might be hard to implement.[2]

Judd (1985) considers an infinite horizon model with two types of agents. In his most general setting, the agents differ with respect to their initial endowments and utility functions. Agents are assumed to derive utility from consumption and leisure. The government has a fixed stream of expenditures over time and raises revenues through capital and labor income taxation, while it redistributes through a non-negative lump-sum transfer which may be different for both types

[2] The influence of budget constraints on optimal capital income tax rates is shown by Stockman (1999), who finds that optimal capital income taxes are highly volatile and generally have a non-zero mean.

of agents. Production is weakly separable between capital and labor. Within this framework, Judd shows that if there exists a steady state then in this steady state it is not optimal to tax capital income.

Chari and Kehoe (1999) build a similar model, the main difference being the absence of lump-sum transfers. Thus, redistribution is a side-effect of revenue raising. They confirm Judd's results and additionally show that the assumption of a weakly separable production function is necessary for the optimality of zero capital income taxes.

While these models show that the optimal tax rate in the steady state is zero, they say nothing about optimal rates off the steady state. Our model, by contrast, considers a finite number of periods and shows that zero taxes on capital income are not optimal if, for example, goods endowments are heterogeneous.

While the models depicted above assume perfectly competitive markets, a number of studies analyzes optimal capital income taxes if there are market imperfections. With few exceptions they find that the optimal capital income tax rate is different from zero. Judd (1997), for example, shows that monopolistic competition among firms can lead to the optimality of a negative tax on capital income. Aiyagari (1995) and Chamley (1998), on the other hand, find that incomplete credit markets can lead to the optimality of a positive capital income tax rate. The present paper refrains from the analysis of market imperfections and shows that even if markets are complete it can be optimal to impose a strictly positive tax on capital income.

The study of optimal capital income taxation is closely related to earlier work on uniform commodity taxation, as for example Atkinson and Stiglitz (1976). Their static analysis can be reinterpreted in terms of a dynamic model where different commodities represent consumption at different points in time. In a setting with heterogeneous agents (differing with respect to their productivity), Atkinson and Stiglitz show that it is not optimal to distort relative prices of consumption goods if a sufficiently flexible income tax scheme is available. In a dynamic interpretation, their result implies that the optimal capital income tax rates are zero. However, Atkinson and Stiglitz consider only one factor of production, while there should be multiple factors of production in a dynamic setting—labor in different periods constitutes different input factors in the intertemporal production function. Thus, taxes on interest earnings not only determine relative prices but also relative wages. Furthermore, their agents have no endowments of consumption goods and identical preferences.

Finally, our model is related to the literature on optimal linear income taxation. Given a one-dimensional heterogeneity in the agents' productivities, Sheshinski (1972) shows in a static setting that the optimal marginal income tax rate is strictly positive and less than 100%. The model presented in this paper considers heterogeneity in two dimensions (in productivities and goods endowments) and shows under which conditions the one-dimensional result holds.

Summarizing, only a few studies have considered distributional aspects in their analysis of optimal taxation in a dynamic setting. At most, they consider two different types of agents and their results are almost exclusively based on steady state analysis. In this paper we allow for an arbitrary number of heterogeneous agents and analyze the influence of different sources of heterogeneity on optimal tax rates in a finite horizon model. Furthermore, we show how the introduction of heterogeneous endowments, non-separable production functions, or non-homothetic preferences changes the optimal tax system away from zero capital income taxation.

The rest of the paper is organized as follows. In chapter 2, we develop a two period model emphasizing the interaction between the households' heterogeneity and optimal taxation. Section 3 derives the welfare effects of capital and labor income taxes depending on the sources of heterogeneity among households and on their joint distribution. The robustness of our results is examined in chapter 5, where we also compare our model with the existing literature. Chapter 6 concludes.

2 The model

Consider an economy existing for two periods, with N households which are heterogeneous with respect to their labor productivities n^j and their endowments e^j, where $j = 1,..,N$. In each period, there is one consumption good (c_1^j, c_2^j) and one type of labor (l_1^j, l_2^j). Second period's utility is discounted with the factor ρ.

All households have identical preferences and live two periods. Their utility function is log-linear and identical in every period and across households. That is, household j's utility in period 1 can be expressed as

$$u_1^j = a \ln(1 - l_1^j) + (1 - a) \ln(c_1^j) \quad , \tag{1}$$

where l_1^j refers to household j's labor supply in the first period, c_1^j is its consumption in the first period, and $a \in (0, 1)$ determines the relative importance of leisure vs. consumption. Each household is endowed with one unit of time in every period. Total utility of household j then is

$$\begin{aligned} U^j &= u_1^j + \rho u_2^j \quad , \\ &= a \ln(1 - l_1^j) + (1 - a) \ln c_1^j + \rho \left[a \ln(1 - l_2^j) + (1 - a) \ln c_2^j \right] \quad . \end{aligned} \tag{2}$$

In the first period, each household has to decide how much to work and how much to consume. It can save an amount k^j and will earn interest r on its savings in the second period. There is no depreciation of capital. The wage of household j is given by its productivity n^j, its labor income is $l_t^j n^j$ in each period $t = 1, 2$. Production is linear with first period's output given by $\sum_j n^j l_1^j$ and second period's by $\sum_j rk^j + n^j l_2^j$.

The government wants to maximize a social welfare function of the form $\sum_j \omega^j U^j$ through linear taxation of labor and capital income, where ω^j is the weight assigned to household j. Since the individual marginal utilities are decreasing in consumption and leisure, redistribution from wealthier households to poorer households increases welfare—unless the government favors richer households, implying that the welfare weights are positively correlated with individual utility. We assume that the government cannot observe endowments and productivities directly, it distinguishes the agents only by their incomes.Capital income is taxed with $(1-\beta_r)$, and labor income with $(1-\beta_w)$.[3]

The households get lump-sum transfers α, which are identical in both periods and across households and may be either positive or negative. We assume that the government has a commitment technology. That is, once the households have made their labor/leisure decisions, the government cannot change the tax rates anymore.[4] The households' budget constraints are:

$$\begin{aligned} &\text{Period 1} \qquad k^j + c_1^j &&= \alpha + \beta_w l_1^j n^j + e^j && (3)\\ &\text{Period 2} \qquad c_2^j &&= \alpha + \beta_w l_2^j n^j + k^j(1+\beta_r r) && (4) \end{aligned}$$

The government earns (pays) the same interest rate as the households on any budget surplus (deficit) in the first period.[5] Letting B denote its budget, the government's budget constraint is

$$B = \sum_{j=1}^{N} \left\{ \left[(1-\beta_w) l_1^j n^j - \alpha \right] (1+r) + (1-\beta_w) l_2^j n^j + (1-\beta_r) r k^j - \alpha \right\} \geq 0 \quad . \tag{5}$$

Given the tax rates and transfers, household j maximizes its utility subject to (3) and (4). The optimal values for its choice variables are

$$\hat{k}^j(\alpha, \beta, n^j, e^j) = \frac{1}{\rho+1} \left[\rho \cdot (\alpha + e^j + \beta_w n^j) - \frac{\alpha + \beta_w n^j}{r\beta_r + 1} \right] \tag{6}$$

$$\hat{l}_1^j(\alpha, \beta, n^j, e^j) = \max \left\{ 0, 1 - \frac{a}{\rho+1} \frac{inc^j}{\beta_w n^j} \right\} \tag{7}$$

$$\hat{l}_2^j(\alpha, \beta, n^j, e^j) = \max \left\{ 0, 1 - \rho(r\beta_r + 1) \cdot \frac{a}{\rho+1} \frac{inc^j}{\beta_w n^j} \right\} \tag{8}$$

[3]If the planner knows the amount of interest earnings and wage income of an individual, it would be reasonable to believe that he could infer the size of the initial endowment. We rule this out by assumption since heterogeneous endowments are meant to represent generic differences between individuals rather than purely monetary ones.

[4]This assumption is crucial for most work on optimal taxation. For an analysis of optimal taxation without commitment see Fischer (1980) or Klein and Rios-Rull (1999).

[5]For ease of exposition, we assume that the government has no expenses besides redistribution. A fixed revenue requirement would not change the results since we allow lump-sum taxes.

$$\hat{c}_1^j(\alpha,\beta,n^j,e^j) = \frac{1-a}{\rho+1}\cdot inc^j \tag{9}$$

$$\hat{c}_2^j(\alpha,\beta,n^j,e^j) = \rho(r\beta_r+1)\cdot\frac{1-a}{\rho+1}\cdot inc^j \quad , \tag{10}$$

where $\beta = (\beta_r, \beta_w)$, and $inc^j = \alpha + e^j + \beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r}$. The term inc^j denotes the potential income of household j, which is its income if the labor supply would be one in both periods. We assume that there are no negative endowments.

The households' reactions to changes in parameter values are in the expected directions. Higher endowments lead to higher savings and to lower labor supply. Higher wage income ($n^j \uparrow, \beta_w \uparrow$) leads to higher labor supply and higher consumption while the effects on savings depend on the magnitude of $\rho(r\beta_r+1)$. The term $\rho(r\beta_r+1)$ determines the ratio of second period's leisure and consumption to first period's leisure and consumption. As long as $\rho(r\beta_r+1) > 1$, the households' discount rate ($\frac{1}{\rho}-1$) is smaller than the net interest rate $\beta_r r$. Thus, households shift more utility to the second period by working less and consuming more. If their wage income increases in both periods while first period's goods endowment remains the same, they have to increase savings to maintain the same relation between first and second period's consumption and leisure (the ratios are constant because of homothetic preferences). Thus, savings increase in wage income if $\rho(r\beta_r+1) > 1$, and decrease otherwise. A higher discount factor ($\rho \uparrow$) leads to higher savings, to higher labor supply in the first period and to a lower one in the second period. Consumption changes the opposite way.

Taking the households' choices (6) to (10) as given, the maximized utility of household j depends only on α, β, n^j and e^j and can be written as $V(\alpha,\beta,n^j,e^j)$. The indirect utility function $V(\cdot)$ is increasing in the household's productivity n^j and in its endowment e^j, since higher values of these variables lead to higher consumption and lower labor supply in both periods.

3 Optimal taxation

This chapter derives properties of the optimal linear tax schedule with a special emphasis on capital income taxation. Given the households' choices, the planner determines α, β_r, and β_w to maximize welfare. Letting W denote the corresponding Lagrangian, the planner's maximization problem can be written as

$$\max_{\alpha,\beta_r,\beta_w,\lambda} W = \sum_{j=1}^{N} \omega^j V(\alpha,\beta,n^j,e^j,\rho) + \lambda B \quad , \tag{11}$$

where λ is the Lagrange multiplier for the planner's budget constraint. Since we want to show that a tax rate of zero is generally not optimal, we focus on the analysis of the planner's first order conditions, evaluated at a tax rate of zero. That is, we calculate the marginal welfare effect of introducing labor or capital income taxes.

Let us first consider labor income taxes. The analysis of the necessary conditions for the maximization in (11) leads to the following proposition:

PROPOSITION 1 *Given the utility and welfare functions specified above, and assuming that the weights ω^j are uncorrelated with productivities and endowments, the labor income tax $1-\beta_w$ is related to welfare as follows:*

i. It is never optimal to tax away all labor income, i.e. $\beta_w^* > 0$.

ii. If productivities are heterogeneous while endowments are not, labor income taxes increase welfare. The increase in welfare rises in the productivities' heterogeneity as measured by $\sum_j \frac{n^j - \bar{n}}{inc^j}$.[6]

iii. If endowments are heterogeneous while productivities are not, labor income taxes decrease welfare. The decrease in welfare rises in the endowments' heterogeneity as measured by $\sum_j \frac{e^j - \bar{e}}{inc^j}$.

Proof. If the labor income tax rate is 100% (*i.e.* $\beta_w = 0$) nobody works and the planner does not collect any revenue from labor taxation. Thus, a labor tax rate of 100% is never optimal, see also Sheshinski (1972).

For parts (*ii*) and (*iii*) consider the marginal welfare effect of labor income taxes. If there is no labor taxation ($\beta_w = 1$), the marginal effect is:[7]

$$\left.\frac{\partial W}{\partial \beta_w}\right|_{\beta_w=1} = \sum_{j=1}^{N} \omega^j (\rho+1) \left\{ \frac{(1-a)(r\beta_r+2)\left[n^j - \bar{n}\right] - a(r\beta_r+1)\left[e^j - \bar{e}\right]}{inc^j} \right\} , \qquad (12)$$

where $inc^j = \alpha + e^j + \beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r}$. A positive value of this derivative implies that an increase in labor income taxes ($\beta_w \downarrow$) decreases welfare, while a negative value implies that labor income taxes increase welfare. If ω is uncorrelated with n and e, we can substitute $\bar{\omega}$ for ω^j in (12). This is because for any Z uncorrelated with ω, $E(\omega Z) = E(\omega) \cdot E(Z)$ or $\sum_j \omega^j Z^j = \bar{\omega} \sum_j Z_j$.

If endowments are homogeneous, $e^j = \bar{e}$ and the second term in the numerator is zero. The sign of (12) then is determined by the sign of $\sum_j (n^j - \bar{n})/(\alpha + \bar{e} + \beta_w n^j + \frac{\alpha+\beta_w n^j}{1+r\beta_r})$. If the numerator is positive ($n^j > \bar{n}$), the denominator is larger than average, that is, positive values obtain a rather low weight. If the numerator is negative ($n^j < \bar{n}$), on the other hand, the denominator is smaller than average, negative values thus obtain a rather high weight. As

[6] Throughout the remainder of the text, $\bar{x}$ refers to the arithmetic average of x, that is, $\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_j$.

[7] For the derivation calculate λ from $\frac{\partial W}{\partial \alpha} = 0$, plug it into $\frac{\partial W}{\partial \beta_w}$ and simplify.

a consequence, the sum is negative. The higher the inequality in n, the lower the sum. This proves part (ii) of the proposition.

If productivities are homogeneous, $n^j = \bar{n}$ and the first term in the numerator is zero. The sign of (12) then is determined by the sign of $-\sum_j (e^j - \bar{e})/(\alpha + e^j + \beta_w \bar{n} + \frac{\alpha + \beta_w \bar{n}}{1 + r\beta_r})$. Using the same arguments as above, it follows that the sum is positive, rising in the inequality in e. This proves part (iii) of the proposition. ■

Heterogeneity in e, for example, is measured by $\sum \frac{e^j - \bar{e}}{inc^j}$ which can be written as $Cov\left(e, \frac{1}{inc}\right)$. If n and e are not correlated, a rising inequality in the sense of a mean preserving spread always implies a decrease in the covariance. A positive correlation between n and e further decreases the value of the covariance while a negative correlation increases its value such that, eventually, the covariance is positive. This is the case if households with high endowments have a lower than average potential income ($inc^j < \overline{inc}$) because of their very low productivity.

Intuitively, if only productivities are heterogeneous more productive households work more than less productive ones and pay more taxes. Thus, at the margin, labor income taxation is redistributive and increases welfare. We can see this from (12) since, given labor taxes are zero, the marginal welfare effect of an increase in labor taxes is positive, independently of capital income taxes. The magnitude of capital income taxes determines the size of the marginal improvement only through its influence on the size of the denominator and on α via the planner's budget constraint. If labor income taxes continue to increase, people work less and the pie to be divided shrinks until eventually this negative effect of labor income taxation dominates. Redistribution thus is limited by the households' response to the increasing tax rates.

If only endowments are heterogeneous, the derivative is positive and, again, the sign is independent of the size of capital income taxes (unless they are larger than 100%). That is, if households have identical earning abilities while their wealth levels differ, then it is optimal to pay a subsidy on wages and to impose lump-sum taxes rather than to tax wages. The intuition goes as follows: since wealthier households generally work less, they benefit less from the subsidies while paying the same lump-sum tax.

If agents differ only with respect to their productivities, more productive agents have a higher income. This property is called "agent monotonicity" and implies that income taxes redistribute from highly productive agents to less productive agents since agents with high income also have a high productivity. Earlier work on optimal income taxes (Sheshinski 1972, for example) has shown that under this assumption the optimal linear tax schedule consists of a positive transfer and a marginal income tax rate which is strictly positive and less than one. This result is consistent with part (ii) of proposition 1, where agent monotonicity holds. It is violated in part (iii), however, since agents with high endowments work less and thus generate less labor income than agents with small endowments.

If both, endowments and productivities, are heterogeneous and uncorrelated the optimality of either taxes or subsidies depends on the relative size of the heterogeneities and on the households' relative valuation of leisure and consumption, a. The value of a determines the relative weights of the heterogeneities in (12).

If the welfare weights vary across households, their correlation with n and e is crucial for the determination of the optimal tax rate. If there is no correlation, the above relationships hold. A negative correlation e.g. between the weights and endowments increases the marginal welfare effects of labor taxes, while a positive correlation decreases them and could even lead to opposite effects. Intuitively, if the government favors wealthier households, who generally work less than poorer households, labor subsidies are less desirable.

Now consider the tax rate on capital income, $1-\beta_r$. The following proposition establishes the main relations between the capital income tax rate and welfare.

PROPOSITION 2 *Given the utility and welfare functions specified above, and assuming that welfare weights, productivities, and endowments are uncorrelated, the capital income tax $1-\beta_r$ is related to welfare as follows:*

i. (a) *An interest income tax exceeding 100% ($\beta_r < 0$) may be optimal if the inequality in endowments exceeds a lower bound for given weights w^j.*

(b) *A confiscation of capital as well as interest is never optimal, i.e. $\beta_r^* > -\frac{1}{r}$.*

ii. Capital income taxes increase welfare if either endowments are heterogeneous or productivities are heterogeneous and $\rho(1+r) > 1$, or both. The larger the heterogeneity as measured by $\sum_j \frac{e^j-\bar{e}}{inc^j}$ and $\sum_j \frac{n^j-\bar{n}}{inc^j}$, the larger the marginal welfare increase.

iii. Capital income taxes decrease welfare if only productivities are heterogeneous and $\rho(1+r) < 1$. The larger the heterogeneity as measured by $\sum_j \frac{n^j-\bar{n}}{inc^j}$, the larger the marginal welfare decrease.

Proof. If capital income taxes are 100%, interest earnings are taxed away. However, a household might still smooth its consumption over time by saving some of its endowments for the second period. If capital income taxes are 100% ($\beta_r = 0$), the marginal welfare impact of lowering capital income taxes ($\beta_r \uparrow$) is

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=0} = r\sum_{j=1}^{N}\omega^j \cdot \left\{\frac{\rho(e^j-\bar{e})-(1-\rho)\beta_w(n^j-\bar{n})}{inc^j} + \frac{r\rho}{1+r+\rho}\cdot\frac{\overline{inc}}{inc^j}\right\} \quad . \tag{13}$$

If ω, n and e are uncorrelated, we can replace ω^j by $\bar{\omega}$, as shown in the proof to proposition 1. The only negative term in (13) is $\sum_j \frac{e^j - \bar{e}}{inc^j}$. This term is large in absolute terms if the inequality in endowments is very high. Thus, the derivative in (13) decreases in the endowments' inequality. Given the weights ω^j, $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=0} = 0$ implicitly defines a lower bound for heterogeneity in endowments which leads to capital income taxes being higher than 100%. This proves part (ia).

While the optimal capital income tax may exceed 100%, it is never optimal to tax away all savings. If $\beta_r = -\frac{1}{r}$, nobody saves and the planner does not collect any revenue from capital taxation. This settles part (ib).

For parts (ii) and (iii), consider the marginal welfare effects if there are no capital income taxes ($\beta_r = 1$):

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = r \sum_{j=1}^{N} \omega^j \frac{\rho(e^j - \bar{e}) + \beta_w \left(\rho - \frac{1}{1+r}\right) [n^j - \bar{n}]}{(r+1) \cdot inc^j} \tag{14}$$

Again, a positive value of the derivative indicates that capital income taxes decrease welfare while a negative value indicates that they increase welfare.

The arguments here are similar to those given for proposition 1. First, we can substitute $\bar{\omega}$ for ω^j. Second, the terms $\sum_j \frac{e^j - \bar{e}}{inc^j}$ and $\sum_j \frac{n^j - \bar{n}}{inc^j}$ are negative if n and e vary across agents and are not correlated. If $\rho > \frac{1}{1+r}$, the partial derivative is negative and increases in the heterogeneity of n and e, proving part (ii). If $\rho < \frac{1}{1+r}$, however, both sources of heterogeneity work in different directions. If only endowments are heterogeneous, (14) is negative and capital income taxes increase welfare. If only productivities are heterogeneous, (14) is positive and capital income taxes decrease welfare. This proves part (iii). ■

Let us spend a few more thoughts on (14). If productivity is the single source of heterogeneity, then $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1}$ is negative as long as $\rho > 1/(1+r)$. The restriction on ρ implies that the households' discount rate $\frac{1}{\rho} - 1$ is less than the interest rate r and is related to the households' optimal savings decision. If $\rho > 1/(1+r)$ more productive households save more to shift more utility to the second period and thus pay a higher amount of capital income taxes than less productive households; redistribution occurs through capital income taxes and transfers. If $\rho < 1/(1+r)$ more productive households save less and thus gain less from capital income subsidies than less productive households; redistribution occurs through capital income subsidies.

If only endowments are heterogeneous the argument is similar: households with higher endowments save more and, thus, pay more capital income taxes than households with lower endowments.

If both, endowments and productivities, are heterogeneous and if $\rho(1+r) > 1$ the effects reinforce each other if n and e are positively correlated. Positive values of $(e^j - \bar{e})$ then go along with an even higher value for inc^j than with no correlation. Negative values of $(e^j - \bar{e})$ go along with an even lower value for inc^j than with no correlation. Thus, the negative sum decreases further if n and e are correlated. We can use the same line of argument for the term $\sum_j \frac{n^j - \bar{n}}{inc^j}$.

If n and e are negatively correlated, the heterogeneities work in different directions and, thus, the marginal welfare increase through capital income taxation decreases. If people have either high endowments or high productivities, redistribution is not welfare enhancing since each household has a different mixture of income sources, leading to roughly the same utility levels.

The effects of varying welfare weights again depend on correlations. If $\rho(1+r) > 1$ and if the weights are positively correlated with productivities and endowments—implying that the government favors wealthy and productive households—the marginal welfare effects of capital income taxes decrease. A negative correlation, on the other hand, increases the marginal welfare effects of capital income taxes.

To summarize, we found that the influence of heterogeneity on the optimal tax rates depends strongly on the source of the heterogeneity and on possible correlations between the different sources. While labor taxes are welfare enhancing if productivities are heterogeneous, they can reduce welfare if endowments are heterogeneous. Capital income taxes increase welfare if endowments are heterogeneous while the effect of heterogeneous productivities depends on the sign of $\rho - \frac{1}{1+r}$.

If welfare weights vary across households, the optimal tax rates crucially depend on the weights' correlation with the households' endowments and productivities. The results confirm the intuition: if the government favors well-to-do households, marginal welfare effects of taxation are lower; if it favors poorer households, they are higher.

4 Optimal taxation with time-dependent labor taxes

Up to now, the planner was restricted to tax labor income in both periods with the same tax rate $1-\beta_w$. The assumption of identical labor income taxes is quite restrictive, however. The present chapter modifies the above analysis so that two different labor income tax rates are analyzed. $(1-\beta_{w1})$ now refers to the first period's labor income tax and $(1-\beta_{w2})$ to the second period's labor income tax. As before, we assume that the government credibly commits to the second period's labor and capital income taxes before households make their labor/leisure choices. What are the effects on the relation between capital income taxes and welfare?

Proposition 3 *Consider the setup as described above, with labor income taxes not restricted to be equal in both periods. If labor taxes are at their optimal values, then, starting from a capital income tax rate of zero:*

i. An increase in the capital income tax rate has no first order effects if endowments are homogeneous.

ii. An increase in the capital income tax rate increases welfare if endowments are heterogeneous and the correlation between n and e is not too negative, given the weights ω^j.

iii. An increase in the capital income tax rate decreases welfare if endowments and productivities are heterogeneous and their correlation is sufficiently negative, given the weights ω^j.

Proof. If there are no capital income taxes and if labor taxes are at their optimal values, the marginal welfare impact of capital income taxation is[8]

$$\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = \frac{r\rho}{1+r} \cdot \sum_{j=1}^{N} \left\{ \frac{\omega^j (e^j - \bar{e})}{inc^j} \right\} \quad . \tag{15}$$

It follows directly that $\left.\frac{\partial W}{\partial \beta_r}\right|_{\beta_r=1} = 0$ if endowments are homogeneous, proving part (i). The derivative is negative if e is heterogeneous and not correlated with n and ω. Only if there is a sufficiently negative correlation, the derivative is positive since negative values of the sum ($e^j < \bar{e}$) go along with very high values of n^j leading to lower than average weights ($inc^j < \overline{inc}$). The critical level is implicitly given by $\sum \frac{\omega^j (e^j - \bar{e})}{inc^j} = 0$. This is captured in parts (ii) and (iii) of the proposition. ■

Proposition 3 indicates that the optimality of a positive tax rate on capital income is solely driven by heterogeneity in endowments. If households are different with respect to their productivities only, it is optimal not to tax capital income.

Although differences in productivities do not call for capital income taxes by themselves, correlations of n and e play an important role. If both variables are not correlated, capital income taxes increase welfare. If they are positively correlated, the marginal welfare effect of capital income taxation increases since households with higher interest earnings tend to be more productive as well and thus have a higher income than others. This can also be seen in (15). For given e^j, a positive correlation between e and n implies that the positive terms of the sum ($e^j - \bar{e} > 0$) obtain an even lower weight than without correlation since not only e^j but also n^j are above their average values and, thus, inc^j is very high. Correspondingly, the negative terms obtain a higher weight if n and e are positively correlated.

If they are negatively correlated, the size of the marginal welfare improvement decreases since households with higher endowments tend to be less productive. For very high levels of negative correlation—where households with higher endowments tend to have lower overall utility than others because of their low productivity—it is desirable to pay interest subsidies instead of

[8]For the derivation of this expression see Appendix A.1.

imposing taxes. Again, this can be seen in (15). If positive values of $e^j - \bar{e}$ go along with very low values of n^j, the corresponding inc^j might be lower than average. That is, the argument given above is reversed: positive values of $e^j - \bar{e}$ obtain a high weight (low inc^j) while negative values obtain a low weight (high inc^j) and the sum is positive; the introduction of capital income taxes decreases welfare.

If the weights are not identical across households, correlations again play an important role. A negative correlation between weights and endowments or productivities—implying that the government favors poorer households—strengthens the marginal welfare improvement of capital income taxation. A positive correlation, on the other hand, lowers the positive impact of capital income taxation. For high levels of positive correlation—implying that the government strongly favors the wealthy and productive households—the marginal welfare impact of capital income taxation can become negative, making it optimal to subsidize interest income.

How does this compare to the literature? For one thing, we can interpret consumption and leisure in different periods as different goods and thus obtain results in line with Atkinson and Stiglitz (1976). Their agents differ only with respect to productivities and they find that uniform commodity taxation is optimal. Furthermore we find that if heterogeneity is one-dimensional (if either only endowments or only productivities are heterogeneous) the tax rate on the respective source of income is positive, which is compatible with the findings in Sheshinski (1972). If the heterogeneity is two-dimensional, this result no longer holds. For example, it might be optimal to subsidize labor income if endowments and productivities are negatively correlated. The comparison with the steady state results of infinite horizon general equilibrium models (Judd 1985, Chari and Kehoe 1999, for example), shows that while the optimal capital income tax rate is zero in the steady state, this result does generally not hold off the steady state. In this respect the optimal tax scheme differs from the analysis of representative agents as e.g. in Chamley (1986), where the tax rate is zero after a certain point in time, independent of the existence of a steady state. For a closer comparison between the different approaches see chapter 5.

Now consider labor income taxes. A question which comes to mind is whether it is optimal to tax labor income in both periods at the same rates. If not, which rate ought to be higher?

PROPOSITION 4 *Consider the setup as described above, with labor income taxes not restricted to be equal in both periods. Assume that the weights ω are not correlated with n, e and that $\frac{\partial W}{\partial \beta_{w2}}$ is monotone in a sufficiently large neighborhood of β_{w2}^*. The labor income tax rates $1 - \beta_{w1}^*$ and $1 - \beta_{w2}^*$ which fulfill the necessary conditions are related as follows:*

i. $1 - \beta_{w1}^* \geq 1 - \beta_{w2}^*$ *if the households' discount rate is lower than the net interest rate ($\rho > \frac{1}{1+r\beta_r}$) and capital income taxes are sufficiently low.*

ii. $1-\beta^*_{w1} \leq 1-\beta^*_{w2}$ *if the households' discount rate is higher than the net interest rate* ($\rho < \frac{1}{1+r\beta_r}$) *or if capital income taxes are sufficiently high.*

iii. $\beta^*_{w1} = \frac{1+r}{1+r\beta_r}\beta^*_{w2}$ *if productivities are homogeneous.*

Proof. The partial derivatives of welfare with respect to labor income taxes are related as follows:

$$\begin{aligned}\left.\frac{\partial W}{\partial \beta_{w2}}\right|_{\beta_{w2}=\tilde{\beta}_w} &= \frac{1}{1+r\beta_r}\left.\frac{\partial W}{\partial \beta_{w1}}\right|_{\beta_{w1}=\tilde{\beta}_w} \\ &+\frac{a}{\tilde{\beta}_w}\sum_j \omega^j\left\{\left[Z\cdot\frac{\overline{inc}}{inc^j}-1\right]\left[\rho-\frac{1}{1+r\beta_r}\right]-Z\cdot\frac{r(1-\beta_r)}{(r\beta_r+1)^2}\frac{\overline{inc}}{inc^j}\right\}\ ,\end{aligned} \tag{16}$$

where $Z=\left[\left(\tilde{\beta}_w+a(1-\tilde{\beta}_w)\right)\left(1+\frac{(1-\beta_r)r}{1+\rho}\frac{1}{r\beta_r+1}\right)\right]^{-1}$.

First assume $\beta_r = 1$. Thus $Z = \frac{1}{\tilde{\beta}_w(1-a)+a} \geq 1\ \forall\tilde{\beta}_w \leq 1$. The second term of the sum is zero. The first term can also be expressed as $N[Z-1-Z\cdot Cov(inc,\frac{1}{inc})]$, which is positive since $Z \geq 1$. Thus, if $\rho > \frac{1}{1+r\beta_r}$, the sum in the above equation is positive. If β_{w1} is at its optimal value at $\tilde{\beta}_w$, $\frac{\partial W}{\partial \beta_{w1}} = 0$ while $\frac{\partial W}{\partial \beta_{w2}} > 0$. Assuming monotonicity of $\frac{\partial W}{\partial \beta_{w2}}$ in a sufficiently large neighborhood of $\tilde{\beta}_w$, second period's labor income is taxed at a lower rate than first period's, settling (i). If, on the other hand, $\rho-\frac{1}{1+r\beta_r} < 0$, second period's labor income is taxed at a higher rate than first period's, settling the first part of (ii).

Now consider what happens if capital income taxes increase, that is, β_r decreases from 1. Z decreases and a negative term is added to the sum in (16). That is, the value of the sum decreases and eventually becomes negative. We can implicitly define a lower limit for β_r through $\sum_j \omega^j\left[Z\cdot\frac{\overline{inc}}{inc^j}-1\right]\cdot\left[\rho-\frac{1}{1+r\beta_r}\right]-Z\cdot\frac{r(1-\beta_r)}{(r\beta_r+1)^2}\frac{\overline{inc}}{inc^j}=0$. If β_r is above this limit, Z is sufficiently high to guarantee that the sum is positive, (i) holds. If β_r is below this limit, however, the sum is negative and (ii) holds.

If n is homogeneous, the partial derivatives can be simplified to $\frac{\partial W}{\partial \beta_{w1}} = \frac{Na}{\beta_{w1}}[\frac{r\beta_r+2}{X\beta_{w1}}\frac{1+r}{r\beta_r+1}(1-Cov(inc,\frac{1}{inc}))-1]$ and $\frac{\partial W}{\partial \beta_{w2}} = \frac{Na\rho}{\beta_{w2}}[\frac{r\beta_r+2}{X\beta_{w2}}(1-Cov(inc,\frac{1}{inc}))-1]$, where X is the derivative of the budget constraint with respect to α. If both terms in parentheses are zero, $\beta_{w1}=\frac{1+r}{1+r\beta_r}\beta_{w2}$, settling part ($iii$) of the proposition. ■

To get some intuition for this result, assume endowments are homogeneous. More productive households then work more than less productive ones in both periods. If $\rho(1+r\beta_r) > 1$, all households work more in the first period than in the second period. As a consequence, taxes in the first period lead to more revenues than taxes in the second period and are more effective for redistribution.

If productivities are homogeneous, labor income taxes are related such that the present-value taxes on labor income in both periods—β_{w1} for the first period and $\beta_{w2}\frac{1+r}{1+r\beta_r}$ for the second period—are identical. In other words, the optimal labor income taxes cancel out the distorting effects of capital income taxes on the labor supply.

How do these results compare with the previous chapter? The expressions determining the marginal welfare impact of capital income taxes are very similar (equations 14 and 15). Optimal capital income taxes are largely determined by the correlation between endowments and productivities in both cases. If the planner is restricted to identical labor taxes in both periods, capital income taxes are employed to substitute for the loss of flexibility. That is, even if endowments are homogeneous, the planner will tax interest income. If he can vary labor income taxes over time, it is optimal for him not to tax capital income

5 Robustness

The preceding chapter has established properties of optimal tax rates in a two period model. Since the analysis has been based on a model with logarithmic preferences and linear production, the present chapter asks whether a relaxation of these assumptions would change these properties. To this purpose, we compare our model with the different sets of assumptions in the existing literature.

Section 5.1 shows under which conditions optimal capital income tax rates off the steady state can be zero, while section 5.2 shows how and why the necessary sets of assumptions differ with time-separable utility and in the steady state. Finally, section 5.3 briefly discusses properties of optimal labor income taxes over time.

5.1 Optimal capital income tax rates off the steady state

The existing literature does not provide an analysis of optimal capital income taxation with heterogeneous agents off the steady state.[9] However, there is a strand of literature based on Atkinson and Stiglitz (1976) that considers optimal taxation with multiple factors and heterogeneous agents. Their results are relevant for our analysis since labor in different periods and endowments can be interpreted as different factors of production.[10] While our results presented in chapters 3 and 4 go beyond the conclusions on optimal tax rates obtained from this literature, it shows under which conditions optimal tax rates are zero. Applying this literature to our problem, we obtain the following proposition.

[9] One exception is Chamley (1998). The focus of his paper, however, is on credit market constraints.

[10] For a detailed account of the relation between static and dynamic models of this kind see the first sections in Judd (1997) and (1999).

PROPOSITION 5 *Suppose there is an arbitrary number of agents who are heterogeneous with respect to their productivities and goods endowments. Utility is separable between leisure and consumption and is strictly concave in all arguments. A social planner sets linear tax rates on labor and capital income and distributes lump-sum transfers to maximize welfare, which is defined as the sum of all agents' utilities. Starting from a capital income tax rate of zero, an increase in the capital income tax rate has no first order effects if the following four conditions hold.*

i. Preferences for consumption are homothetic.

ii. Preferences for consumption are identical for all households.

iii. Goods endowments are homogeneous or proportional to actual consumption.[11]

iv. Production is weakly separable between labor and capital or agents have identical productivities.

Items (i) to (iii) correspond to the findings in Bassetto (1999). In a setting similar to Atkinson and Stiglitz (1976), he analyzes properties of optimal commodity taxes for general homothetic and separable utility functions. His model considers two types of agents of whom only one works and he finds that homothetic and identical preferences as well as homogeneous or proportional endowments are necessary for the optimality of uniform commodity taxes or, in a dynamic interpretation, for a zero tax on capital income. Thus, the relevant assumption in our model is the not log-linearity of preferences per se but the implicit homotheticity.

Item (iv) of proposition 5 corresponds to Naito (1999). Re-examining optimal commodity taxation in a setting close to Atkinson and Stiglitz (1976), he shows that the optimality of uniform commodity taxes is not robust to the introduction of production functions that are not weakly separable between labor and consumption. Thus, the linearity of production by itself—as assumed in our model—is not necessary for the optimality of a zero tax rate on capital income. The optimal tax rate is zero as long as production is weakly separable between capital and leisure.

In the two period model delineated in chapter 2, items (i), (ii), and (iv) of proposition 5 hold. Proposition 3 shows that the optimal capital income tax rate is zero only if endowments are homogeneous (corresponding to item iii), while it is generally non-zero for heterogeneous endowments—with the exception of a negative correlation between endowments and productivities that is just large enough to cancel out the effects of capital income taxes on welfare. Thus, our earlier findings are compatible with the literature summarized in proposition 5. For a formal derivation of proposition 5, see appendix A.2.1.

[11] We use the term endowments here to be consistent with the rest of the paper. The endowments can be any exogeneous and deterministic revenues in current or future periods.

The following paragraphs discuss the conditions of proposition 5 and show why each of them is necessary for the optimality of a zero tax on capital incom.

Homogeneous goods endowments Heterogeneous endowments distort the optimality of uniform taxation since they lead to different intertemporal trading patterns among agents. If wealthier agents have higher capital holdings due to higher endowments, a taxation of capital income extracts high revenues from the wealthier agents which can be used for redistribution. The tax payments are directly related to differences in endowments that are generally not captured by the revenues of linear labor income taxes which are proportional to productivities.

Homothetic and identical preferences If preferences are not homothetic, luxury goods may exist. That is, wealthier agents consume disproportionately more of the luxury goods than poorer agents. Thus, while labor income taxes are proportional to the agents' productivities, higher tax rates on the luxury good disproportionately tax the wealthy and thus provide a means to redistribute.

In a dynamic interpretation this example translates as follows. If, for example, the desire for consumption in later periods increases in income, wealthier households save disproportionately more than poorer ones. A tax on capital income thus disproportionately affects the wealthy. If preferences differ across households, a similar mechanism works.

Homothetic preferences, however, are required only in the presence of linear labor income taxes. In the setup of Atkinson and Stiglitz (1976), labor income taxes are non-linear and the authors find that uniform commodity taxation is optimal if labor income taxes are sufficiently flexible. Homotheticity is not required for this result since non-linear labor income taxes already provide a means to tax the wealthy disproportionately. Even if there are luxury goods, an additional tax on these will not improve welfare if labor income taxes follow an optimal disproportionate scheme.

Weakly separable production The last requirement in proposition 5 concerns the production side of the economy. To see why this assumption is important, consider the following example. Let the production function be such that a rising capital usage in production implies that the relative productivity of low productivity households decreases. If this is the case the government might want to discourage capital accumulation by taxing capital income in order to prevent a higher discrepancy in relative productivities. In other words, if production is not weakly separable between labor and capital, capital income taxes might influence relative productivities. As a consequence, non-zero capital income taxes can be optimal even if endowments are homogeneous and preferences are homothetic and identical across households.

5.2 Time separable utility and optimal taxation in the steady state

The results presented in chapters 3 and 4 are based on time separable utility functions. In how far does this assumption change the requirements for the optimality of a zero tax on capital income as presented in proposition 5? Which further changes occur in the steady state?

PROPOSITION 6 *Consider the same setting as in proposition 5 and assume that utility is time-separable. Starting from a capital income tax rate of zero, an increase in the capital income tax rate in period t has no first order effects if the following conditions hold.*

i. Preferences for consumption are homothetic and identical for all households.

ii. Goods endowments in periods t and $t-1$ are homogeneous or proportional to actual consumption.

iii. Production is weakly separable between labor and capital or agents have identical productivities.

In the steady state, item (iii) is sufficient for an increase in the capital income tax rate to have no first order effects.

While homotheticity and weakly separable production are required for the same reasons as before, time separability limits the effects of heterogeneous endowments to two periods. Capital income taxes in the period with heterogeneous endowments are used to redistribute while the following period's capital income taxes ensure that relative prices in all other periods are not affected. Longer lasting effects occur only if the capital income tax rate is restricted e.g. to be no larger than 100%. This intuition lies behind the result in Chamley (1998) who finds that the optimal capital income tax rate is zero in finite time, even if endowments are heterogeneous.

The mechanisms working in the steady state are similar to those off the steady state. Weakly separable production ensures that capital income taxes do not change relative productivities. The necessity of this assumption for the optimality of a zero tax on capital income has been shown by Stiglitz (1985) and Chari and Kehoe (1999). Homotheticity is not required in the steady state since consumption is constant by definition, that is, it is not possible to disproportionately tax some households by varying the tax rate in different periods. By definition, there are no heterogeneous endowments in the steady state. Because of time-separable utility, heterogeneous endowments in earlier periods have no effect on the optimal capital income tax rate in the steady state. For a formal treatment, see appendix A.2.3.

5.3 Optimal labor income taxes off the steady state

In chapter 2 we have shown that the optimal labor income taxes are generally non-zero and vary over time. These results are in line with most of the literature since optimal labor income taxes are generally found to be positive and varying over time, see for example Chari and Kehoe (1999), Chari, Christiano, and Kehoe (1994), and Chamley (1986).

If productivities are homogeneous, however, we have found that the present value labor income taxes are constant for all periods. The following proposition shows under which conditions the last conclusion remains valid in a more general setting.

PROPOSITION 7 *A departure from uniform labor income taxation has no first order effects on welfare if the following conditions hold.*

i. Preferences for labor are linearly homogeneous.

ii. Preferences for labor are identical for all households.

iii. Productivities are homogeneous.

The intuition behind the conditions is similar to the discussion in the previous sections. Homothetic and identical preferences ensure that the relative labor supply in different periods does not vary across households with different wealth. Thus, varying labor income taxes over time would not tax wealthy households disproportionately. If productivities are heterogeneous, homothetic preferences are not sufficient to ensure a proportional labor supply for all households. Homogeneity of endowments, however, is irrelevant for the optimality of uniform labor income taxation. For a formal analysis see appendix A.2.2.

Proposition 7 confirms the findings from our two period model that it is not optimal to distort relative wages if productivities are homogeneous (see proposition 4 (*iii*)). Again, we find that not the assumption of log-linear preferences per se but the implicit homotheticity drives the result.

6 Conclusion

The analysis in the previous chapters shows that heterogeneity among households considerably influences optimal tax rates. While steady state analyses as, for example, Judd (1985) have found that the optimal capital income tax rate in the steady state is zero, we show that off the steady state this is generally not the case.

In a two period model with arbitrarily many heterogeneous households, we find that if households are heterogeneous with respect to productivities and endowments, capital income taxes generally increase welfare. If they are heterogeneous only with respect to productivities, endowments being identical for all, it is optimal not to tax capital income. The extent of the inequality and the joint distribution of its different components (productivities and endowments in our model) are crucial for the size of the marginal welfare effects of taxation. A positive correlation between endowments and productivities increases the marginal welfare effects of capital income taxation, while a negative correlation decreases the effects. Thus correlation of the households' characteristics plays an important role in determining the optimal tax policy.

Throughout the paper we emphasize the analogy between commodity taxes and capital income taxes, which effectively tax consumption goods at different points in time. Checking the robustness of our results with the help of the literature on optimal commodity taxation, we find that zero capital income taxation is optimal if endowments are homogeneous, if production is weakly separable between labor and capital, and if utility functions are homothetic and identical across agents. When we link our model to the infinite horizon steady state analyses as found in Judd (1985) and Chari and Kehoe (1999), we find that the models are compatible if we make appropriate assumptions.

In sum, we find evidence that if the planner maximizes a social welfare function and wants to redistribute, capital income taxes might be a good way to do so.

A Appendix

A.1 Derivation of the Partial Derivative with Respect to β_r

The partial derivatives of the planner's welfare function are modified the following way. Firstly, calculate λ from $\frac{\partial W}{\partial \alpha} = 0$, yielding

$$\lambda = \frac{\sum_{j=1}^{N} \frac{(\rho+1)(r\beta_r+2)}{(r\beta_r+1)\cdot inc^j}}{\sum_{j=1}^{N} 2 + r + \frac{a}{\rho+1}\left(\frac{1-\beta_{w1}}{\beta_{w1}}\frac{1+r}{1+r\beta_r} + \frac{1-\beta_{w2}}{\beta_{w2}}\rho\right)(r\beta_r+2) - (1-b_r)r\frac{\rho - \frac{1}{r\beta_r+1}}{\rho+1}} \quad ,$$

where $inc^j = \alpha + e^j + \beta_w n^j + \frac{\alpha + \beta_w n^j}{1+r\beta_r}$. Secondly, manipulate the other partial derivatives and substitute for λ to get

$$\frac{\partial W}{\partial \beta_r}\frac{1+r\beta_r}{r} + \frac{\partial W}{\partial \beta_{w2}}\frac{\beta_{w2}}{\rho+1} - \frac{\partial W}{\partial \beta_{w1}}\frac{\rho\beta_{w1}}{\rho+1} =$$

$$\rho \cdot \sum_{j=1}^{N}\left\{\frac{e^j - \bar{e}}{inc^j} + \frac{r(1-\beta_r)}{X(\rho+1)}\left(\frac{r\beta_r+2}{r\beta_r+1}\right)\left[(1-a)\frac{\overline{inc}}{inc^j} + \frac{2a}{inc^j}\left[\alpha + \bar{e} + \frac{\alpha}{r\beta_r+1}\right]\right]\right\} \quad ,$$

$$X = (2+r) + \frac{a}{\rho+1}\left[\frac{(1-\beta_{w1})}{\beta_{w1}}\frac{1+r}{r\beta_r+1} + \frac{1-\beta_{w2}}{\beta_{w2}}\rho\right](r\beta_r+2) - r(1-\beta_r)\frac{\rho - \frac{1}{1+r\beta_r}}{\rho+1} \quad .$$

If α, β_{w1}, and β_{w2} are at their optimal values, the respective partial derivatives are zero and one can easily calculate $\frac{\partial W}{\partial \beta_r}$ from the above equation.

A.2 Derivation of Propositions 5 to 7

The model in this section is an extension of Bassetto (1999).[12] There are N households, preferences are homothetic and separable between consumption and leisure. Household j's utility is given by

$$U^j(G^j(c^j), H^j(1-l^j)) \quad , \tag{17}$$

where $G^j(\cdot)$ is its subutility from consumption, and $H^j(\cdot)$ is its subutility from leisure.[13] $c^j = (c_1^j, c_2^j, ..)$ is the vector of its consumption with c_t^j being household j's consumption in period t, while $l^j = (l_1^j, l_2^j, ..)$ is the vector of its time spent working with l_t^j being household j's work time in period t. The endowment of time is one for each household in every period. The intertemporal technology constraint is given by

$$F\left(\sum_{j=1}^{N} c^j + g, l^1, .., l^N\right) \leq 0 \quad , \tag{18}$$

[12] Bassetto (1999) considers two agents, of whom only one works. The present paper extends his model to an arbitrary number of heterogeneous agents who are all working.

[13] In what follows, t, k refer to different periods, while j refers to households.

where g is government consumption and $F(\cdot)$ is assumed to be twice continuously differentiable, increasing in the first argument and decreasing in labor. In the following, we use the "primal approach" or "Ramsey approach" to determine properties of optimal tax rates.[14]

The firm produces consumption goods c_t, sells them at a price q_t and pays wages w_t^j. Wages are per unit of time and differ across households. The firm solves

$$\max_{c_t, l_t^j} \sum_t \left\{ q_t c_t - \sum_j w_t^j l_t^j \right\} \quad \text{s.t.} \quad F\left(\sum_{j=1}^{N} c^j + g, l^1, .., l^N \right) \leq 0 \quad . \tag{19}$$

From the firm's first order condition, wages and prices are determined as $q_t = -F_{c_t}/F_{l_1^1}$ and $w_t^j = F_{l_t^j}/F_{l_1^1}$, where w_1^1 is normalized to 1.[15]

Letting $\tau_{wt} w_t^j$ denote household j's after tax wages for labor in period t, p_t the consumer price for goods in period t, e_t^j household j's endowment in period t, and α the government transfers (which are distributed in the initial period only), then household j maximizes

$$\max_{c_t^j, l_t^j} U^j(G^j(c^j), H^j(1-l^j)) \ \text{ s.t. } \ \sum_t p_t(c_t^j - e_t^j) - \sum_t \tau_{wt} w_t^j l_t^j - \alpha = 0 \quad . \tag{20}$$

From the first order conditions, it follows that $\frac{G_t^j}{G_1^j} = \frac{p_t}{p_1}$, $\frac{H_t^j}{H_1^j}\frac{F_{l_t^1}}{F_{l_t^j}} = \frac{\tau_{wt}}{\tau_{w1}}$, and $\frac{\tau_{wt}}{p_t} = \frac{U_2^j H_t^j}{U_1^j G_t^j}\frac{F_{l_1^1}}{F_{l_t^j}}$, where functions with subscripts (U_k, H_k, G_k) denote partial derivatives and w_t^j has been substituted by $\frac{F_{l_t^j}}{F_{l_1^1}}$. The implementability constraint is derived as follows (normalizing $p_1 = 1$): $U_1^j G_t^j = U_1^j G_1^j p_t \Rightarrow \sum_t U_1^j G_t^j (c_t^j - e_t^j) = \sum_t U_1^j G_1^j p_t (c_t^j - e_t^j) = U_1^j G_1^j \alpha + U_1^j G_1^j \sum_t \tau_{wt} w_t^j l_t^j = U_1^j G_1^j \alpha + U_2^j \sum_t H_t^j l_t^j$. Thus,

$$U_1^j \sum_t G_t^j (c_t^j - e_t^j) - U_2^j \sum_t H_t^j l_t^j - U_1^j G_1^j \alpha \leq 0 \quad . \tag{21}$$

Since all agents face the same taxes on consumption goods, they face the same prices and, in competitive equilibrium, marginal rates of substitution must be equal for all households. That is,

$$G_t^1 G_1^j = G_1^1 G_t^j \quad \forall t, j \quad . \tag{22}$$

---

[14]For a comprehensive overview see Chari and Kehoe (1999).

[15]In the setup of the chapter 2, we would have $U^j(G^j, H^j) = (\rho+1)[(1-a)\ln(G^j) + a\ln(H^j)]$, $G^j(c_1^j, c_2^j) = [c_1^j c_2^{j\rho}]^{\frac{1}{\rho+1}}$, and $H^j(1-l_1^j, 1-l_2^j) = [(1-l_1^j)(1-l_2^j)^\rho]^{\frac{1}{\rho+1}}$. (18) would correspond to $\sum_j (e^j + w_1 n^j l_1^j - c_1^j)(1+r) + w_2 n^j l_2^j - c_2^j \geq 0$. Since we have no government spending besides redistribution, $g = 0$. The representative firm maximizes $\sum_j q_1(c_1^j - e^j) + q_2 c_2^j - w_1 n^j l_1^j - w_2 n_j l_2^j$ s.t. $F(c_1^j, c_2^j, n_j l_1^j, n_j l_2^j) = 0$. Normalizing $w_1 = 1$, we get $w_2 = \frac{1}{1+r}, q_1 = 1, q_2 = \frac{1}{1+r}$. Notice that prices and taxes are present-value in the setup of this chapter and that we tax consumption and labor instead of capital and labor. If we denote the tax rates here by τ_{ct} and τ_{wt} they are related to chapter 2 by $\tau_{w1} = \beta_{w1}$, $\tau_{w2} = \beta_{w2}\frac{1+r}{1+r\beta_r}$, $\tau_{c1} = 1$, and $\tau_{c2} = \frac{1+r}{1+r\beta_r}$.

A similar argument can be made for labor taxes; if labor taxes are equal for all households it follows from the households' necessary conditions that

$$\frac{H_t^1 F_{l_1^1}}{H_t^j F_{l_1^j}} = \frac{H_1^1 F_{l_t^1}}{H_1^j F_{l_t^j}} \quad \forall\ t, j \quad . \tag{23}$$

Since (22) and (23) ensure that tax rates as well as prices are identical for all agents, it also follows from the households' necessary conditions that

$$\frac{U_1^1 G_t^1}{U_2^1 H_t^1} \cdot F_{l_t^1} = \frac{U_1^j G_t^j}{U_2^j H_t^j} \cdot F_{l_t^j} \quad , \quad \forall j, t\ . \tag{24}$$

We use this equation in combination with (22) to manipulate the implementability constraint (21) for all households except household 1. First, from (22), substitute $\frac{G_1^j}{G_1^1} \cdot G_t^1$ for G_t^j. Second, multiply the modified constraint by $\frac{G_1^1}{G_1^j}\frac{U_1^1}{U_1^j}$. Third, from (24), substitute $\frac{G_t^j}{G_t^1 U_2^1 H_t^1} \cdot F_{l_t^j}/F_{l_t^1}$ for $\frac{H_t^j U_2^j V1^1}{U_1^j}$. The planner's maximization problem then is

$$\begin{aligned}
\max_{c_t^j, l_t^j, \alpha} \quad & \sum_j \omega^j U^j \left(G^j(c^j), H^j(1-l^j)\right) \\
& + \sum_j \lambda^j \left[U_1^1 \sum_k G_k^1 (c_k^j - e_k^j) - U_2^1 \sum_k \frac{F_{l_k^j}}{F_{l_k^1}} H_k^1 l_k^j - U_1^1 G_1^1 \alpha \right] \\
& + \sum_{j>1} \sum_{k>1} \nu_k^j \left[G_1^j G_k^1 - G_k^j G_1^1 \right] + \sum_{j>1}\sum_{k>1} \eta_k^j \left[\frac{H_k^1 F_{l_1^1}}{H_k^j F_{l_1^j}} - \frac{H_1^1 F_{l_k^1}}{H_1^j F_{l_k^j}} \right] \\
& - \mu F \left(\sum_j c^j + g, l^1, .., l^N \right) \quad ,
\end{aligned}$$

where $\lambda^j, \nu_k^j, \eta_k^j$, and μ are Lagrange multipliers.

A.2.1 Proof of Proposition 5

This section derives the properties of optimal commodity tax rates over time and closely follows Bassetto (1999). Commodity taxes are determined by $\frac{p_t}{q_t} = \frac{G_t^j}{F_{c_t}} \frac{F_{c_1}}{G_1^j} \frac{1}{q_1} \quad \forall j, t$, and commodity taxation is uniform if $\frac{G_t^j}{F_{c_t}}$ is independent of t. The following paragraphs examine under which conditions uniform commodity taxation is compatible with the planner's necessary conditions.

The first order condition w.r.t. α implies $\sum_j \lambda^j = 0$. Thus, all terms containing α drop out of the remaining first order conditions. If we rearrange the planner's first order conditions to isolate the terms violating the independence of $\frac{G_t^j}{F_{c_t}}$ from t on the right hand side of the equation, the derivative of the welfare function with respect to c_t^1 becomes

$$
\begin{aligned}
c_t^1: \quad & \frac{G_t^1}{F_{c_t}}\left\{\left(\omega^1+\lambda^1\right)U_1^1+U_{11}^1\sum_j\sum_k\lambda^j G_k^1\left(c_k^j-e_k^j\right)\right\} \qquad (25)\\
&=\mu-\frac{1}{F_{c_t}}\left\{U_1^1\sum_j\sum_k\lambda^j G_{kt}^1\left(c_k^j-e_k^j\right)-U_2^1\sum_{j>1}\sum_k\lambda^j H_k^1 l_k^j\Lambda_{k,c_t}^j\right.\\
&\left.+\sum_{j>1}\sum_{k>1}\nu_k^j\left[G_1^jG_{kt}^1-G_k^jG_{1t}^1\right]+\sum_{j>1}\sum_{k>1}\eta_k^j\left[\frac{H_k^1}{H_k^j}\Lambda_{1,c_t}^j-\frac{H_1^1}{H_1^j}\Lambda_{k,c_t}^j\right]\right\} \quad ,
\end{aligned}
$$

where $\Lambda_{k,c_t}^j=\frac{\partial\left(\frac{F_{l_t^1}}{F_t^j}\right)}{\partial c_t}$, which is the derivative with respect to c_t of the difference in productivity between households 1 and j. The r.h.s is independent of t if the following terms are equal to zero:

- $\sum_j\sum_k\lambda^jG_{tk}^1c_k^j$.

If G^1 is linearly homogeneous, G_t^1 is homogeneous of degree 0. Thus $\sum_k G_{tk}^1c_k^1=0$. If, in addition, the functions G^1 and G^j are identical, c^1 and c^j are proportional because of the equality of the marginal rates of substitution and homotheticity of preferences. Thus, the above sum is zero. An alternative possibility would be identical consumption for all types of agents.

- $\sum_j\sum_k\lambda^jG_{tk}^1e_k^j$.

Since $\sum_j\lambda^j=0$, this expression is zero if a) endowments are homogeneous or b) endowments are proportional to actual consumption and the previous term is zero.

- $\Lambda_{k,c_t}^j=\partial\left(\frac{F_{l_k^1}}{F_{l_k^j}}\right)/\partial c_t$.

This term equals zero if either productivities are identical for all agents or if the resource constraint (production function) is weakly separable between $(l_k^1,..,l_k^N)$ and c_t, which is equivalent to a production function that is weakly separable between labor and capital.

- $\sum_{j>1}\sum_{k>1}\nu_k^j(G_1^jG_{tk}^1-G_k^jG_{1t}^1)$.

ν_k^j is the multiplier on the equality of marginal rates of substitution. We can show that the constraint is not binding if the previous conditions hold and, thus, that $\nu_k^j=0\ \forall k,j$.

Assume that the above conditions hold and that $\nu_k^j=0\ \forall k,j$. The first order conditions with respect to consumption can then be expressed as

$$
c_t^1: \quad \frac{G_t^1}{F_{c_t}}\left\{\left(\omega^1+\lambda^1\right)U_1^1+U_{11}^1\sum_j\sum_k\lambda^jG_k^1c_k^j\right\}=\mu \qquad (26)
$$

$$c_t^j: \qquad \frac{G_t^j}{F_{c_t}}\left\{\omega^j U_1^j\right\} + \frac{G_t^1}{F_{c_t}}\left\{\lambda^j U_1^1\right\} = \mu \qquad \forall j > 1 \quad . \tag{27}$$

As a consequence of (26), G_t^1/F_{c_t} does not depend on t. From (27) we then see that, for any j, G_t^j/F_{c_t} does not depend on t either. Thus, G_t^1 and G_t^j are proportional or G_t^1/G_t^j is constant $\forall t, j$. Consequently, the constraint on the equality of marginal rates of substitution is not binding and indeed $\nu_k^j = 0 \ \forall k, j$.

Thus we have shown that the necessary conditions for the Ramsey problem outlined above hold with a capital income tax of zero if conditions (i) to (iv) of proposition 5 hold.

A.2.2 Proof of Proposition 7

This section derives properties of the optimal labor income taxes over time. Labor income taxes are determined by $\tau_{wt} = \frac{U_2^j}{U_1^j}\frac{p_1}{w_1^1}\frac{F_{l_1^1}}{G_1^j} \cdot \frac{H_t^j}{F_{l_t^j}}$. They are uniform if $H_t^j/F_{l_t^j}$ is independent of t. To analyze the properties of optimal labor income taxes, consider the first derivative of welfare with respect to l_t^1.

$$\begin{aligned} l_t^1 \ : \qquad & -\frac{H_t^1}{F_{l_t^1}}\left\{\omega^1 U_2^1 + \lambda^1 U_2^1 - U_{22}^1 \sum_j \lambda^j \sum_k H_k^1 l_k^j \Lambda_k^j\right\} \\ & = \mu + \frac{1}{F_{l_t^1}}\left\{U_2^1\left[\sum_j \lambda^j \sum_k H_{kt}^1 l_k^j \Lambda_k^j + \sum_{j>1}\lambda^j \sum_k H_k^1 l_k^j \Lambda_{k,l_t^1}^j\right]\right. \\ & \left. -\sum_{j>1}\sum_{k>1}\eta_k^j\left[\frac{H_{kt}^1}{H_k^j}\Lambda_1^j + \frac{H_k^1}{H_k^j}\Lambda_{1,l_t^1}^j - \frac{H_1^1}{H_1^j}\Lambda_{k,l_t^1}^j - \frac{H_{1t}^1}{H_1^j}\Lambda_k^j\right]\right\} \end{aligned} \tag{28}$$

The independence of $\frac{H_t^1}{F_{l_t^1}}$ from t is violated by the following terms.

- $\sum_{j>1}\sum_k \lambda^j H_k^1 l_k^j \Lambda_{k,l_t^1}^j$.

If the difference in productivities is constant, $\Lambda_{k,l_t^1}^j$ is zero and so is this sum.

- $\sum_j \sum_k \lambda^j H_{kt}^1 l_k^j \Lambda_k^j$.

If the difference in productivities is identical for all goods, $\Lambda_k^j = \Lambda^j \ \forall j, k$. If H is linearly homogeneous and identical across agents, this term can be modified to $-\sum_j \lambda^j \Lambda^j \sum_k H_{kt}^1$ since $\sum_k H_{kt}^1(1 - l_t^j) = 0$ (the argument is the same as given for consumption in the previous section). From the maximization with respect to α we know that $\sum_j \lambda^j = 0$. Thus, if Λ^j is independent of j, that is, if productivities are homogeneous, this term equals zero.

- $\sum_{j>1}\sum_{k>1}\eta_k^j\left[\frac{H_{kt}^1}{H_k^j}\Lambda_1^j + \frac{H_k^1}{H_k^j}\Lambda_{1,l_t^1}^j - \frac{H_1^1}{H_1^j}\Lambda_{k,l_t^1}^j - \frac{H_{1t}^1}{H_1^j}\Lambda_k^j\right]$.

This term is zero if $\eta_k^j = 0 \; \forall k, j$. In what follows, we show that if the above conditions hold the respective constraint is not binding, that is, $\eta_k^j = 0$.

Assume that productivities are homogeneous, that is, $\Lambda_k^j = 1 \; \forall j, k$, and that $\eta_k^j = 0 \; \forall k, j$. The first order conditions with respect to labor then are:

$$l_t^1 : \qquad -\frac{H_t^1}{F_{l_t^1}} \left\{ \left(\omega^1 + \lambda^1\right) U_2^1 - U_{22}^1 \sum_j \lambda^j \sum_k H_k^1 l_k^j \right\} = \mu \tag{29}$$

$$l_t^j : \qquad -\frac{H_t^j}{F_{l_t^j}} \left\{ \omega^j U_2^j \right\} - \frac{H_t^1}{F_{l_t^1}} \left\{ \lambda^j U_2^1 \right\} = \mu \qquad \forall j > 1 \quad . \tag{30}$$

From (29) it follows that $H_t^1 / F_{l_t^1}$ is independent of t. In combination with (30) it follows that $H_t^j / F_{l_t^j}$ is independent of t as well. Thus, H_t^1 and H_t^j are proportional and the constraint (23) is not binding since $F_{l_t^j} = F_{l_t^1} \; \forall j$ and, thus, $\eta_k^j = 0 \; \forall k, j$.

Thus we have shown that the necessary conditions for the Ramsey problem outlined above are fulfilled with uniform labor income taxation if conditions (i) to (iii) of proposition 7 are fulfilled.

A.2.3 Proof of Proposition 6

This section shows how the assumption of time-separable utility influences optimal tax rates. In particular, we show under which assumptions the steady state results of Judd (1985) and Chari and Kehoe (1999) can be derived in the setup of the previous sections.

If utility is time separable, $U_{11} G_t G_k + U_1 G_{tk} = 0$ or $G_{tk} = -G_k G_t \frac{U_{11}}{U_1} \; \forall \; k \neq t$. From (25), the terms violating the optimality of zero capital income taxes are

- $\sum_j \lambda^j \sum_k G_{kt}^1 e_k^j$.

If utility is time separable, $\sum_k G_{kt}^1 = -\sum_k \left\{ G_k^1 G_t^1 \frac{U_{11}^1}{U_1^1} \right\} + (G_t^1)^2 \frac{U_{11}^1}{U_1^1} + G_{tt}^1$ and

$$\begin{aligned} \sum_j \lambda^j \sum_k G_{kt}^1 e_k^j &= -G_t^1 \frac{U_{11}^1}{U_1^1} \sum_j \lambda^j \sum_k G_k^1 e_k^j \\ &\quad + \left[(G_t^1)^2 \frac{U_{11}^1}{U_1^1} + G_{tt}^1 \right] \sum_j \lambda^j e_t^j \quad . \end{aligned}$$

The first term on the r.h.s is of the form $G_t^1 \cdot X$, where X is independent of t. Thus, this term moves to the l.h.s. of (25), not violating the optimality of zero capital income taxes any more.

For the second term, remember that $\sum_j \lambda^j = 0$. Thus, if endowments in period t are homogeneous, then this term equals zero. In other words, if endowments are distributed only in period one, time separable utility implies that heterogeneous endowments have no effects on capital income taxes beyond period two.

- $\sum_j \lambda^j \sum_k G^1_{kt} c^j_k$.

There are two cases when this expression is zero. Firstly, if utility is homothetic and identical across agents, $\sum_k G^1_{kt} c^j_k = 0$ as shown in section A.2.1.[16] Secondly, if utility is time separable it follows that

$$\sum_j \lambda^j \sum_k G^1_{kt} c^j_k = -G^1_t \frac{U^1_{11}}{U^1_1} \sum_j \lambda^j \sum_k G^1_k c^j_k + \left[(G^1_t)^2 \frac{U^1_{11}}{U^1_1} + G^1_{tt} \right] \sum_j \lambda^j c^j_t \quad .$$

Again, the first term moves to the l.h.s. of (25), not violating the optimality of zero capital income taxes any more. The second term is zero if we are in a steady state, since consumption is constant there. Thus, if we are in a steady state and utility is time-separabel, the optimal capital income tax rate is zero even if preferences are not homothetic.

- $\sum_k H^1_k l^j_k \Lambda^j_{k,ct}$ and $\sum_{j>1} \sum_{k>1} \eta^j_k \left[\frac{H^1_k}{H^j_k} \Lambda^j_{1,ct} - \frac{H^1_1}{H^j_1} \Lambda^j_{k,ct} \right]$.

These terms are zero if production is weakly separable. This requirement has been shown in Stiglitz (1985) and Chari and Kehoe (1999).

- $\sum_{j>1} \sum_{k>1} \nu^j_k \left[G^j_1 G^1_{tk} - G^j_k G^1_{1t}) \right]$.

If preferences are homothetic and all above conditions hold, this term is zero as shown in the proof of proposition 5.

If we are in a steady state after period κ and preferences are not homothetic, we can divide the term into two parts. The first term is $\sum_{j>1} \sum_{k=2}^{\kappa} \nu^j_k \left[G^j_1 G^1_{tk} - G^j_k G^1_{1t}) \right]$, where $t > \kappa$. Since utility is time separable and $t \neq k$, we can substitute G^1_{tk} by $-G^1_k G^1_t \frac{U^1_{11}}{U^1_1}$ and G^1_{1t} by $-G^1_1 G^1_t \frac{U^1_{11}}{U^1_1}$. The remaining term can be simplified to $\sum_{j>1} \sum_{k=2}^{\kappa} G^1_t \nu^j_k \frac{U^1_{11}}{U^1_1} \left[-G^j_1 G^1_k + G^j_k G^1_1 \right]$. The term $-G^j_1 G^1_k + G^j_k G^1_1$ is zero since it is identical to the planner's constraint (equality of marginal rates of substitution). Given that this term is zero and that all above conditions hold, this condition is not binding in the steady state and thus $\nu^j_\kappa = 0 \ \forall \ \kappa, j$ as shown in the proof of proposition 5.

The preceding paragraphs have shown that if preferences are time-separable, heterogeneous endowments do not influence optimal capital income taxes in later periods. If preferences are homothetic, this implies that, starting from zero capital income taxes, an increase in capital income taxes has no first order effects. If we are in a steady state, there are no first order effects even if preferences are not homothetic. Weakly separable production again is needed for the absence of first order effects (proposition 6).

[16] This case corresponds to the findings of Chamley (1998), who shows that if preferences are time-separable and homothetic, if endowments are heterogeneous, then a zero tax on capital income is optimal after a finite time.

References

Aiyagari, S. R. (1995): "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Public Economics*, 103(6), 1158–75.

Atkinson, A. B., and J. Stiglitz (1976): "The Design of Tax Structure: Direct versus Indirect Taxation," *Journal of Public Economics*, 6, 55–75.

Bassetto, M. (1999): "Optimal Fiscal Policy with Heterogeneous Agents," Mimeo, Northwestern University.

Chamley, C. (1986): "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 54(3), 607–22.

——— (1998): "Capital Income Taxation, Wealth Distribution and Borrowing Constraints," Discussion Paper 9811, INSEE.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994): "Optimal Fiscal Policy in a Business Cycle Model," *Journal of Political Economy*, 102(4), 617–52.

Chari, V. V., and P. J. Kehoe (1999): "Optimal Fiscal and Monetary Policy," *NBER Working Paper*, #6891.

Fischer, S. (1980): "Dynamic Inconsistency, Cooperation and the Benevolent Dissembling Government," *Journal of Economic Dynamics and Control*, 2, 93–107.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1997): "On the Optimal Taxation of Capital Income," *Journal of Economic Theory*, 73, 93–117.

Judd, K. L. (1985): "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28, 59–83.

——— (1997): "The Optimal Tax Rate for Capital Income is Negative," *NBER Working Paper*, #6004.

——— (1999): "Optimal Taxation and Spending in General Competitive Growth Models," *Journal of Public Economics*, 71, 1–26.

Klein, P., and J.-V. Rios-Rull (1999): "Time Consistent Optimal Fiscal Policy," Mimeo, Institute for International Economic Studies, Stockholm University.

Milesi-Ferretti, G. M., and N. Roubini (1995): "Growth Effects of Income and Consumption Taxes: Positive and Normative Analysis," *NBER working paper*, #5317.

Naito, H. (1999): "Re-Examination of Uniform Commodity Taxes under a Non-Linear Income Tax System and its Implication for Production Efficiency," *Journal of Public Economics*, 71, 165–88.

Sheshinski, E. (1972): "The Optimal Linear Income-Tax," *Review of Economic Studies*, 39, 297–302.

Stiglitz, J. E. (1985): "Inequality and Capital Taxation," Discussion Paper 457, IMSSS, Stanford University.

Stockman, D. R. (1999): "Balance-Budget Rules: Welfare Loss and Optimal Policies," Mimeo, University of Delaware.