

Discussion Papers

409

Pio Baake
Rainald Borck
Andreas Löffler

Complexity and Progressivity in Income
Tax Design: Deductions for Work-
Related Expenses

Berlin, March 2004



DIW Berlin

German Institute
for Economic Research



Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

DIW Berlin

German Institute
for Economic Research

Königin-Luise-Str. 5
14195 Berlin,
Germany

Phone +49-30-897 89-0

Fax +49-30-897 89-200

www.diw.de

ISSN 1619-4535

Complexity and Progressivity in Income Tax Design: Deductions for Work-Related Expenses

Pio Baake*

DIW Berlin, Berlin, Germany

Rainald Borck

DIW Berlin, Berlin, Germany

Andreas Löffler

University of Hannover, Department of Economics, Hannover, Germany

Revised Version, February 24, 2004

Abstract

We analyze optimal income taxes with deductions for work-related or consumptive goods. We consider two cases. In the first case (called a complex tax system) the tax authorities can exactly distinguish between consumptive and work-related expenditures. In the second case (called a simple tax system) this distinction is not exact. Assuming additively separable utility functions, we show that work-related expenditures should be fully deductible in the first case while deduction rates should be less than 100 percent in the second case. Under further simplifying assumptions, we also show that the simple system can be characterized by higher tax burdens on low income earners and less redistribution.

Keywords: optimal taxation, tax deductions, work-related expenditures.

JEL classification: H21

*Corresponding author. DIW Berlin, 14191 Berlin, Germany, email: pbaake@diw.de

1. Introduction

Starting in the 1980s, income tax reforms in many countries focused on lowering marginal tax rates combined with the attempt to reduce the complexity of tax systems, e.g., by simplifying the regulations for admissible tax deductions. Most notable were differing proposals for a flat tax, whose proponents argued that taxpayers would need only a postcard to file their returns (see Atkinson, 1995). A move in this direction was the 1986 Tax Reform Act in the US: it introduced a tax schedule with only two brackets and increased the standard deduction, which meant that fewer households had to itemize their deductions. Similarly, recent German tax reforms were intended to decrease marginal tax rates and to standardize deductions for work-related expenditures.

Our focus is on the observation that the complexity and the rate structure of the tax systems were reformed simultaneously. We consider an optimal income tax system which may incorporate tax deductions for work-related expenditures. Following the literature on optimal taxation, we assume a simple utilitarian welfare function and heterogeneous individuals. We model complexity of the tax system by distinguishing two cases. In the first case, an individual's tax payment is based on her income and on a detailed distinction of whether income is used for consumptive or work-related purposes. We call this the case of a complex tax system. Our second case deals with a simplified tax system where deductible expenditures may include expenditures for consumption goods which are not (or cannot be) distinguished from work-related goods. That is, the simplified tax system requires less information concerning the use of an individual's income.

The distinction between these two cases is motivated by the observation that for some goods it may be easy to ascertain whether or not they are used for work-related purposes, for instance, advanced medical equipment. Others, however, may be used both for consumption and production, say, personal computers or company cars. In this case, it may be prohibitively costly for the tax authorities to monitor which part of expenditure is for work-related use and which part consumptive.

Under the assumption of additively separable utility functions, it turns out that work-related expenditures should be fully deductible in case of the complex system, while deduction of work-related and consumptive expenditures should be allowed at a rate of less than 100 per cent in the simple system. With utility functions which are separable in all goods, the simplified system leads to higher average income and to a shift of the tax burden to

low income earners. Hence, the simplified system provides higher (average) incentives to work and is less redistributive.

The intuitive explanation for these results is based on the incentive constraints which are at the heart of the problem of optimal taxation. Monitoring work-related expenditures and allowing for tax deductions mitigates the restrictions implied by incentive compatibility and thus allows for high marginal tax rates on income. Under a simplified tax system, tax deductions imply an additional negative effect by forcing individuals to increase those consumptive expenditures which are deductible. Hence, optimal tax deductions are lower. In order to satisfy incentive compatibility the optimal income tax system must be less progressive.¹

Our results are based on well known contributions to the optimal taxation literature. Starting with the seminal paper of Mirrlees (1971), several authors established that optimal tax schemes are non-linear and are characterized by marginal tax rates equal to zero at both the upper and lower end of the income distribution (see for example Atkinson and Stiglitz, 1980). General nonlinear taxation problems with income and excise taxes were considered by Atkinson and Stiglitz (1976), Mirrlees (1976) and Cooter (1978). When utility is weakly separable between leisure and consumption, the optimal tax system is characterized by a nonlinear income tax with zero excise taxes (Atkinson and Stiglitz, 1976; Cooter, 1978). Our model also assumes weak separability, but in addition we distinguish work-related and consumptive goods. We also effectively study two different nonlinear tax problems, where the complex system is similar to that of Mirrlees (1976) and the simple system has the additional constraint that two of the goods have the same tax schedule. Kaplow (1990) and Slemrod (1994) study optimal taxation in models where individuals can reduce their tax burden by spending part of their income on tax avoidance and evasion activities. Whereas evasion is illegal, and avoidance may be thought of as an implicit part of the tax code we consider tax deductions, which are an explicit part of the tax code and characterize the optimal income tax scheme when tax deductions are incorporated.

In the next section we present our model. In section 3 we focus on tax deductions. In section 4, we consider the progressivity of the optimal tax system. Under strict separability on the utility function, we show that the simplified system leads to higher incentives to work and to a higher tax burden on low incomes. The last section contains a short conclusion.

¹We use the term progressivity here in the sense that there is an interval of low income earners who pay higher taxes under the simple than under the complex system.

2. The Model

For simplicity we assume that all individuals have identical utility functions but they differ in a parameter θ which measures the individual's ability to work. Utility is increasing in the quantities c and s of two distinct consumption goods but decreases in the effort, e , which an individual must exert in order to earn income y . The distinction between the two consumption goods c and s is used to characterize the different informational requirements of the two tax systems considered below. Good s stands for a good whose consumptive use may not be easily separated from its work-related use (e.g., a personal computer or company car). While effort is increasing in y , it is decreasing in the quantity q of a work-related good and in the individual parameter θ . In the case of a computer, q thus measures the work-related use versus the consumptive use which is captured by s . We assume that utility is additively separable in consumption and labor:²

$$u(c, s) - e(y, q)h(\theta) \text{ with } u_c, u_s > 0 \text{ and } e_y > 0 > e_q; h' < 0 < h. \quad (1)$$

where subscripts denote partial derivatives. We assume that u is strictly concave and that e is strictly convex in y and q . Furthermore, we will impose the following assumptions on the utility function:

$$u_{cs} > u_{cc}, u_{ss} \text{ and } e_{yq} > -e_{qq}.$$
³

The ability to work parameter θ is distributed on an interval $[\underline{\theta}, \bar{\theta}]$ according to the distribution function $F(\theta)$ with density $f(\theta) > 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$. $F(\theta)$ is common knowledge, but only the agents know their individual parameter θ .

An individual's budget constraint depends on her income y and tax payment t . We assume linear production technologies for all goods and normalize all prices to one. The tax payment is a function of y and of the composition of expenditures. We distinguish two cases: In the first case the tax payment depends on expenditures for q and s separately. The tax function is written as $t(y, s, q)$ and we call this the case of a complex tax system. In the second case the tax system is simplified in that the tax function takes into account only the sum of the expenditures for q and s . We define $k := s + q$ and write the tax function as $t(y, k)$. That is, if the tax system allows for tax deductions on k , deductible expenditures may include expenditures for consumptive purposes.

²In the following we omit the arguments of the functions where this does not lead to any confusion.

³These are sufficient conditions for c and s to be normal goods and for the optimal y to be increasing in θ .

Given either $t(y, s, q)$ or $t(y, k)$ an individual solves

$$\max_{c, s, q, y} u(c, s) - e(y, q)h(\theta) \quad \text{s.t.} \quad y = c + q + s + t(y, \cdot). \quad (2)$$

In the following we assume that the (optimal) tax functions are differentiable and involve no bunching. This allows us to characterize the solutions $c^*(\theta)$, $s^*(\theta)$, $q^*(\theta)$ and $y^*(\theta)$ of (2) by the corresponding first order conditions. For the complex tax system we get

$$u_c(1 - t_y) - e_y h = 0, \quad \frac{u_s}{u_c} = 1 + t_s, \quad e_q = -e_y \frac{1 + t_q}{1 - t_y}, \quad (3)$$

while the simplified tax system implies

$$u_c(1 - t_y) - e_y h = 0, \quad u_s = -e_q h, \quad e_q = -e_y \frac{1 + t_k}{1 - t_y}. \quad (4)$$

Furthermore, defining $v^*(\theta) := u(c^*, s^*) - e(y^*, q^*)h$ as the indirect utility function the envelope theorem leads to

$$\frac{d}{d\theta} v^*(\theta) = -e(y^*, q^*)h'. \quad (5)$$

The government's aim is to design the tax functions $t(y, s, q)$ for the complex system and $t(y, k)$ for the simple system such that the sum of individual utilities is maximized subject to (5) and to some minimum tax requirement T :

$$\max_{t(y, \cdot)} W = \int_{\underline{\theta}}^{\bar{\theta}} [u(c^*, s^*) - e(y^*, q^*)h(\theta)] f d\theta \quad \text{s.t.} \quad \int_{\underline{\theta}}^{\bar{\theta}} t(y, \cdot) f d\theta = T. \quad (6)$$

3. Optimal Tax Deductions

We first focus on the optimal tax deductions, i.e., the optimal relation between the marginal tax rates t_y and t_q in the complex tax system and t_y and t_k in the simplified tax system. With positive marginal tax rates on income it turns out that while the complex tax system entails full tax deduction, the simplified tax system is characterized by less than full deduction. The intuition for both results is due to the observation that deductions serve to decrease the individuals' efforts. In the complex tax system, with full tax deductions consumers' efforts are not distorted. Since any redistribution achieved through taxing income cannot be improved upon by taxing work-related expenditures, full deduction is optimal. The simplified system, however, taxes work-related expenditures for q and consumptive expenditures for s equally. Therefore, tax deductions aimed at increasing q also increase s . To offset the implied negative effects with respect to efforts and to distortions in the consumptive expenditures, less than full deduction is optimal.

3.1. Tax Deductions in the Complex System

Considering the complex tax system $t(y, s, q)$ we start by specifying the optimal control problem for (6). We use y, s and q as control variables and the individuals' utility levels v as state variable. Let $\tilde{c}(s, q, y, v, \theta)$ be implicitly defined by the solution of $v(\theta) := u(c, s) - e(y, q)h$. Then the Hamiltonian for the government's maximization problem can be written as (λ_1 and $\lambda_2(\theta)$ are the multipliers for the tax requirement and the incentive compatibility restriction (5), respectively)

$$\begin{aligned} \mathcal{H}^c &= [v + \lambda_1 [y - s - q - \tilde{c}(s, q, y, v, \theta)]] f - \lambda_2 e(y, q)h' & (7) \\ \text{s.t. } R(\underline{\theta}) &= 0, R(\bar{\theta}) = T, v'(\theta) = -e(y, q)h' \\ \text{with } R(\theta) &:= \int_{\underline{\theta}}^{\theta} [y - s - q - \tilde{c}(s, q, y, v, \theta)] f d\tilde{\theta}. \end{aligned}$$

Evaluating the optimality conditions for (7) and letting $y^c(\theta), c^c(\theta), s^c(\theta)$ and $q^c(\theta)$ denote the optimal solutions, we get the following result.

Lemma 1 *The optimal $s^c(\theta)$ and $q^c(\theta)$ satisfy*

$$u_c(c^c(\theta), s) = u_s(c^c(\theta), s) \text{ and} \quad (8)$$

$$e_y(y^c(\theta), q) = -e_q(y^c(\theta), q). \quad (9)$$

Proof See Appendix A.1.

Using lemma 1 and employing (3) to characterize the marginal tax rates we get $t_s = 0$ and $t_q = -t_y$.⁴ Hence, we have

Proposition 1 *The optimal complex tax system $t^c(y, s, q)$ can be written as*

$$t^c(y, s, q) = t^c(y - q)$$

and implies full tax deduction.

Proof Substituting (8) and (9) in (3) leads to $t_s = 0$ and $t_y = -t_q$. ■

⁴Furthermore, it is easy to show that $0 < t_y < 1$ (see for example Mirrlees (1971) and (1976), Seade (1977) and (1982) and Ebert (1992)).

That t_s must be zero is of course well known from the literature (e.g., Atkinson and Stiglitz, 1976). Given the assumption of additively separable utility, an optimal income tax cannot be improved upon by excise taxes. The result that $t_y = -t_q$ is optimal is also in line with this literature, but we stress the implication for deductibility of work-related expenditures in order to have a benchmark for our simple tax system. The intuition for the result is simple. Since individual effort decisions are not distorted if $t_y = -t_q$, there are no further welfare gains from imposing $t_y \neq -t_q$. Any redistribution from taxing income cannot be improved upon by not fully exempting work-related expenditures.

3.2. Tax Deductions in the Simplified System

Turning to the simplified tax system, the tax function $t(y, k)$ with $k = s + q$ does not allow the government to influence q and s separately. Instead, the allocation of k between q and s is determined by the individuals according to their optimal private decisions, i.e. $u_s = -e_q h$ (see (4)). To incorporate this observation into the government's problem to design an optimal tax function we first define

$$\bar{s}(c, k, y, \theta) := \arg \max_s [u(c, s) - e(y, k - s)h].$$

Let $\bar{c}(k, y, v, \theta)$ be implicitly determined by the solution of $v(\theta) := u(c, \bar{s}(c, k, y, \theta)) - e(y, k - \bar{s}(c, k, y, \theta))h$. Choosing y and k as control variables and v as state variable, the Hamiltonian for the government's maximization problem can be written as (again, λ_1 and $\lambda_2(\theta)$ are the multipliers for the tax requirement and the incentive compatibility restriction)

$$\mathcal{H}^s = [v + \lambda_1 [y - k - \bar{c}(k, y, v, \theta)]] f - \lambda_2 e(y, k - \bar{s}(k, y, v, \theta))h' \quad (10)$$

$$\text{s.t. } R(\underline{\theta}) = 0, R(\bar{\theta}) = T, v'(\theta) = -e(y, k - \bar{s}(k, y, v, \theta))h'$$

$$\text{with } R(\theta) := \int_{\underline{\theta}}^{\theta} [y - \bar{c}(k, y, v, \theta) - k] f d\tilde{\theta}$$

$$\text{and } \bar{s}(k, y, v, \theta) := \tilde{s}(\bar{c}(k, y, v, \theta), k, y, \theta).$$

Evaluating the optimality conditions for (10) and letting $y^s(\theta)$ and $k^s(\theta)$ denote the solutions we get:

Lemma 2 *The optimal $y^s(\theta)$ and $k^s(\theta)$ satisfy*

$$e_y(y^s, k^s - \bar{s}(k^s, y^s, v^s, \theta)) = \Phi e_q(y^s, k^s - \bar{s}(k^s, y^s, v^s, \theta)) \quad (11)$$

$$\text{with } \Phi := \frac{(u_c + e_q h)e_{yq} h + u_c u_{ss} + u_{cs} e_q h}{(u_c + e_q h)e_{qq} h - u_c u_{ss} - u_{cs} e_q h}.$$

(i) With $u_c - e_y h > 0$, we have $-1 < \Phi < 0$.

(ii) With $u_{cs} \geq 0$, we have $u_c - e_y h \geq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$ and $u_c - e_y h = 0 \Rightarrow e_y = -e_q$.

Proof See Appendix A.2.

Using lemma 2 and (4) to characterize the marginal tax rates t_y and t_k implies the next proposition.⁵

Proposition 2 (i) With positive marginal tax rates the optimal simplified tax system $t^s(y, k)$ is characterized by less than full tax deduction:

$$t_y^s + t_k^s > 0 \text{ for } t_y^s > 0.$$

(ii) For $u_{cs} \geq 0$ the optimal tax rates on income are positive, i.e., $t_y^s \geq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$, and $t_y = 0$ implies $t_k = 0$.

Proof (i) Using (4) we get $u_c - e_y h > 0 \Leftrightarrow t_y > 0$ and $\text{sign}[t_y + t_k] > 0$ for $e_y < -e_q$. (ii) Obviously, $u_c = e_y h \Leftrightarrow t_y = 0$ and $e_y = -e_q \Leftrightarrow t_y = -t_k$. Therefore, lemma 2 implies the proposition. ■

The first part of the Proposition shows that less than full deduction is optimal if marginal taxes on income are positive. Since work-related and consumptive expenditures cannot be distinguished perfectly, any attempt to decrease e by increasing k will also increase the individual's choice of s .⁶ Furthermore, comparing $u_s = -e_q h$ with the condition that characterizes the optimal structure of consumptive expenditures, i.e., $u_c = u_s$,⁷ shows that with $t_y > 0$ full tax deductions would lead to a distorted consumption structure. Therefore, with positive marginal taxes on income full tax deductions cannot be optimal.

The second part of Proposition 2 establishes that the optimal marginal tax rates on income are positive as long as $u_{cs} \geq 0$ holds. Intuitively, with $u_{cs} \geq 0$, negative (marginal) tax rates on income would induce an increase in y and c which in turn provides a strong incentive to increase s at the expense of q . Hence, negative marginal income taxes would

⁵Generally, the marginal tax rates along the optimal expansion path satisfy $-1 < t_y, t_k < 1$. Furthermore, the marginal tax rates for $\underline{\theta}, \bar{\theta}$ are zero.

⁶Using $u_s(\cdot) + e_q(\cdot)h(\theta) = 0$ and $e_{yq} > -e_{qq}$ simple comparative statics shows that an increase in k leads to an increase of s .

⁷Noting that $t_s = 0$ in the optimal complex tax system.

imply rather strong distortions with respect to the choice of y and q . On the other hand, taxing income positively (at the margin) and allowing for tax deductions reduces y and c and thus leads to relatively lower distortions with respect to y and q . Note, however, that this also implies that income which is finally used for work-related expenditures is taxed positively.

4. Progression in a Simple Example

The last observation together with the incentive compatibility condition (5) leads to the conjecture that the simplified tax system should imply less redistribution. The incentive compatibility constraints show that the higher the individuals' effort levels the more their utilities must increase with their ability to work and the lower the tax induced redistribution should be.

In the following we will explore this conjecture by comparing the optimal complex and simplified tax system. Using further separability assumptions it turns out that the simplified system implies that low income earners have to pay higher taxes while the average tax payments of high income earners decrease. Since the simplified system also induces higher average income, it can be characterized as less progressive and less redistributive than the complex tax system.

We assume that the utility from consumption $u(c, s)$ is additively separable in c and s and the effort function e linear in y and additively separable in y and q :

$$u_{cs} = e_{yy} = e_{yq} = 0. \quad (12)$$

While (12) is quite restrictive, it allows us to compare the allocations induced by the two systems $t^c(y - q)$ and $t^s(y, k)$ directly. Using the optimality conditions for (7) and (10) and defining $\Delta z := z^s(\theta) - z^c(\theta)$ for $z = y, c, s,$ and q , we get:

Lemma 3 *With (12), the optimal tax systems $t^c(y - q)$ and $t^s(y, k)$ imply*

$$(i) \Delta c = 0 \text{ and } t_y^c(y^c(\theta) - q^c(\theta)) = t_y^s(y^s(\theta), k^s(\theta)) \forall \theta \in [\underline{\theta}, \bar{\theta}]$$

$$(ii) \Delta s > 0 > \Delta q; \Delta q + \Delta s > 0 \text{ and}$$

$$e(c^s + k^s, k^s - s^s) > e(c^c + s^c + q^c, q^c) \quad (13)$$

$$u(c^s, s^s) - e(c^s + k^s, k^s - s^s)h(\theta) > u(c^c, s^c) - e(c^c + s^c + q^c, q^c)h(\theta) \quad (14)$$

for all $\theta \in (\underline{\theta}, \bar{\theta})$.

Proof See Appendix A.3.

The first part of the Lemma shows that the optimal marginal distortion between c and income does not vary under the complex and the simplified tax system. As marginal utility u_c and marginal effort e_y are not affected by s and q , the two tax systems induce the same distortions with respect to c and y . While the different tax systems may imply different income levels $y^c(\theta)$ and $y^s(\theta)$, the optimal income decisions are governed by the same marginal tax rates.

The second part of the lemma quantifies the distortions with respect to s and q . Note that these distortions imply that the individuals' efforts as well as their utility levels increase with the simplified system if tax payments are disregarded.

However, we know that welfare must be higher under the complex tax regime, that is ($v^c(\theta)$ and $v^s(\theta)$ denote the individuals' utilities with the optimal complex and the optimal simplified tax system):

$$\int_{\underline{\theta}}^{\bar{\theta}} v^s(\theta) f(\theta) d\theta < \int_{\underline{\theta}}^{\bar{\theta}} v^c(\theta) f(\theta) d\theta. \quad (15)$$

Therefore, combining (13) and (15), defining $\Delta t(\theta) := t^c(y^c(\theta) - q^c(\theta)) - t^s(y^s(\theta), k^s(\theta))$ and using $\int_{\underline{\theta}}^{\bar{\theta}} \Delta t(\theta) f(\theta) d\theta = 0$ and again (12), the weighted sum of aggregate tax payments is higher with the simplified tax system, i.e.,

$$\int_{\underline{\theta}}^{\bar{\theta}} \Delta t(\theta) h(\theta) f(\theta) d\theta > 0. \quad (16)$$

Finally, employing the incentive compatibility condition (5) we get:

Proposition 3 *Under the assumptions in (12), the simplified tax system is less progressive than the complex tax system in the sense that*

(i) *the simplified tax system induces higher aggregate income,*

(ii) *individuals with low incomes pay strictly higher taxes and are strictly worse off compared to the complex tax system.*

Proof See Appendix A.4.

Proposition 2 confirms the conjecture that the simplified system is less redistributive and less progressive than the complex system. Redistribution is ultimately targeted at decreasing the differences between the individuals' utility levels, but the Proposition shows that the difference between the utility levels at the lower and upper end of the ability distribution is higher with the simplified system. With respect to progressivity, there exists an interval of low ability individuals whose tax payments increase with the simplified system. Hence, the average tax payment of the individuals with higher ability levels and thus higher incomes must be lower than with the complex system.

5. Conclusion

The paper has studied the interaction of the rate schedule and tax complexity in an optimal income tax model. Complexity was defined with respect to the informational requirements of the tax system: the more information authorities require from taxpayers, the more complex the tax system. We obtained two basic results from our model: First, tax deductions are optimal even if authorities cannot perfectly distinguish between work-related and consumptive expenditures. Second, the example showed that the optimal simple system may be characterized by higher tax burdens on the poor and less redistribution.

In the 1980s, many countries reformed their tax systems towards less progressivity and less complexity. Of course, this may be primarily due to the distributional objectives of the conservative parties that came to power at this time. Our model has shown that this history could also be rationalized in an optimal taxation framework. Our analysis implies that an increase in the costs of administering a complex tax system and a corresponding simplification of the tax rules might imply less progressive tax rates even if the government's distributional objectives remain the same.

Acknowledgments

We thank the editor John Wilson and two referees for very helpful comments and suggestions.

Appendix

A.1 Proof of Lemma 1

The optimality conditions for (7) with respect to y, s and q are

$$\mathcal{H}_y^c = \lambda_1 [1 - \tilde{c}_y] f - \lambda_2 e_y h' = 0, \quad (\text{A.1})$$

$$\mathcal{H}_s^c = \lambda_1 [-1 - \tilde{c}_s] f = 0, \quad \mathcal{H}_q^c = \lambda_1 [-1 - \tilde{c}_q] f - \lambda_2 e_q h = 0. \quad (\text{A.2})$$

In addition we have the multiplier equations, the state equations and the boundary conditions:

$$0 = [1 - \lambda_1 \tilde{c}_v] f(\theta) + \lambda_2'(\theta), \quad (\text{A.3})$$

$$0 = v'(\theta) + e(y, q) h', \quad R'(\theta) = [y - s - q - \tilde{c}(s, q, y, v, \theta)] f(\theta), \quad (\text{A.4})$$

$$0 = \lambda_2(\underline{\theta}) = \lambda_2(\bar{\theta}), \quad R(\bar{\theta}) = T. \quad (\text{A.5})$$

Simple comparative statics for $\tilde{c}(y, s, q, v, \theta)$ yields

$$\tilde{c}_y = \frac{e_y h}{u_c}, \quad \tilde{c}_q = \frac{e_q h}{u_c}, \quad \tilde{c}_s = -\frac{u_s}{u_c}, \quad \tilde{c}_v = \frac{1}{u_c}. \quad (\text{A.6})$$

Substituting (A.6) in (A.2) and solving for u_s and e_q leads to

$$e_q = -e_y \text{ and } u_s = u_c. \quad (\text{A.7})$$

A.2 Proof of Lemma 2

In order to prove the Lemma we first calculate the optimality conditions for (10). We then prove part (i) and part (ii).

The optimality conditions for (10) with respect to y and k are

$$\mathcal{H}_y^s = \lambda_1 [1 - \bar{c}_y] f - \lambda_2 (e_y + e_q (-\bar{s}_y)) h' = 0, \quad (\text{A.8})$$

$$\mathcal{H}_k^s = \lambda_1 [-1 - \bar{c}_k] f - \lambda_2 e_q (1 - \bar{s}_k) h' = 0. \quad (\text{A.9})$$

In addition we have the multiplier equations, the state equations and the boundary conditions:

$$0 = [1 - \lambda_1 \bar{c}_v] f(\theta) - \lambda_2(\theta) e_q (-\bar{s}_v) h'(\theta) + \lambda_2'(\theta), \quad (\text{A.10})$$

$$0 = v'(\theta) + e(y, k - \bar{s}(k, y, v, \theta)) h', \quad R'(\theta) = [y - \bar{c}(k, y, v, \theta) - k] f(\theta), \quad (\text{A.11})$$

$$0 = \lambda_2(\underline{\theta}) = \lambda_2(\bar{\theta}), \quad R(\bar{\theta}) = T. \quad (\text{A.12})$$

Comparative statics of $\bar{c}(k, y, v, \theta)$ with respect to y, k and v gives

$$\bar{c}_y = \frac{e_y h}{u_c}, \quad \bar{s}_y = \frac{h(e_y u_{cs} + u_c e_{yq})}{u_c(e_{qq}h - u_{ss})}, \quad (\text{A.13})$$

$$\bar{c}_k = \frac{e_q h}{u_c}, \quad \bar{s}_k = \frac{h(e_q u_{cs} + u_c e_{qq})}{u_c(e_{qq}h - u_{ss})}, \quad \bar{c}_v = \frac{1}{u_c}, \quad \bar{s}_v = \frac{u_{cs}}{u_c(e_{qq}h - u_{ss})}. \quad (\text{A.14})$$

Substituting (A.13) and (A.14) in (A.8) and (A.9) and rearranging yields

$$u_c - e_y h = \frac{\lambda_2 h'}{\lambda_1 f} \frac{e_q h(e_y u_{cs} + e_{yq} u_c) - u_c e_y (e_{qq} h - u_{ss})}{u_{ss} - e_{qq} h}, \quad (\text{A.15})$$

$$u_c + e_q h = \frac{\lambda_2 h'}{\lambda_1 f} \frac{e_q (u_c u_{ss} + e_q h u_{cs})}{e_{qq} h - u_{ss}}. \quad (\text{A.16})$$

Solving (A.15) and (A.16) implies

$$e_y = e_q \Phi \quad \text{with} \quad \Phi := \frac{(u_c + e_q h) e_{yq} h + u_c u_{ss} + u_{cs} e_q h}{(u_c + e_q h) e_{qq} h - u_c u_{ss} - u_{cs} e_q h} \quad (\text{A.17})$$

which confirms (11).

To prove part (i) of the lemma note first that $e_y > 0 > e_q$ leads to $\Phi < 0$. To show that $u_c - e_y h > 0$ implies $\Phi > -1$ assume to the contrary that $\Phi \leq -1$ holds. Then, using (A.17) we have to consider two cases:

$$(i) \quad (u_c + e_q h) e_{yq} h + u_c u_{ss} + u_{cs} e_q h > 0 > (u_c + e_q h) e_{qq} h - u_c u_{ss} - u_{cs} e_q h$$

$$\text{and} \quad (u_c + e_q h) (e_{qq} + e_{yq}) \geq 0,$$

$$(ii) \quad (u_c + e_q h) e_{yq} h + u_c u_{ss} + u_{cs} e_q h < 0 < (u_c + e_q h) e_{qq} h - u_c u_{ss} - u_{cs} e_q h$$

$$\text{and} \quad (u_c + e_q h) (e_{qq} + e_{yq}) \leq 0.$$

Case (i) requires $u_c + e_q h \geq 0$. Using $u_{cs} > u_{ss}$ this leads to $(u_c + e_q h) e_{qq} h - u_c u_{ss} - u_{cs} e_q h > 0$ and thus to a contradiction with $0 > (u_c + e_q h) e_{qq} h - u_c u_{ss} - u_{cs} e_q h$. In case (ii) we have $u_c + e_q h \leq 0$. But since $e_y \geq -e_q$ and $u_c - e_y h > 0$ lead to $u_c + e_q h > 0$ we again have a contradiction. Hence, $u_c - e_y h > 0 \Rightarrow 0 > \Phi > -1$.

The proof of part (ii) is again based on a contradiction. We first show that $u_c - e_y h \leq 0$ implies $u_c + e_q h \leq 0$. Using $u_{cs} \geq 0$ we get that $u_c + e_q h < 0 \Rightarrow \lambda_2 > 0$. However, solving the multiplier equation (A.10) incentive compatibility requires that $\lambda_2(\theta) \leq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$. Hence, we must also have $u_c - e_y h \geq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$ and $u_c - e_y h = 0 \Rightarrow u_c + e_q h = 0$.

Assuming $u_c - e_y h < 0$, substituting e_y (see (A.17)) on the right hand sides of (A.15) and (A.16) and dividing these equations reveals

$$\text{sign}(u_c + e_q h) = \text{sign} \left[\frac{e_q h (e_{yq} h + u_{cs}) - u_c (e_{qq} h - u_{ss})}{e_q h (e_{qq} h - u_{cs}) + u_c (e_{qq} h - u_{ss})} \right]. \quad (\text{A.18})$$

Now, assume to the contrary that $u_c + e_q h > 0$. Then we have to consider two cases:

$$(i) \frac{e_q h}{u_c}(e_{yq} h + u_{cs}) > e_{qq} h - u_{ss} \text{ and } \frac{e_q h}{u_c}(e_{qq} h - u_{cs}) > -(e_{qq} h - u_{ss}),$$

$$(ii) \frac{e_q h}{u_c}(e_{yq} h + u_{cs}) < e_{qq} h - u_{ss} \text{ and } \frac{e_q h}{u_c}(e_{qq} h - u_{cs}) < -(e_{qq} h - u_{ss}).$$

Since $u_c + e_q h > 0 \Leftrightarrow e_q h/u_c > -1$, $u_{cs} > u_{ss}$ and $e_{yq} > -e_{qq}$ imply that in case (i) the first inequality and in case (ii) the second inequality is violated, we must have $u_c - e_y h < 0 \Rightarrow u_c + e_q h < 0$.

Considering $u_c - e_y h = 0$, note that (A.15) either implies $\lambda_2 = 0$ and thus $u_c = -e_q h$ (see (A.16)) or it leads to

$$e_y = -e_q \frac{u_{cs} + e_{yq} h}{u_{ss} - e_{qq}}. \quad (\text{A.19})$$

Since $e_y > 0 > e_q$ requires $u_{cs} + e_{yq} h < 0$ for (A.19) to hold, using $u_{cs} > u_{ss}$ and $e_{yq} h > -e_{qq}$ shows that (A.19) leads to $e_y < -e_q$ and therefore to $u_c + e_q h < 0$.

Turning to the sign of λ_2 and using $u_c u_{ss} + e_q h u_{cs} < 0$ (by the assumption that $u_{cs} \geq 0$), $e_{qq} h - u_{ss} > 0$ and $e_q h' > 0$, inspection of (A.16) reveals $\text{sign}(u_c + e_q h) = -\text{sign} \lambda_2/\lambda_1$. Solving (A.10) for $\lambda_2(\theta)$, using the boundary conditions (A.12) and substituting \bar{c}_v (see (A.14)) we get

$$\lambda_2(\theta) = - \int_{\underline{\theta}}^{\theta} \left(1 - \frac{\lambda_1}{u_c}\right) f(\tilde{\theta}) E \int_{\underline{\theta}}^{\tilde{\theta}} \Gamma(\hat{\theta}) d\hat{\theta} d\tilde{\theta} \text{ with } \Gamma(\hat{\theta}) := e_q \bar{s}_v h'. \quad (\text{A.20})$$

Since $\lambda_2(\bar{\theta}) = 0$ we must have $\lambda_1 > 0$ and thus $\text{sign}(u_c + e_q h) = -\text{sign} \lambda_2$. Hence, $u_c - e_y h < 0 \Rightarrow u_c + e_q h < 0 \Rightarrow \lambda_2 > 0$.

We complete the proof by showing that incentive compatibility requires $\lambda_2(\theta) \leq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$.⁸ Using (A.20) and evaluating all functions at the optimal solutions $y^s(\theta)$, $k^s(\theta)$ we get

$$\lambda_2^s(\theta) = - \left(1 - \frac{\lambda_1}{u_c}\right) f(\theta) - \lambda_2^s(\theta) e_q \bar{s}_v h'(\theta). \quad (\text{A.21})$$

Now, assume to the contrary that there exists an interval $[\underline{\theta}, \theta_1]$ with $\underline{\theta} < \theta_1 < \bar{\theta}$ such that $\lambda_2^s(\theta_1) = 0$, $\lambda_2^s(\theta) > 0 \forall \theta \in (\underline{\theta}, \theta_1)$ and $\lambda_2^s(\theta) \leq 0$ for all $\theta \in (\theta_1, \theta_1 + \varepsilon)$ with $\varepsilon > 0$

⁸This part of the proof is analogous to Seade (1982).

but small enough. These assumptions imply $\lambda_2^s '(\underline{\theta}) \geq 0 \geq \lambda_2^s '(\theta_1)$ which by using (A.21), $\lambda_2^s(\underline{\theta}) = \lambda_2^s(\theta_1) = 0$ and (A.15), (A.16) also leads to

$$\lambda_1 \geq u_c|_{\theta=\underline{\theta}} = u_s|_{\theta=\underline{\theta}} = e_y|_{\theta=\underline{\theta}} h(\underline{\theta}) = -e_q|_{\theta=\underline{\theta}} h(\underline{\theta}) \quad \text{and} \quad (\text{A.22})$$

$$\lambda_1 \leq u_c|_{\theta=\theta_1} = u_s|_{\theta=\theta_1} = e_y|_{\theta=\theta_1} h(\theta_1) = -e_q|_{\theta=\theta_1} h(\theta_1). \quad (\text{A.23})$$

(A.22), (A.23) and $h(\underline{\theta}) > h(\theta_1)$ obviously imply $u_c|_{\theta=\underline{\theta}} \leq u_c|_{\theta=\theta_1}$ and $e_y|_{\theta=\underline{\theta}} > e_y|_{\theta=\theta_1}$. From strict concavity of $u(c, s)$ and strict convexity of $e(y, q)$ we get $u_c|_{\theta=\underline{\theta}} \leq u_c|_{\theta=\theta_1} \Rightarrow c^s|_{\theta=\underline{\theta}} \geq c^s|_{\theta=\theta_1}$ and $e_y|_{\theta=\underline{\theta}} > e_y|_{\theta=\theta_1} \Rightarrow y^s|_{\theta=\underline{\theta}} < y^s|_{\theta=\theta_1}$. Define $\hat{s}(c)$ and $\hat{q}(y)$ such that

$$u_c(c, \hat{s}(c)) \equiv u_s(c, \hat{s}(c)) \quad \text{and} \quad e_y(y, \hat{q}(y)) + e_q(y, \hat{q}(y)) \equiv 0,$$

and note that $\hat{s}(c) = \bar{s}$ and $\hat{q}(y) = k^s - \bar{s}$ for $\theta = \underline{\theta}$ and for $\theta = \theta_1$. Simple comparative statics reveals

$$\frac{d^2 u(c, \hat{s}(c))}{dc^2} = \frac{u_{cc}u_{ss} - u_{cs}^2}{u_{ss} - u_{cs}} < 0 \quad \text{and} \quad \frac{d^2 e(y, \hat{q}(y))}{dy^2} = \frac{e_{yy}e_{qq} - e_{yq}^2}{e_{yq} + e_{qq}} > 0$$

where the sign conditions follow from $u_{cs} > u_{ss}$ and strict concavity of $u(c, s)$ and from $e_{yq} > -e_{qq}$ and strict convexity of $e(y, q)$. Hence, (A.22) and (A.23) imply $c^s|_{\underline{\theta}} \geq c^s|_{\theta_1}$ and $y^s|_{\underline{\theta}} < y^s|_{\theta_1}$. But then the θ_1 consumer prefers $(y^s(\underline{\theta}), k^s(\underline{\theta}))$ to $(y^s(\theta_1), k^s(\theta_1))$ which contradicts incentive compatibility. Repeating this argument shows that with any interval $[\theta_1, \theta_2]$ with $\underline{\theta} \leq \theta_1 < \theta_2 \leq \bar{\theta}$ such that $\lambda_2^s(\theta_1) = \lambda_2^s(\theta_2) = 0$ and $\lambda_2^s(\theta) > 0 \forall \theta \in (\theta_1, \theta_2)$ incentive compatibility would be violated. Therefore, we must have $\lambda_2(\theta) \leq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$ which contradicts $u_c - e_y h < 0$ for any $\theta \in [\underline{\theta}, \bar{\theta}]$.

A.3 Proof of Lemma 3

To show $c^c(\theta) = c^s(\theta)$ we first use lemma 1 and (A.3)–(A.5) and (A.20), respectively, to calculate $\lambda_2^c(\theta)$ and $\lambda_2^s(\theta)$. Defining $\bar{h} := \int_{\underline{\theta}}^{\theta} h(\tilde{\theta}) f(\tilde{\theta}) d\tilde{\theta}$ and using (12) we obtain

$$\lambda_2^c(\theta) = \lambda_2^s(\theta) = \frac{1}{h(\theta)} \int_{\underline{\theta}}^{\theta} [\bar{h} - h(\tilde{\theta})] f(\tilde{\theta}) d\tilde{\theta} < 0 \quad \text{and} \quad \lambda_1^c = \lambda_1^s = e_y \bar{h} > 0, \quad (\text{A.24})$$

which by using (A.1) and (A.8) also reveals $c^c(\theta) = c^s(\theta)$. Furthermore, employing (3) and (4) shows that the marginal tax rates on income are the same under both tax systems, i.e.,

$$t_y^c(y^s(\theta), k^s(\theta)) = t_y^s(y^c(\theta) - q^c(\theta)) = \frac{e_y \lambda_2^c(\theta) h'(\theta)}{\lambda_1^c f(\theta)}.$$

To prove part (ii) we can combine (3) and (4) such that with $\phi \in [0, 1]$

$$u_s = \phi u_c - (1 - \phi)e_q h, \quad e_q = -e_y(\phi + (1 - \phi)x) \quad (\text{A.25})$$

$$\text{with } x := \frac{u_c(e_{qq}h - u_{ss})}{e_{qq}e_y h^2 - u_c u_{ss}},$$

where the definition of x follows from lemma 1 and lemma 2. Solving $u_{ss}s_\phi = e_y[u_c + x]h$ and $e_{qq}q_\phi = -e_y(1 - x)$ for q_ϕ and s_ϕ leads with $u_c > e_y h \Leftrightarrow t_y > 0$ to

$$q_\phi = \frac{e_y(x - 1)}{e_{qq}} > 0, \quad s_\phi = \frac{e_y h(u_c + x)}{u_{ss}} < 0 \text{ and } s_\phi + q_\phi < 0. \quad (\text{A.26})$$

Substituting s_ϕ and q_ϕ in (A.25) and using $u_c > e_y h$ we obtain

$$0 > u_s s_\phi - (e_y(q_\phi + s_\phi) + e_q q_\phi)h, \quad (\text{A.27})$$

$$0 > e_y(q_\phi + s_\phi) + e_q q_\phi h. \quad (\text{A.28})$$

Inequalities (A.26) and (A.27)–(A.28) confirm (13) and (14).

A.4 Proof of Proposition 3

Part (i) is implied by lemma 3 and $\int_{\underline{\theta}}^{\bar{\theta}} \Delta t(\theta) f(\theta) d\theta = 0$. To prove part (ii) note first that (16) implies that there exists a $\theta_{\Delta t}$ defined by

$$\theta_{\Delta t} := \min \{ \theta \mid \Delta t(\theta) = 0, \theta \in [\underline{\theta}, \bar{\theta}] \} \in (\underline{\theta}, \bar{\theta}).$$

Using $\theta_{\Delta t}$ the following two inequalities which are implied by incentive compatibility and lemma 3

$$\Delta t(\theta) > 0 \Rightarrow \frac{d}{d\theta} v^s(\theta) > \frac{d}{d\theta} v^c(\theta), \quad (\text{A.29})$$

$$\Delta t(\theta) \leq 0 \Rightarrow v^s(\theta) > v^c(\theta) \quad (\text{A.30})$$

lead to $\Delta t(\theta) > 0$ for all $\theta < \theta_{\Delta t}$. Assume to the contrary that $\Delta t(\theta) \leq 0$ for $\theta \in [\underline{\theta}, \theta_{\Delta t}]$. Since both tax systems are characterized by zero marginal tax rates at the lowest income levels lemma 3 then implies

$$v^s(\theta) \geq v^c(\theta) \text{ for } \theta \in [\underline{\theta}, \theta_{\Delta t}]$$

with strict inequality for $\underline{\theta} < \theta \leq \theta_{\Delta t}$. Now assume that there exists a θ_1 with $\theta_{\Delta t} < \theta_1 \leq \bar{\theta}$ such that $\Delta t(\theta) > 0$ for $\theta \in (\theta_{\Delta t}, \theta_1]$. Then (A.29) implies that

$$\frac{d}{d\theta} v^s(\theta) > \frac{d}{d\theta} v^c(\theta) \text{ for all } \theta \in (\theta_{\Delta t}, \theta_1].$$

Hence we have $v^s(\theta) > v^c(\theta)$ for all $\theta \in (\theta_{\Delta t}, \theta_1]$. Alternatively, assume that there exists a θ_2 with $\theta_{\Delta t} < \theta_2 \leq \bar{\theta}$ such that $\Delta t(\theta) < 0$ for $\theta \in (\theta_{\Delta t}, \theta_2]$. Then, (A.30) reveals $v^s(\theta) > v^c(\theta)$ for all $\theta \in (\theta_{\Delta t}, \theta_2]$. Repeating these two arguments for the entire range $[\theta_{\Delta t}, \bar{\theta}]$ and all possible intervals with either $\Delta t(\theta) > 0$ or $\Delta t(\theta) \leq 0$ shows that $\Delta t(\theta) \leq 0$ for $\theta \in [\underline{\theta}, \theta_{\Delta t}]$ would imply

$$\int_{\underline{\theta}}^{\bar{\theta}} v^s(\theta) f(\theta) d\theta > \int_{\underline{\theta}}^{\bar{\theta}} v^c(\theta) f(\theta) d\theta \quad (\text{A.31})$$

which contradicts (15). Hence, $\Delta t(\theta) > 0$ for all $\theta < \theta_{\Delta t}$.

Finally, applying the same reasoning and using $v^s(\theta_{\Delta t}) > v^c(\theta_{\Delta t})$, (A.29) and (A.30) imply that there exists a $\theta_{\Delta v} \in (\underline{\theta}, \theta_{\Delta t})$ such that $v^s(\theta) < v^c(\theta)$ for all $\theta < \theta_{\Delta v}$ and $v^s(\theta) \geq v^c(\theta)$ for all $\theta \geq \theta_{\Delta v}$.

References

- Atkinson, Anthony B. (1995). *Public Economics in Action. The Basic Income/Flat Tax Proposal*. Oxford: Clarendon Press.
- Atkinson, Anthony B. and Joseph E. Stiglitz. (1976). “The design of tax structure: Direct versus indirect taxation.” *Journal of Public Economics* 6, 55–75.
- Cooter, Robert. (1978). “Optimal tax schedules and rates: Mirrlees and Ramsey.” *American Economic Review* 68, 756–768.
- Ebert, Udo. (1992). “A reexamination of the optimal nonlinear income tax.” *Journal of Public Economics* 49, 47–73.
- Kaplow, Louis. (1990). “Optimal taxation with costly enforcement and evasion.” *Journal of Public Economics* 43, 221–236.
- Mirrlees, James A. (1971). “An exploration in the theory of optimum income taxation.” *Review of Economic Studies* 38, 175–208.
- Mirrlees, James A. (1976). “Optimal tax theory: A synthesis.” *Journal of Public Economics* 6, 327–358.
- Seade, Jesus K. (1977). “On the shape of optimal tax schedules.” *Journal of Public Economics* 7, 203–236.

Seade, Jesus K. (1982). "On the sign of the optimum marginal income tax." *Review of Economic Studies* 49, 637–643.

Slemrod, Joel. (1994). "Fixing the leak in Okun's bucket. Optimal tax progressivity when avoidance can be controlled." *Journal of Public Economics* 55, 41–51