J Econ (2013) 110:107–124
DOI 10.1007/s00712-012-0314-4

Optimal taxes on wealth and consumption in the presence of tax evasion

Johann K. Brunner · Paul Eckerstorfer · Susanne Pech

Received: 10 April 2012 / Accepted: 5 October 2012 / Published online: 26 October 2012

Abstract This article incorporates tax evasion into an optimum taxation framework with individuals differing in earning abilities and initial wealth. We find that despite the possibility of its evasion a tax on initial wealth should supplement the optimal nonlinear income tax, given a positive correlation between initial wealth and earning abilities. Further, even if income and initial wealth are taxed optimally, it is still desirable to levy a tax on commodities, though it can be evaded as well. Thus, our result provides a rationale for a comprehensive tax system. Optimal tax rates on commodities differ in general, however for the special case of a uniform evasion technology equal rates are optimal if preferences are homothetic and weakly separable.

Keywords Optimal taxation · Tax evasion · Initial wealth

JEL Classification D82 · H21 · H24 · H26

1 Introduction

Most countries rely on a comprehensive tax system and impose, in addition to an income tax, also taxes on commodities and on wealth or wealth transfers. This is in contrast to the well-known result by Atkinson and Stiglitz (1976) which tells us that, essentially, an income tax alone is a sufficient instrument for raising and redistributing funds, given that individuals differ in abilities to earn income. The imposition of other taxes can therefore only be explained in a model with additional sources of heterogeneity of individuals.

J. K. Brunner (✉) · P. Eckerstorfer · S. Pech
Department of Economics, University of Linz, Altenbergerstr. 69, 4040 Linz, Austria
e-mail: johann.brunner@jku.at

An obvious idea is to extend the standard optimal-taxation model by considering differences in initial wealth.[1] Indeed, there exist some papers which do pay attention to the fact that initial wealth creates a second distinguishing characteristic, in addition to earning abilities. However, these papers are based on rather strong assumptions concerning the observability of initial wealth. In particular, Cremer et al. (2001) consider a static economy with exogenously given initial wealth, which is completely unobservable to the social planner. They find that commodity taxation is a useful instrument in such an economy, in addition to an optimal nonlinear income tax, as it allows implicit taxation of unobservable initial wealth.[2] In contrast to this, Brunner and Pech (2012a,b) assume that inherited wealth is observable to the social planner. They analyze the optimal taxation of inheritances and show that a redistributive motive for an inheritance tax arises, given that initial wealth increases with earning abilities. Moreover, this tax is equivalent to a uniform expenditure tax in their model.

In the present article we also study an extended optimal-taxation model, with initial wealth as a second distinguishing characteristic of individuals in addition to heterogeneous ability levels. But we drop the extreme assumptions of complete observability and unobservability, respectively. In our model the social planner is aware that she can tax wealth only to the extent which is reported to tax authorities, but not the true size of initial wealth. This reflects the fact that when designing the tax system, governments have to take into account the possibility that individuals and firms try to find legal or even illegal ways in order to escape (part of) the tax liability. For instance, in the current discussion about wealth taxation, it is often claimed that financial wealth can be concealed from tax authorities by moving assets offshore or by simply not reporting the true amount of wealth.[3] Then most notably the rich, who are intended to bear the major burden, given that wealth is distributed unequally in most societies, do not contribute as a much as one might expect. Another important issue in tax policy is the evasion of indirect taxes, such as the value added tax in the European Union.[4] For example, firms can hide part of their tax base by understating sales to tax authorities. But this does not imply complete unobservability, because hiding the tax base comes at some costs and the degree to which these are borne depends on the size of the tax burden.

Accordingly, we formulate an optimal-taxation model which accounts for the possibility that taxes on initial wealth and on commodities are subject to tax evasion. This framework allows us to study the trade-off between redistribution and the efficiency losses due to tax evasion, and to analyze the question of whether taxes on wealth and on the consumption of goods are adequate instruments in an optimal tax system, in addition to a tax on labor income.

1 The relevance of this extension is underscored by the fact that wealth differences between individuals are significant and have risen steadily over the last decades in many societies, see for example Atkinson et al. (2011).

2 Similarly, Boadway et al. (2000) and Cremer et al. (2003) demonstrate in a dynamic model the desirability of a tax on capital income as a surrogate for the taxation of inheritances, which again are assumed to be unobservable. They argue that in an open economy initial inherited wealth is hardly observable due to the high mobility of capital.

3 The existence of tax havens represents a particular opportunity for this, as is revealed by the recent debate on information transmission from foreign banks to the tax authorities in their customers' home countries.

4 For a recent overview on evasion of the VAT see Keen and Smith (2006).

The problem of tax evasion for the design of an optimal tax system has already been recognized by some previous contributions to the literature (e.g. Cremer and Gahvari 1993, 1995; Boadway et al. 1994; Pestieau et al. 2004). However, none of these studies has taken up the issue of wealth tax evasion, nor—with the exception of Cremer and Gahvari (1993)—the issue of commodity tax evasion. So far the focus has been laid on the evasion of income taxes and its implication for the optimal tax structure. For example, Boadway et al. (1994) analyze the optimal tax problem in a Mirrlees model, if individuals can evade income taxes but not commodity taxes. They show that in such a setting indirect taxes are a useful supplement to an optimal tax system and argue that this can explain the direct-indirect tax mix observed in most countries. Yet, one may object to this reasoning that at least for employees, who make up the majority of the tax payers, income taxes are rather difficult to evade as they are deducted by the employers in most countries. Arguably wealth and commodity taxes are more in danger to be evaded, in particular if one thinks of financial wealth, which is highly mobile, and of services, which are difficult to monitor.

We obtain three main results from our analysis. First, given a positive correlation between ability and wealth it is optimal to supplement the income tax with a tax on initial wealth, despite the possibility of its evasion. The reason is that if high-able individuals also own more initial wealth, taxing wealth allows for further redistribution which dominates the efficiency loss due to tax evasion for low tax rates. Second, we find that also a role for commodity taxes arises in our model even if they are exposed to tax evasion and even if wealth and income are already taxed optimally. Again, taxing commodities allows for further redistribution to what is possible through the income and the wealth tax. This redistributive effect dominates the efficiency losses due to tax evasion and distorted prices for low tax rates. Third, in our model commodities should in general be taxed at different rates. Uniform taxation of commodities is optimal if evasion costs for commodities are uniform and if preferences are homothetic for the consumption goods and weakly separable between labor supply and consumption.

The rest of the paper is organized as follows. Section 2 presents the model. Section 3 analyzes the optimal taxation of initial wealth and Sect. 4 deals with the taxation of commodities, including the analysis of the optimal structure of commodity taxes. Finally, Sect. 5 concludes.

2 The model

In the economy there exist two types of individuals $i = L, H$. The types differ in two characteristics, namely in earning abilities $\omega_L < \omega_H$ and exogenously given initial wealth $e_L < e_H$, that is, we assume a fixed relationship between abilities and initial wealth.[5] There is a large number of individuals of each type $i = L, H$; we normalize

[5] Indeed, in view of some empirical evidence that individuals with higher labor income also own more wealth (e.g., Diaz-Giménez et al. 2011) and that a substantial part of wealth is due to inheritances (Davies and Shorrocks 2000), a positive correlation between initial wealth and abilities appears quite plausible. It is also corroborated by the fact that both, the level of a person's education and her inheritances can be expected to increase with accumulated wealth of the parents.

the size of each group to one.[6] Individuals live for one period, in which they work and consume. By providing labor supply l_i they obtain a pre-tax income $z_i = \omega_i l_i$. It is subject to a nonlinear income tax and the resulting net income is denoted by x_i. Initial wealth is taxed at a proportional rate τ_e. Individuals use after-tax income x_i together with net initial wealth for the consumption c_{ji} of two commodities $j = 1, 2$, which are subject to per-unit taxes denoted by τ_j. Commodities are produced by a large number of firms in two industries, with perfect competition among the identical firms in each industry. Again the size (the set of firms) of each industry is normalized to one. We assume technologies to be linear, that is, the marginal product of labor is constant. Quantities are chosen in such a way that the (constant) marginal costs of production are equal to one for both commodities.

Individual behavior

Taxes on initial wealth can be evaded by individuals at some cost. In the modeling of tax evasion we follow the riskless approach first introduced by Usher (1986).[7] An individual of type i conceals a fraction α_{ei} of initial wealth from tax authorities at the expense of evasion costs depending on the fraction α_{ei}. Once individuals have incurred those costs they cannot be detected by tax authorities, thus there is no decision under uncertainty in the model. The cost of misreporting depends on the evaded amount in the following way: evading one unit of e_i entails a resource cost described by the function $K_{ei}(\alpha_{ei})$, depending on the evaded share α_{ei}.[8] Total evasion costs for individual i are then equal to $\alpha_{ei} e_i K_{ei}(\alpha_{ei})$. The function $K_{ei}(\alpha_{ei})$ is assumed to be an increasing and convex function of the evaded share α_{ei}, with $K_{ei}(0) = 0$ and $K_{ei}(1) < 1$. Observe that with these assumptions the total cost of evading all wealth $e_i K_{ei}(1)$ cannot exceed e_i. To simplify notation we define $k_{ei}(\alpha_{ei}) \equiv \alpha_{ei} K_{ei}(\alpha_{ei})$.[9]

Both types of individuals have the same strictly concave utility function, $u(c_{1i}, c_{2i}, l_i)$, with $\partial u/\partial c_{ji} > 0$, $\partial u/\partial l_i < 0$. For any given labor supply and net income individuals maximize their utility subject to the private budget constraint

$$p_1 c_{1i} + p_2 c_{2i} \leq x_i + R_{ei} e_i \tag{1}$$

with consumer prices denoted by p_j, $j = 1, 2$, and where $R_{ei} \equiv 1 - \tau_e(1 - \alpha_{ei}) - k_{ei}(\alpha_{ei})$. For an individual i, $R_{ei} e_i$ represents the amount that remains after the deduction of the tax on the reported wealth and of the evasion costs; thus $R_{ei} e_i$ can

[6] More formally, the members of each group are represented by the real numbers in the interval [0,1], whose size (Lebesgue measure) is one.

[7] This approach has also been used subsequently in Boadway et al. (1994), Slemrod (2001), Chetty (2009) and Blomquist et al. (2011), among others.

[8] In order to be as general as possible, we allow the cost function $K_{ei}(\alpha_{ei})$ to differ between the two types. As it depends on the evaded share, identical cost functions imply that the evasion of any given amount is cheaper for the wealthier individual, which in our model is also the more able.

[9] Note that convexity of $K_{ei}(\alpha_{ei})$ implies strict convexity of $k_{ei}(\alpha_{ei})$ because $\partial^2 k_{ei}(\alpha_{ei})/\partial \alpha_{ei}^2 = 2\partial K_{ei}(\alpha_{ei})/\partial \alpha_{ei} + \alpha_{ei} \partial^2 K_{ei}(\alpha_{ei})/\partial \alpha_{ei}^2 > 0$.

be interpreted as her effective net initial wealth, which together with net labor income x_i can be spent on consumption.

From the first-order condition for an interior solution of α_{ei} to maximize R_{ei} one gets

$$\tau_e = \partial k_{ei}(\alpha_{ei})/\partial \alpha_{ei}, \tag{2}$$

i.e. individuals conceal wealth until the marginal cost of evasion equals the marginal benefit. Let $R_{ei}(\tau_e)e_i$ denote the optimally chosen effective net initial wealth and observe that the optimal fraction α_{ei} is determined solely by the tax rate and is independent of total initial wealth e_i, which follows from the assumption that per-unit evasion costs depend only on the fraction α_{ei}. Due to convexity $\partial \alpha_{ei}(\tau_e)/\partial \tau_e = (\partial^2 k_{ei}(\alpha_{ei})/\partial \alpha_{ei}^2)^{-1} > 0$ holds. That is, the model has the realistic property that the fraction α_{ei} of misreported wealth rises with the tax rate. Moreover, note that in general both types evade different fractions of e_i, as the evasion technology may differ between types.

Further observe that the assumption $K_{ei}(1) < 1$ (and consequently $k_{ei}(1) < 1$) guarantees that effective net initial wealth $R_{ei}(\tau_e)e_i$ is strictly positive for all tax rates τ_e. This follows immediately from the fact that for $\alpha_{ei} = 1$, $R_{ei}(\tau_e) = 1 - k_{ei}(1) > 0$, which implies that for any tax rate (possibly larger than one) where individuals would report no wealth at all, $R_{ei}(\tau_e)$ is positive. Moreover, for lower tax rates, where individuals report a positive share of wealth ($\alpha_{ei} < 1$), $R_{ei}(\tau_e)$ is even larger, as $\partial R_{ei}(\tau_e)/\partial \tau_e = -(1 - \alpha_{ei}) < 0$.[10]

Firm behavior

Commodity taxes are levied on firms, which have access to an evasion technology similar to that of the households. Each firm of industry j evades commodity taxes by understating its sales at a cost depending on the evaded fraction α_j. Let c_j denote the produced quantity of a firm in industry j. Concealing one unit of good j entails a resource cost described by the function $K_j(\alpha_j)$, with $\partial K_j(\alpha_j)/\partial \alpha_j > 0$, $\partial^2 K_j(\alpha_j)/\partial \alpha_j^2 \geq 0$ and $K_j(0) = 0$. Again we define $k_j(\alpha_j) \equiv \alpha_j K_j(\alpha_j)$.

Then the maximization problem of a firm in industry j can be written as

$$\max \pi_j = p_j c_j - c_j - (1 - \alpha_j)\tau_j c_j - c_j k_j(\alpha_j). \tag{3}$$

The right-hand side of (3) consists of sale revenues minus the cost of production (remember that the constant marginal costs are equal to one), minus taxes paid to tax authorities and minus the cost of tax evasion. From the first-order condition for the optimal fraction α_j one gets for an interior solution

$$\tau_j = \partial k_j(\alpha_j)/\partial \alpha_j. \tag{4}$$

---

10 Use the definition for R_{ei} following (1) and note that due to the Envelope Theorem, only the partial derivative with respect to τ_e needs to be considered.

Thus, firms also evade taxes until the marginal cost of evasion equals its marginal benefit. As before the optimal α_j is determined solely by the tax rate τ_j and is thus independent of the amount c_j sold by the firm. Also $\partial \alpha_j(\tau_j)/\partial \tau_j > 0$ still holds. In equilibrium firms supply good j at a consumer price

$$p_j = 1 + \tau_j(1 - \alpha_j) + k_j(\alpha_j), \tag{5}$$

with α_j chosen optimally according to (4). The corresponding producer price is equal to the constant marginal costs (equal to one). Observe that without tax evasion the consumer price is equal to $1 + \tau_j$. Hence, all gains of tax evasion $(\tau_j \alpha_j - k_j(\alpha_j))$ are transmitted to consumers via a lower consumer price, and firms make zero profits (as usual under the assumption of a linear technology). Finally, for prices given by (5) the equilibrium quantity of each commodity $j = 1, 2$ is determined by aggregate household demand $c_{jL} + c_{jH}$.

The social planner's problem

In a next step we describe the social planner's maximization problem. For this purpose we first introduce the indirect utility function of household i for given tax rates τ_e, τ_1, τ_2 and given net and gross income x_i, z_i,

$$\begin{aligned} &v_i(x_i, z_i, e_i, \tau_e, \tau_1, \tau_2) \\ &\quad \equiv max \, \{u(c_{1i}, c_{2i}, z_i/\omega_i) \, | p_1 c_{1i} + p_2 c_{2i} \leq x_i + R_{ei} e_i \} \, . \end{aligned} \tag{6}$$

Note that the demand functions $c_{ji}(\cdot)$ determined by the maximization problem in (6) depend on the same arguments as the indirect utility function.

First-best taxes referring to abilities are not implementable, because abilities are not observable to the social planner. Nor can the social planner infer abilities from initial wealth, as only reported wealth is observable.[11] Therefore, the tax authority imposes an income tax as a second-best instrument. As there are no restrictions on the functional form of the income tax the standard way of solving such a problem is to maximize a social welfare function with respect to the individuals' net and gross income bundles (x_L, z_L), (x_H, z_H), subject to a self-selection constraint and a resource constraint. By this the optimal income tax for the two types of individuals is determined implicitly as the difference $z_i - x_i$, $i = L, H$. The other available tax instruments τ_e, τ_1 and τ_2 are taken as fixed at some rate for the moment.

[11] One may argue that with a fixed relationship between abilities and initial wealth the social planner could in principle identify individuals by observing reported wealth, and impose a lump-sum tax on abilities. This would make the high-able individual worse off than the low-able individual (Mirrlees 1974) and, thus, be a further incentive for the H-type to conceal wealth. We avoid this complexity by assuming that the social planner cannot impose a lump-sum tax on abilities by using information transmitted through reported wealth. Note, that a related problem arises also in the standard Mirrlees optimum income tax model (Mirrlees 1971): given a tax schedule, gross income of high-able individuals is higher than gross income of low-able individuals. Thus, ex-post the social planner could identify individuals as well.

The utilitarian social welfare function, which is the objective function of the maximization problem, reads

$$\max_{x_i, z_i, i=L,H} f_L v_L(x_L, z_L, e_L, \tau_e, \tau_1, \tau_2) + f_H v_H(x_H, z_H, e_H, \tau_e, \tau_1, \tau_2), \tag{7}$$

where f_L and f_H, with $f_L \geq f_H \geq 0$, represent the weights of the two types of individuals. As usual we assume that preferences fulfill the condition of "agent monotonicity" (Seade 1982). Formally this means that $MRS^L_{zx} > MRS^H_{zx}$ holds at any vector (x, z), where MRS^i_{zx} is defined as $MRS^i_{zx} \equiv -(\partial v_i/\partial z_i)/(\partial v_i/\partial x_i)$. This assumption—also known as the single crossing condition—implies that for any income tax function the high-able individual does not choose to earn less income than the low-able.[12]

Whereas the objective function is standard, the resource constraint has to be modified in our setting. It reads

$$\begin{aligned} x_L + x_H \leq z_L + z_H &+ \tau_e((1-\alpha_{eL})e_L + (1-\alpha_{eH})e_H) \\ &+ \tau_1(1-\alpha_1)(c_{1L}+c_{1H}) + \tau_2(1-\alpha_2)(c_{2L}+c_{2H}) - g. \end{aligned} \tag{8}$$

The social planner has to collect tax revenues in order to finance public spending g. One can see that the base for a tax on initial wealth is reported wealth and not the true amount of wealth. The same holds for the base of commodity taxes τ_j, $j = 1, 2$. As reported wealth and reported consumption decrease with increasing tax rates, an increase of τ_e, τ_1 and τ_2 might even reduce tax revenues. For the taxation of initial wealth the possibility of a Laffer effect only arises due to the existence of tax evasion. This is not the case for the taxation of a single good j, where a Laffer effect might also arise without the possibility of tax evasion due to the existence of substitution effects. Further note that resources spent on the evasion activity represent pure waste, as they are not included in the resource constraint.

The self-selection constraint is again standard. We restrict the analysis to cases, where the social planner wants to redistribute from high- to low-ability persons and, due to the assumption of agent monotonicity, only the self-selection constraint, which prevents the high-able individual from mimicking the low-able individual (i.e. from choosing the bundle which is intended for the low-able) is binding in the optimum. The self-selection constraint reads

$$v_H(x_H, z_H, e_H, \tau_e, \tau_1, \tau_2) \geq v_H(x_L, z_L, e_H, \tau_e, \tau_1, \tau_2). \tag{9}$$

Maximizing (7) subject to (8) and (9) with respect to x_i and z_i yields the first-order conditions for the optimal bundles of net and gross income. The Lagrange multipliers of the resource constraint and the self-selection constraint are denoted by λ for the former and by μ for the latter. The marginal utility of income of the high-able individual

[12] It should be mentioned that the existence of initial endowments makes the assumption of agent-monotonicity more problematic than in the standard case. The reason is that with e_H larger than e_L the high-able individuals' marginal utility of net income might be sufficiently low such that they might demand at least as much additional net income as the low-ability individuals in order to be compensated for achieving an additional unit of gross income.

in the case of mimicking is written as $\partial v_H[L]/\partial x_L$. The first-order conditions for x_i and z_i read

$$f_L\frac{\partial v_L}{\partial x_L}-\lambda+\lambda\tau_1(1-\alpha_1)\frac{\partial c_{1L}}{\partial x_L}+\lambda\tau_2(1-\alpha_2)\frac{\partial c_{2L}}{\partial x_L}-\mu\frac{\partial v_H[L]}{\partial x_L}=0, \tag{10}$$

$$f_H\frac{\partial v_H}{\partial x_H}-\lambda+\lambda\tau_1(1-\alpha_1)\frac{\partial c_{1H}}{\partial x_H}+\lambda\tau_2(1-\alpha_2)\frac{\partial c_{2H}}{\partial x_H}+\mu\frac{\partial v_H}{\partial x_H}=0, \tag{11}$$

$$f_L\frac{\partial v_L}{\partial z_L}+\lambda+\lambda\tau_1(1-\alpha_1)\frac{\partial c_{1L}}{\partial z_L}+\lambda\tau_2(1-\alpha_2)\frac{\partial c_{2L}}{\partial z_L}-\mu\frac{\partial v_H[L]}{\partial z_L}=0, \tag{12}$$

$$f_H\frac{\partial v_H}{\partial z_H}+\lambda+\lambda\tau_1(1-\alpha_1)\frac{\partial c_{1H}}{\partial z_H}+\lambda\tau_2(1-\alpha_2)\frac{\partial c_{2H}}{\partial z_H}+\mu\frac{\partial v_H}{\partial z_H}=0. \tag{13}$$

The optimal income tax is described implicitly by these conditions for given tax rates τ_1, τ_2, τ_e (possibly zero). In the next two sections we analyze the role of taxes on initial wealth and on commodities in an optimal tax system in the presence of evasion opportunities. Clearly, in an economy without tax evasion a tax on exogenously given wealth would be lump-sum and therefore a desirable tax instrument for redistributive reasons, given that ability and wealth are positively correlated. The same holds true for a uniform expenditure tax, which Brunner and Pech (2012a,b) have shown to be equivalent to a tax on initial wealth in an economy without tax evasion possibilities. However, our model has the realistic property that these taxes also cause efficiency losses which have to be weighted against their positive redistributive effect.

3 Optimal taxation of initial wealth

When analyzing whether a proportional tax on initial wealth is a useful supplement to an optimal income tax, the social planner has to take into account two aspects. On the one hand a tax on initial wealth allows for further redistribution, as the high-able individual also owns a larger amount of wealth. But on the other hand a higher tax rate leads to an increase in taxes evaded and might at some point even reduce tax revenues. Proposition 1 addresses this trade-off faced by the social planner. Let the optimal value function of the maximization problem (7)–(9) be denoted by $S(\tau_e, \tau_1, \tau_2)$.

Proposition 1 *The welfare effect of a marginal increase of a tax τ_e on initial wealth, given that x_i, z_i are chosen optimally, is described by*

$$\frac{\partial S}{\partial \tau_e}=\mu\frac{\partial v_H[L]}{\partial x_L}((1-\alpha_{eH})e_H-(1-\alpha_{eL})e_L)-\lambda\tau_e\left(e_L\frac{\partial \alpha_{eL}}{\partial \tau_e}+e_H\frac{\partial \alpha_{eH}}{\partial \tau_e}\right). \tag{14}$$

Despite the existence of tax evasion a positive tax rate on initial wealth is always optimal in our model, as $\frac{\partial S}{\partial \tau_e}\mid_{\tau_e=0}>0$.

Proof The derivation of Eq. (14) is provided in the Appendix.

One can see that the welfare effect of a marginal increase of τ_e consists of two different parts with opposite signs. The first term on the right-hand side of Eq. (14) describes the effect on the self-selection constraint. It is unambiguously positive given that the high-able individual reports a higher amount of wealth to tax authorities than the low-able individual. That is, if $(1-\alpha_{eH})e_H > (1-\alpha_{eL})e_L$ an increase of τ_e relaxes the self-selection constraint.[13] To understand the intuition behind this mechanism, assume in a first step that after a marginal increase of τ_e each individual i is fully compensated for the loss of net initial wealth $\partial(R_{ei}e_i)/\partial\tau_e = -(1-\alpha_{ei})e_i$ by an increase of net income x_i of the same amount. This makes mimicking less attractive, as long as the compensation $(1-\alpha_{ei})e_i$ is higher for the high-able individual, and relaxes the self-selection constraint. Then in a second step further redistribution via the income tax becomes possible and this in turn increases social welfare.[14]

However, due to tax evasion individuals cannot be fully compensated, as an increase of τ_e also increases the fraction of wealth concealed by individuals. This is taken into account by the second part on the RHS of (14), which affects the resource constraint. It is always negative except for a zero tax rate, where it obviously is zero. It can be interpreted as the marginal deadweight loss of tax evasion. The intuition behind this effect is that an increase of τ_e leads to a decline in the fraction of wealth reported and therefore has a negative influence on attainable wealth tax revenues.

Although individuals have the possibility to underreport wealth (at some cost), a positive tax on initial wealth is always optimal in our model, as $\frac{\partial S}{\partial \tau_e}|_{\tau_e=0} > 0$. This is due to the fact that the effect on the self-selection constraint is positive at $\tau_e = 0$, because $\alpha_{ei} = 0$ and $e_H > e_L$, while the effect on the resource constraint is zero at $\tau_e = 0$.

Next, we briefly characterize the optimal tax on initial wealth τ_e^* and the resulting net effective wealth $R_{ei}(\tau_e^*)e_i$. Obviously the optimum occurs when $\frac{\partial S}{\partial \tau_e} = 0$; thus it is obtained by setting the RHS of (14) equal to zero, which is the first-order condition of the maximization problem (7)–(9) for τ_e. Intuitively, the social planner should increase τ_e as long as the positive redistributive effect is larger than the negative deadweight loss effect and set the optimal tax rate τ_e^* such that both effects have the same size.

We know from above that the optimal tax rate is greater than zero. On the other hand, it cannot be optimal to set such a high tax rate that both types of individuals conceal all their wealth ($\alpha_{ei} = 1$ for $i = L, H$), because then government revenues are the same as at a tax rate of zero while evasion costs are wasted. Thus, the optimal tax rate is also bounded from above. At which tax rate individuals would decide to report no wealth at all to tax authorities obviously depends on the evasion technology, but clearly τ_e^* has to be below this tax rate.

An open question remains whether net effective wealth $R_{ei}(\tau_e^*)e_i$ at the optimal tax rate is larger for the H- or for the L-type. If $R_{ei}(\tau_e^*)$ is higher for the L-type then it could be that $R_{eL}(\tau_e^*)e_L > R_{eH}(\tau_e^*)e_H$ holds even with $e_H > e_L$. Loosely speaking

[13] Observe that if $\alpha_{eL} \geq \alpha_{eH}$ the effect becomes zero only if both individuals evade all their wealth. If, however, $\alpha_{eH} > \alpha_{eL}$ it becomes zero for positive amounts of wealth reported to tax authorities.

[14] Note that the validity of this argument does not hinge on our assumption of a perfect correlation of wealth and ability. One can show that the redistributive potential of a wealth tax persists if wealth endowments are stochastic and positively correlated with abilities Brunner and Pech (2008).

this can occur if tax evasion is cheaper for the low-able.[15] We abstract from this case and assume for the rest of the analysis that $R_{eH}(\tau_e^*)e_H > R_{eL}(\tau_e^*)e_L$ holds. There are two arguments which corroborate this assumption. First, it is plausible that wealthier individuals also have an at least as good access to tax evasion activities as the low-able, implying lower (or equal) marginal cost of tax evasion and hence $R_{eH}(\tau_e^*) \geq R_{eL}(\tau_e^*)$ (see footnote 15). Second, even in the opposite case of $R_{eL}(\tau_e^*) > R_{eH}(\tau_e^*)$ the inequality $R_{eH}(\tau_e^*)e_H > R_{eL}(\tau_e^*)e_L$ may hold, given $e_H > e_L$. Altogether, the case of $R_{eH}(\tau_e^*)e_H > R_{eL}(\tau_e^*)e_L$ is certainly the more realistic scenario.

4 Optimal taxation of commodities

A classical result on the role of indirect taxes in an optimal tax system is due to Atkinson and Stiglitz (1976). They showed that when preferences are weakly separable in labor supply and consumption, nonlinear income taxation does not need to be supplemented by commodity taxes. However, this result is derived in a model, where individuals differ in only one characteristic, namely in earning abilities. In a more recent paper Cremer et al. (2001) have shown that even if preferences are weakly separable between labor and consumption, commodity taxation is a useful instrument of tax policy if individuals differ not only in abilities but also in endowments. The role of commodity taxation in their model is to indirectly tax initial endowments which are assumed to be unobservable. This is in contrast to our model where reported endowments are subject to taxation. We proof in the following that even if the latter are taxed optimally, the remaining differences in (net) wealth provide a case for commodity taxes, though these can be partly evaded as well. Furthermore, we analyze the role of tax evasion on the optimal structure of commodity taxes, similar to Cremer and Gahvari (1993). Their study analyzes the influence of tax evasion on optimal commodity taxes in a representative agent model à la Ramsey and is, thus, in contrast to our study, not concerned with redistribution.

4.1 The welfare effect of commodity taxes

How commodity taxation affects social welfare is determined by differentiating the optimal value function of (7)–(9) with respect to τ_j.

Proposition 2 *The welfare effect of a marginal increase of a tax τ_j on commodity $j = 1, 2$, given that x_i, z_i and τ_e are chosen optimally is described by*

$$\frac{\partial S}{\partial \tau_j} = \mu \frac{\partial v_H[L]}{\partial x_L}(1-\alpha_j)(c_{jH}[L] - c_{jL}) - \lambda \tau_j \frac{\partial \alpha_j}{\partial \tau_j}(c_{jL} + c_{jH})$$
$$+\lambda(1-\alpha_j)\sum_{k=1}^{2} \tilde{\tau}_k \left(\frac{\partial c_{kL}^{com}}{\partial p_j} + \frac{\partial c_{kH}^{com}}{\partial p_j} \right). \tag{15}$$

[15] More precisely it is the marginal cost of tax evasion $\partial k_{ei}(\alpha_{ei})/\partial \alpha_{ei}$ which is decisive. If $\partial k_{ei}(\alpha_{ei})/\partial \alpha_{ei}$ is larger (smaller) for the low-able individual at any α_{ei}, then $R_{eH}(\tau_e^*) > R_{eL}(\tau_e^*)$ ($R_{eH}(\tau_e^*) < R_{eL}(\tau_e^*)$) holds as in this case the per-unit rent of tax evasion $\tau_e \alpha_{ei} - k_{ei}(\alpha_{ei})$ is larger (smaller) for the high-able.

Given that commodity j is a normal good and that preferences are weakly separable between labor supply and consumption, welfare can be further increased by an introduction of a commodity tax, as $\frac{\partial S}{\partial \tau_j} |_{\tau_1=\tau_2=0} > 0$.

Proof The derivation of Eq. (15) is provided in the Appendix.

In formula (15), compensated demand for good k is denoted by c_{ki}^{com} and $\tilde{\tau}_k \equiv \tau_k(1-\alpha_k)$. Proposition 2 states that despite the existence of tax evasion positive commodity taxes on normal goods are optimal in our model, even if preferences between labor supply and consumption are weakly separable. From Eq. (15) one can see that the total welfare effect of a marginal increase of τ_j consists of three effects, one redistributive effect on the self-selection constraint (multiplier μ) and two efficiency effects on the resource constraint (multiplier λ). First, consider the effect on the self-selection constraint. The intuition is similar to that for the redistributive effect of τ_e mentioned above: consider a marginal increase of $\Delta\tau_j$ with both types being compensated by an increase of net income $\Delta x_i = \Delta\tau_j c_{ji}$, $i = L, H$, such that welfare of both types remains approximately unchanged. However, if now the H-type mimics the L-type, the former has to bear the tax increase $\Delta\tau_j c_{jH}[L]$, which is more than the compensating increase $\Delta x_L = \Delta\tau_j c_{jL}$ in net income of the mimicked individual (observe that in the case of normal goods $c_{jH}[L] > c_{jL}$ because of $R_{eH}e_H > R_{eL}e_L$).[16] Consequently, a marginal increase of $\Delta\tau_j$ makes mimicking less attractive, as long as $\alpha_j < 1$ (firms report a positive amount of sales, thus the consumer price increases with τ_j).[17] In other words, it relaxes the self-selection constraint and allows for additional redistribution. Hence, the redistributive effect of τ_j is positive.

The efficiency effects in Eq. (15) describe the deadweight loss due to tax evasion and the distorting effects on compensated demand. The interpretation of the deadweight loss effect induced by tax evasion, which is described by the second term on the right hand side of (15), is quite similar to the one given in the preceding section for the case of wealth taxes. Higher tax rates lead to an increase of the fraction of hidden sales and therefore reduce the tax base for the commodity tax. Finally, the last term in Eq. (15) represents the effects on compensated demand associated with the distortion of the consumer price p_j due to an increase of τ_j. Both efficiency effects are of second order, thus for a zero tax rate they are zero, while the effect on the self-selection constraint is positive at $\tau_j = 0$. Hence, an introduction of a commodity tax increases welfare.[18] To summarize, a role for commodity taxes arises in our model even if they induce tax evasion and even if exogenous initial wealth is taxed optimally. Thus, in the economy we describe, it is optimal to supplement the income tax by all other available tax instruments. The reason is that this allows to balance the deadweight loss effects created by tax evasion in the best way.

A further interesting point is that even the introduction of a uniform expenditure tax improves welfare. This follows from the fact that the welfare effect of introducing

[16] In an economy without initial wealth $c_{jH}[L] = c_{jL}$ holds, if preferences are weakly separable in labor supply and consumption. Then one obviously is back at the classical Atkinson-Stiglitz result.

[17] Note that $\partial p_j/\partial\tau_j = 1 - \alpha_j$. With $\alpha_j = 1$ we have $p_j = 1 + k_j(1)$ and a further increase of τ_j has no effect.

[18] For an inferior good the introduction of a subsidy improves welfare, as then $c_{jH}[L] < c_{jL}$.

a uniform tax on all expenditures ($\tau_1 = \tau_2 = \tau$) consists of the sum over $j = 1, 2$ of the welfare effects given by (15). That is, also in this case there is a positive first-order redistributive effect while the negative efficiency effects are of second-order. Still, differential commodity taxes are in general optimal in our framework, in particular if evasion technologies differ between the two commodities. In the rest of the paper we analyze the optimal structure of commodity taxes and elaborate a condition when uniform taxation of commodities indeed turns out to be optimal. In this discussion we only consider the case of normal goods.

The optimal tax structure is described by the following relationship, which is obtained by setting the first-order-conditions for τ_1 and τ_2 from Eq. (15) equal to zero:

$$\frac{c_{1H}[L] - c_{1L}}{c_{2H}[L] - c_{2L}} = \frac{\frac{\tau_1}{1-\alpha_1}\frac{\partial \alpha_1}{\partial \tau_1}(c_{1L} + c_{1H}) - \sum_i \left(\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_1} + \tilde{\tau}_2 \frac{\partial c_{2i}^{com}}{\partial p_1}\right)}{\frac{\tau_2}{1-\alpha_2}\frac{\partial \alpha_2}{\partial \tau_2}(c_{2L} + c_{2H}) - \sum_i \left(\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_2} + \tilde{\tau}_2 \frac{\partial c_{2i}^{com}}{\partial p_2}\right)}. \tag{16}$$

Observe that the optimal tax structure depends on the relative size of the redistributive and efficiency effects, which in turn depend on preferences and on the evasion technologies for both commodity taxes. However, as the size of those effects remains rather arbitrary it is impossible to draw any precise conclusion on the optimal structure of τ_1 and τ_2 for the general case. To gain more insight we now turn to two special cases. The first one deals with uniform evasion costs for both commodities, which implies that for $\tau_1 = \tau_2$ firms would conceal the same fraction α_j of c_1 and c_2. In the second special case we assume that only the tax on one of the two goods can be evaded.

4.2 Uniform evasion costs for both commodity taxes

Assume now that evasion costs are uniform, i.e. $k_1(\alpha_1) = k_2(\alpha_2)$, the cost of concealing a fraction α_j is the same for both goods. Then obviously, for any uniform tax rate τ, the same fraction $\alpha_1 = \alpha_2 \equiv \alpha$ is concealed. For this special case we can show that no substitution effects occur with the introduction of a uniform commodity tax, and we can draw some conclusions under what circumstances uniform commodity taxation is optimal.

Proposition 3 *Let evasion costs for both commodities be identical.*

(a) *The welfare effect of a marginal increase of a uniform tax τ on both commodities, given that x_i, z_i and τ_e are chosen optimally reads as*

$$\frac{\partial S}{\partial \tau} = \mu \frac{\partial v_H[L]}{\partial x_L}(1-\alpha)\left(\sum_{j=1}^{2} c_{jH}[L] - c_{jL}\right) + \lambda\tau\frac{\partial \alpha}{\partial \tau}\sum_{j=1}^{2}\sum_{i=L,H} c_{ji}. \tag{17}$$

This effect is positive at $\tau = 0$, given that both commodities are normal goods and that preferences are weakly separable between labor supply and consumption.

(b) *If preferences are homothetic with respect to c_1 and c_2 and weakly separable between consumption and labor, then uniform commodity taxes are optimal.*

Proof The proof of Proposition 3 is provided in the Appendix.

Note that with homothetic preferences both goods are normal, thus the case of a uniform tax being optimal implies that a positive tax rate is desirable. Clearly with identical evasion costs a uniform tax increase induces a uniform price increase for both goods [see (5)]. As a consequence, no substitution effects (compensated price effects) occur, compared to the increase of a single tax (Proposition 2). Then a motive for differential commodity taxation can only arise from some asymmetry of the households' preferences with respect to the two commodities. Proposition 3 states that for homothetic preferences (i. e., when the relative importance of the two goods does not vary with the available budget) a uniform commodity tax is optimal. This result is especially interesting as it is related to a standard result of the optimum taxation literature, which states that if income is subject to an optimal linear income tax, uniform commodity taxation is optimal if preferences between consumption and labor supply are weakly separable and if Engel curves for goods are linear (but need not go through the origin).[19] We find that if individuals differ in initial wealth and if commodity taxes can be evaded at a uniform cost, preferences have to be weakly separable between consumption and labor and homothetic in consumption for uniform commodity taxes to be optimal, even if income can be taxed nonlinearly. Note that in our model initial wealth exists and can only be taxed at a proportional rate, which is the analogy to the restriction in Deaton (1979) that income can only be taxed linearly.

If uniform taxes are not optimal, no general conclusion can be drawn as to the relative size of the tax rates in the optimum, even in case of linear Engel curves (not crossing the origin) and identical evasion costs. However, Proposition 2 shows that the welfare-enhancing effect of an introduction of a commodity tax is larger for that good whose consumption increases relatively more with income.[20] This conforms to the intuition that taxing a luxury good more heavily than a necessity may be desirable. But for higher tax rates the shape of the evasion-cost function and compensated price reactions enter in a complex way which prevents the derivation of a definite result.

4.3 The tax on good two cannot be evaded

Assume now that the tax on good two cannot be evaded because marginal evasion costs are infinitely high, i.e. $k_2'(0) = \infty$. A possible illustration for such a scenario could be that c_1 represents services while c_2 represents goods, as it is plausible that taxes on services can be evaded more easily than taxes on goods. Clearly, the assumption that a tax on some good cannot be evaded at all is too strict, but it helps to illustrate the point we want to make. Intuitively one might expect that such a situation would call for taxing the commodity that cannot be evaded higher than the commodity for which tax evasion is possible. However, it turns out that this need not be the case as one also has to take into account the distorting effects on compensated demand. This

[19] See for example Deaton (1979).

[20] For $\tau_j = 0$ and, thus, $\alpha_j = 0$ the relative size of $c_{jH}[L] - c_{jL}$ determines whether the marginal welfare effect is larger for good 1 or good 2 [see Eq. (15)]. $c_{jH}[L] - c_{jL}$ is the difference in consumption of good j between the mimicking individual - who has a larger initial wealth—and the mimicked individual.

can be seen from Eq. (18), which is an adapted version of Eq. (16). Note that now we have $\alpha_2 = 0$ for all τ_2:

$$\frac{c_{1H}[L] - c_{1L}}{c_{2H}[L] - c_{2L}} = \frac{\frac{\tau_1}{1-\alpha_1}\frac{\partial \alpha_1}{\partial \tau_1}(c_{1L} + c_{1H}) - \sum_i (\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_1} + \tau_2 \frac{\partial c_{2i}^{com}}{\partial p_1})}{-\sum_i (\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_2} + \tau_2 \frac{\partial c_{2i}^{com}}{\partial p_2})} \tag{18}$$

Optimal tax rates for τ_1 and τ_2 have to satisfy this condition. We can conclude that given normal goods the right-hand side of (18) must be positive. It is well-known that compensated demand has the property (homogeneity) that $\sum_j p_j \frac{\partial c_{ji}^{com}}{\partial p_k} = 0$, for $k = 1, 2$ and any $i = L, H$. This and negativity (positivity) of own (cross, respectively) compensated price effects imply that the summation terms $\sum_i (\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_k} + \tau_2 \frac{\partial c_{2i}^{com}}{\partial p_k})$, $k = 1, 2$, in the numerator and denominator, respectively, of the RHS of (18) have opposite signs for arbitrary $\tilde{\tau}_1, \tau_2$. They clearly cannot be zero. This in turn means that the denominator must be positive because otherwise the RHS of (18) would be negative (the numerator would be positive). Next, $\sum_i (\tilde{\tau}_1 \frac{\partial c_{1i}^{com}}{\partial p_2} + \tau_2 \frac{\partial c_{2i}^{com}}{\partial p_2}) < 0$ implies, again due to homogeneity, that $\frac{\tilde{\tau}_1}{\tau_2} \leq \frac{p_1}{p_2}$. Finally, using (5) and $\alpha_2 = 0$, we get $\tau_1(1 - \alpha_1) \leq \tau_2(1 + k_1(\alpha_1))$, which obviously holds for $\tau_1 = \tau_2$ as $\alpha_1 < 1$ and $k_1(\alpha_1) > 0$. Thus, without specifying in more detail the cost function $k_1(\alpha_1)$ and preferences we cannot tell whether $\tau_1 \gtrless \tau_2$ is optimal.

5 Conclusion

In this paper we have extended the standard model of optimal income taxation by an important aspect which conforms to reality: individuals differ not only in earning abilities, but also in initial wealth. The government can impose a rather comprehensive set of taxes: a nonlinear tax on labor income and proportional taxes on wealth and on commodities. Moreover, we have introduced the restriction that the latter two taxes can—at some cost—be evaded by individuals and firms, respectively. We analyzed the question of whether there is a role for these taxes in a welfare-maximizing tax system.

It turned out that, given the essential condition that abilities and initial wealth are positively correlated, a tax on wealth—in addition to an optimal nonlinear income tax—is desirable, even if it can be evaded. Further, even if income and wealth are taxed optimally, taxes on commodities still raise social welfare, given that consumption increases with income. Thus, the result in the Atkinson–Stiglitz model that an optimal income tax does not need to be supplemented by commodity taxes if preferences are weakly separable between labor and consumption, does not arise in our model. The main reason for this clearly comes from the existence of initial wealth as a second characteristic, which distinguishes individuals and calls for redistribution via the wealth tax. As the deadweight loss of evasion is of second order, it does not outweigh the redistributive effect as long as the tax rate is not too high. On the other hand, a tax on commodities can, in principle, perform the same task as the tax on wealth (Brunner and Pech 2012a,b). However, due to the second-order effect of tax evasion it is optimal to impose taxes on both wealth and commodities in our model, because then the overall deadweight loss is smaller. Thus, our model provides a rationale for the existence of a comprehensive tax system, as we find it in most industrialized economies.

In our study, taxes on wealth and on commodities are restricted to be linear. For commodity taxes this is clearly justified by the fact that they are levied on every single transaction. Total consumption is unobservable for the tax authority, which makes it impossible to have a tax rate varying with the total amount. This argument does not apply for the wealth tax, which can indeed be constructed as a nonlinear function. The main reason for our linearity assumption was tractability of the model in order to derive definite results; modeling how the extent of redistribution is limited by the possibility of evasion is quite complex with an arbitrary nonlinear wealth tax. Note, however, that a nonlinear wealth tax would allow even higher social welfare by a more differentiated treatment of the individuals, that is, by performing more redistribution. Thus, its desirability is implied by Proposition 1 stating the desirability of a linear wealth tax in our model.

Finally, it should be mentioned that obviously in reality also the income tax can be evaded, which is excluded in our model. However, as mentioned in the introduction, we may conclude from Boadway et al. (1994) that allowing for income tax evasion would only reinforce the case for commodity taxation.

Acknowledgments The authors gratefully acknowledge valuable suggestions by two anonymous referees. This research was supported by the Austrian Science Funds (FWF).

6 Appendix

Proof of Proposition 1 The Lagrangian for the maximization problem (7)–(9) reads

$$\begin{aligned} L = {} & f_L v_L(x_L, z_L, e_L, \tau_e, \tau_1, \tau_2) + f_H v_H(x_H, z_H, e_L, \tau_e, \tau_1, \tau_2) \\ & -\lambda(x_L + x_H - z_L - z_H - \tau_e((1-\alpha_{eL})e_L + (1-\alpha_{eH})e_H) \\ & -\tau_1(1-\alpha_1)(c_{1L} + c_{1H}) - \tau_2(1-\alpha_2)(c_{2L} + c_{2H}) + g) \\ & +\mu(v_H(x_H, z_H, e_H, \tau_e, \tau_1, \tau_2) - v_H(x_L, z_L, e_H, \tau_e, \tau_1, \tau_2)). \end{aligned} \tag{19}$$

To abbreviate notation we write $v_H[L] \equiv v_H(x_L, z_L, e_H, \tau_e, \tau_1, \tau_2)$. Using the Envelope Theorem we get for the optimal value function $S(\tau_e, \tau_1, \tau_2)$

$$\begin{aligned} \frac{\partial S}{\partial \tau_e} = {} & f_L \frac{\partial v_L}{\partial \tau_e} + f_H \frac{\partial v_H}{\partial \tau_e} + \lambda((1-\alpha_{eL})e_L + (1-\alpha_{eH})e_H) \\ & -\lambda\tau_e\left(\frac{\partial \alpha_{eL}}{\partial \tau_e}e_L + \frac{\partial \alpha_{eH}}{\partial \tau_e}e_H\right) + \lambda\tau_1(1-\alpha_1)\left(\frac{\partial c_{1L}}{\partial \tau_e} + \frac{\partial c_{1H}}{\partial \tau_e}\right) \\ & +\lambda\tau_2(1-\alpha_2)\left(\frac{\partial c_{2L}}{\partial \tau_e} + \frac{\partial c_{2H}}{\partial \tau_e}\right) + \mu\frac{\partial v_H}{\partial \tau_e} - \mu\frac{\partial v_H[L]}{\partial \tau_e}. \end{aligned} \tag{20}$$

In a next step we use $\frac{\partial v_i}{\partial \tau_e} = \frac{\partial R_{ei}}{\partial \tau_e} e_i \frac{\partial v_i}{\partial x_i}$, $\frac{\partial v_H[L]}{\partial \tau_e} = \frac{\partial R_{eH}}{\partial \tau_e} e_H \frac{\partial v_H[L]}{\partial x_L}$ and $\frac{\partial c_{ji}}{\partial \tau_e} = \frac{\partial R_{ei}}{\partial \tau_e} e_i \frac{\partial c_{ji}}{\partial x_i}$, where $\frac{\partial R_{ei}}{\partial \tau_e} = -(1-\alpha_{ei})$. Plugging those expressions into (20) and substituting for $f_i \frac{\partial v_i}{\partial x_i}$ from (10) and (11) yields equation (14) in the text.

Proof of Proposition 2 We make again use of the Envelope Theorem. The derivative of the optimal value function $S(\tau_e, \tau_1, \tau_2)$ from the maximization problem represented by the Lagrangian in (19) with respect to τ_j, $j = 1, 2$, reads

$$\begin{aligned}\frac{\partial S}{\partial \tau_j} &= f_L \frac{\partial v_L}{\partial \tau_j} + f_H \frac{\partial v_H}{\partial \tau_j} + \lambda \Bigg[\left(1 - \alpha_j - \tau_j \frac{\partial \alpha_j}{\partial \tau_j}\right)(c_{jL} + c_{jH}) \\ &\quad + \tau_j (1-\alpha_j) \left(\frac{\partial c_{jL}}{\partial \tau_j} + \frac{\partial c_{jH}}{\partial \tau_j}\right) + \tau_k (1-\alpha_k) \left(\frac{\partial c_{kL}}{\partial \tau_j} + \frac{\partial c_{kH}}{\partial \tau_j}\right) \Bigg] \\ &\quad + \mu \left(\frac{\partial v_H}{\partial \tau_j} - \frac{\partial v_H[L]}{\partial \tau_j}\right),\end{aligned} \tag{21}$$

with $k = 3 - j$. We find $\frac{\partial v_i}{\partial \tau_j} = -\frac{\partial p_j}{\partial \tau_j} c_{ji} \frac{\partial v_i}{\partial x_i}$, with $\frac{\partial p_j}{\partial \tau_j} = 1 - \alpha_j$. Note also that $\frac{\partial c_{ji}}{\partial \tau_j} = \frac{\partial c_{ji}}{\partial p_j} \frac{\partial p_j}{\partial \tau_j}$. By use of these expressions and substituting for $f_i \frac{\partial v_i}{\partial x_i}$ from (10) and (11), (21) can be transformed to

$$\begin{aligned}\frac{\partial S}{\partial \tau_j} &= \mu \frac{\partial v_H[L]}{\partial x_L}(1-\alpha_j)(c_{jH}[L] - c_{jL}) - \lambda \tau_j \frac{\partial \alpha_j}{\partial \tau_j}(c_{jL} + c_{jH}) \\ &\quad + \lambda (1-\alpha_j) \left[\tilde{\tau}_j \sum_{i=L}^{H} \left(\frac{\partial c_{ji}}{\partial p_j} + c_{ji} \frac{\partial c_{ji}}{\partial x_i}\right) + \tilde{\tau}_k \sum_{i=L}^{H} \left(\frac{\partial c_{ki}}{\partial p_j} + c_{ji} \frac{\partial c_{ki}}{\partial x_i}\right)\right]\end{aligned} \tag{22}$$

By use of the Slutsky equation one obtains Eq. (15) in the text.

Proof of Proposition 3 (a) First, observe that $\frac{\partial S}{\partial \tau} = \sum_{j=1}^{2} \frac{\partial S}{\partial \tau_j}$, with the latter effects shown in Proposition 2. Next, observe that with $k_1(\alpha_1) = k_2(\alpha_2)$ for $\alpha_1 = \alpha_2$, we have $p_1 = p_2$ [use (5)] and $\tilde{\tau}_1 = \tilde{\tau}_2$, if $\tau_1 = \tau_2$. As is well-known, compensated demand is homogeneous of degree zero in the price vector. Note that in our model compensated demand [as used in (15) and (16)] refers to consumer choice concerning consumption of the two goods, for given net and gross income (hence, in particular for given labor time). Thus, by Euler's law, homogeneity implies, for $j = 1, 2$

$$p_1 \frac{\partial c_{1i}^{com}}{\partial p_j} + p_2 \frac{\partial c_{2i}^{com}}{\partial p_j} = 0 \tag{23}$$

for any $p_1, p_2 \geq 0$. In particular $p(\frac{\partial c_{1i}^{com}}{\partial p_j} + \frac{\partial c_{2i}^{com}}{\partial p_j}) = 0$ for $p_1 = p_2 \equiv p$, thus $\tilde{\tau}(\frac{\partial c_{1i}^{com}}{\partial p_j} + \frac{\partial c_{2i}^{com}}{\partial p_j}) = 0$, for $\tilde{\tau}_1 = \tilde{\tau}_2 \equiv \tilde{\tau}$. Thus, the last effect on the RHS of (15) drops out for $j = 1, 2$.

(b) Moreover, $\tilde{\tau}(\frac{\partial c_{1i}^{com}}{\partial p_j} + \frac{\partial c_{2i}^{com}}{\partial p_j}) = 0$, for $\tilde{\tau}_1 = \tilde{\tau}_2 \equiv \tilde{\tau}$, implies that on the RHS of (16) the effects on compensated demand are zero. Then for $\tau_1 = \tau_2$ we have $\frac{\tau_1 \frac{\partial \alpha_1}{\partial \tau_1}}{1-\alpha_1} = \frac{\tau_2 \frac{\partial \alpha_2}{\partial \tau_2}}{1-\alpha_2}$, and thus (16) reduces to

$$\frac{c_{1H}[L] - c_{1L}}{c_{2H}[L] - c_{2L}} = \frac{c_{1L} + c_{1H}}{c_{2L} + c_{2H}}. \tag{24}$$

Finally, if preferences of each individual i for good 1 and 2 are homothetic and weakly separable between consumption and labor, each individual i spends the same constant share g_j of her budget b_i on each commodity $j = 1, 2$. Then (24) can be rewritten as

$$\frac{g_1(b_H[L] - b_L)}{g_2(b_H[L] - b_L)} = \frac{g_1(b_L + b_H)}{g_2(b_L + b_H)}, \tag{25}$$

which is obviously true. Altogether we have shown that with uniform evasion costs, weakly separable and homothetic preferences the optimality condition (16) is fulfilled for $\tau_1 = \tau_2$.

References

Atkinson AB, Stiglitz JE (1976) The design of tax structure: direct versus indirect taxation. J Public Econ 6:55–75

Atkinson AB, Piketty T, Saez E (2011) Top incomes in the long run of history. J Econ Lit 49:3–71

Blomquist S, Christiansen V, Micheletto L (2011) Public provision of private goods, self-selection and income tax avoidance. Working paper. http://www.nek.uu.se/Pdf/wp201106ucfs.pdf

Boadway R, Marchand M, Pestieau P (1994) Towards a theory of the direct-indirect tax mix. J Public Econ 55:71–88

Boadway R, Marchand M, Pestieau P (2000) Redistribution with unobservable bequests: a case for taxing capital income. Scand J Econ 102:253–267

Brunner JK, Pech S (2008) Optimum taxation of inheritances. Working paper. http://www.econ.jku.at/papers/2008/wp0806.pdf

Brunner JK, Pech S (2012a) Optimal taxation of bequests in a model with initial wealth. Scand J Econ (forthcoming)

Brunner JK, Pech S (2012b) Optimal taxation of wealth transfers when bequests are motivated by joy of giving. The BE J Econ Anal Pol 12(1) (Art 8)

Chetty R (2009) Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. Am Econ J Econ Pol 1(2):31–52

Cremer H, Gahvari F (1993) Tax evasion and optimal commodity taxation. J Public Econ 50:261–275

Cremer H, Gahvari F (1995) Tax evasion and optimum general income tax. J Public Econ 60:235–249

Cremer H, Pestieau P, Rochet JCh (2001) Direct versus indirect taxation: the design of the tax structure revisited. Int Econ Rev 42:781–799

Cremer H, Pestieau P, Rochet JCH (2003) Capital income taxation when inherited wealth is not observable. J Public Econ 87:2475–2490

Deaton A (1979) Optimally uniform commodity taxes. Econ Lett 2:357–361

Davies JB, Shorrocks AF (2000) The distribution of wealth. In: Atkinson AB, Bourguignon F (eds) Handbook of income distribution, vol 1. North-Holland, Amsterdam, pp 605–675

Diaz-Giménez J, Glover A, Rios-Rull JV (2011) Facts on US distributions of earning, income and wealth: 2007 Update. Federal Reserve Bank of Minneap Quart Rev 34(1):2–31

Keen M, Smith S (2006) VAT fraud and evasion: what do we know and what can be done? Nat Tax J 59:861–887

Mirrlees JA (1971) An Exploration in the theory of optimal income taxation. Rev Econ Stud 38:175–208
Mirrlees JA (1974) Notes on welfare economics, information, and uncertainty. In: Balch MS, McFadden DL, Wu SY (eds) Essays on economic behavior under uncertainty. North Holland, Amsterdam, pp 243–258
Pestieau P, Possen UM, Slutsky SM (2004) Jointly optimal taxes and enforcement policies in response to tax evasion. J Public Econ Theory 6:337–374
Seade J (1982) On the sign of the optimum marginal income taxation. Rev Econ Stud 49:637–643
Slemrod J (2001) A general model of the behavioral response to taxation. Int Tax Public Fin 8:119–128
Usher D (1986) Tax evasion and the marginal cost of public funds. Econ Inquiry 24:563–586