BONN ECON DISCUSSION PAPERS

Discussion Paper 14/2000

Income taxation and production efficiency in a simple two-sector economy

by

Thomas Gaube

October 2000

Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post World Net

MAIL EXPRESS LOGISTICS FINANCE

Income taxation and production efficiency in a simple two-sector economy

Thomas Gaube*

October 2000

Abstract

In a recent contribution, H. Naito (1999) has shown that production efficiency may be violated in the optimum with non-linear income taxation. Using a slightly simpler framework, this paper complements Naito's analysis in showing that production efficiency does not hold in the optimum with (i) non-linear and (ii) linear income taxation provided that second best and first best do not coincide. These findings indicate that income taxation generally implies the desirability to complement the distortion between consumer and producer prices by means of a corresponding distortion in input prices.

Keywords: income taxation, production efficiency

JEL-Classification: H21, H23

*Department of Economics, University of Bonn, Adenauerallee 24-42, D-53113 Bonn, Germany, phone: +49-228-739236, fax: +49-228-739239, e-mail: thomas.gaube@wiwi.uni-bonn.de

1 Introduction

The production efficiency theorem of Diamond and Mirrlees (1971) is one of the main results in the theory of (linear) commodity taxation. Since this finding holds under quite general and reasonable conditions,[1] one is tempted to conclude that it is satisfied in models of non-linear taxation as well. In a recent contribution, however, H. Naito (1999) has shown that production efficiency may not hold in the optimum with non-linear income taxation. This result demonstrates that inefficient production can be desirable in a well-founded and broadly accepted framework and points to an important difference[2] between non-linear income and linear commodity taxation, i.e. the two main models of second-best tax analysis.

Employing the two-type self-selection approach to income taxation (Stiglitz 1982, Stern 1982), Naito (1999) investigates a model with three sectors of production and endogenous factor prices. The analysis concentrates on those allocations where at least one of the two self-selection constraints is slack in the optimum. These are the relevant regimes in the basic version of the income tax model where a single production sector and constant factor prices are assumed and where production efficiency holds in second best.[3] In contrast to that basic version, however, the general framework allows also for situations where both self-selection constraints are binding. Since it is not clear which regime prevails in the optimum, the question emerges whether production efficiency is violated in general or just in specific regimes of the second best.

The present paper is meant as a first step towards clarifying this question. In concentrating on a complete description of possible equilibria, I restrain the analysis to the simplest framework which allows to investigate potential inefficiencies in production. Therefore, two sectors and endogenous factor prices are assumed. Within this framework, it is shown that production efficiency is always violated in the optimum with non-linear income taxation provided that this optimum is not first best as well. Second, a similar result is established for the special case of linear income taxation. These results complement Naito's finding and indicate that distortionary income

[1] Production efficiency is desirable if profits can be taxed at the full rate and if no exogenous restrictions on the tax rates are imposed. See section 4 for a discussion of these issues.

[2] Note that various results which have been obtained within the Diamond-Mirrlees framework are based on (extensions of) the production efficiency theorem. For some recent examples, see project evaluation (Drèze and Stern, 1990), international taxation (Eggert and Haufler, 1999; Keen and Wildasin, 2000), and the provision of public inputs (Feehan and Matsumoto, 1999).

[3] See Guesnerie and Seade (1982) for a proof of production efficiency in a one-sector model with exogenous factor prices. Note, however, that the analysis of Diamond and Mirrlees (1971) is related to more general situations where at least two units of production are available.

taxation generally provides an incentive to introduce inefficiencies of production.

The intuition behind the first result is similar to that of Naito (1999). Due to the assumption of a strictly quasiconcave technology, deviations from production efficiency affect relative factor prices which in turn allows to weaken the individuals' self-selection constraints. Hence, inefficient production can be used to mitigate the cost of asymmetric information. With linear income taxation, however, the individuals' self-selection constraints are not binding in the optimum. Still, inefficient production is desirable because a linear income tax - in contrast to linear commodity taxation - does not allow to vary the ratio of the individuals' net wages independently of the corresponding ratio of the producers' gross wages.[4]

The remainder of this paper is organized as follows. Section 2 presents the model. In sections 3 and 4, the optimal structure of production is analyzed with non-linear and linear income taxation respectively. Section 5 concludes.

2 The model

The economy consists of two types of households $i = 1, 2$ and two sectors of production. The N_1 households of type 1 may differ from the N_2 households of type 2 with respect to their wage rates ω_i, but not with respect to preferences or endowments. Preferences are represented by a strictly monotone, strictly quasiconcave, and twice continuously differentiable utility function $U(X_i, L_i)$, where X_i and L_i denote private consumption and labor supply respectively. Conventionally, it is assumed that the term $L_i(-U_L^i/U_X^i)$ is increasing in L_i. This (agent monotonicity) condition holds if private consumption is normal and ensures that the single crossing property is satisfied for all allocations with $\omega_1 \neq \omega_2$. Without loss of generality, I also assume $N_1 = N_2 = 1$.

The commodity X is produced in the private sector according to the aggregate production function $F(L_{1X}, L_{2X})$, where L_{1X} and L_{2X} denote the sector specific labor inputs. Normalizing the price of X to unity, the industry's profits can be written in the form $\Pi_X = F(L_{1X}, L_{2X}) - \omega_1 L_{1X} - \omega_2 L_{2X}$. The function $F(L_{1X}, L_{2X})$ is assumed to be linear homogeneous, twice continuously differentiable, and strictly quasiconcave.[5] The same assumptions are made with respect to the public sector where the commodity G is produced according to the aggregate production function

[4]Hence, a linear income tax represents an exogenous restriction on the powers to tax. Note that such restrictions have also been analyzed by Stiglitz and Dasgupta (1971) and Munk (1980). However, while they concentrate on a representative consumer framework, the present analysis emphasizes the role of distributional objectives for introducing inefficiencies in production.

[5]Note that the basic model of income taxation assumes a single production sector with a linear

$H(L_{1G}, L_{2G})$. This commodity can be interpreted as public consumption (affecting the individuals' utilities) or public investments (affecting the private sector's production function). In order to simplify the analysis, however, a specific interpretation of G is avoided by making the assumption that an exogenous amount $\bar{G}$ of the commodity has to be produced. Consequently, the set of technically feasible allocations can be described by the conditions

$$F(L_{1X}, L_{2X}) - X_1 - X_2 \geq 0, \qquad H(L_{1G}, L_{2G}) - \bar{G} \geq 0, \tag{1}$$

$$L_1 - L_{1X} - L_{1G} \geq 0, \qquad L_2 - L_{2X} - L_{2G} \geq 0. \tag{2}$$

Let F_i and H_i denote the partial derivatives of the production functions with respect to the labor inputs L_{iX} and L_{iG} respectively. Restricting attention to strictly positive vectors $(\{L_{iX}, L_{iG}\}_{i=1}^2)$, it is clear that overall production efficiency can only be fulfilled if the condition $F_1/F_2 = H_1/H_2$ is satisfied. Note that profit maximization implies cost minimization in terms of the market prices ω_i. We thus have

$$f(L_{1X}, L_{2X}) := \frac{F_1(\cdot)}{F_2(\cdot)} = \frac{\omega_1}{\omega_2} \tag{3}$$

in a competitive equilibrium. In order to investigate whether production efficiency is desirable, it is assumed that public firms minimize costs with respect to the shadow prices $\tilde{\omega}_i$ set by the government. This implies $h(L_{1G}, L_{2G}) := H_1(\cdot)/H_2(\cdot) = \tilde{\omega}_1/\tilde{\omega}_2$. Consequently, production efficiency holds in the optimum only if $\Omega := \omega_1/\omega_2 = \tilde{\omega}_1/\tilde{\omega}_2 := \tilde{\Omega}$.

Public expenditures are financed by taxing the individuals' wage incomes $Y_i := \omega_i L_i$ according to a possibly non-linear function $T(Y_i)$. Income taxation is motivated by the assumption that the government can observe the pre-tax incomes Y_i, but not their components ω_i and L_i. Referring to the revelation principle, this informational restriction is usually formalized by means of the individuals' self-selection constraints. Based on the definition $V^i(X_i, Y_i) := U(X_i, Y_i/\omega_i)$, these constraints can be written in the form $V^1(X_1, Y_1) \geq V^1(X_2, Y_2)$ and $V^2(X_2, Y_2) \geq V^2(X_1, Y_1)$, i.e. in terms of the observable variables X_i and Y_i. These inequalities mean that an individual of type i prefers to choose the tax contract $(X_i, Y_i) = (Y_i - T(Y_i), Y_i)$ instead of mimicking the other type $j \neq i$ by choosing (X_j, Y_j). Within the framework analyzed here, however, an alternative version of the incentive constraints is more

technology. With two sectors of production, however, strict quasiconcavity has to be assumed in order to avoid corner solutions. Consequently, the factor prices ω_i are endogenous. For the analysis of a one-sector model with a strictly quasiconcave technology, see Stiglitz (1982) and Stern (1982).

appropriate (see Stiglitz 1982, p. 222). Reformulating the self-selection constraints in commodity space (X_i, L_i), we get

$$U(X_1, L_1) \geq U\left(X_2, L_2\frac{\omega_2}{\omega_1}\right) \quad \text{and} \quad U(X_2, L_2) \geq U\left(X_1, L_1\frac{\omega_1}{\omega_2}\right). \tag{4}$$

Taking care of the restrictions (1) - (4), the government seeks to implement a Pareto efficient allocation. Since the instruments $(L_i, X_i, \tilde{\omega}_i)$, $i = 1, 2$ allow full command of the sector specific inputs L_{iX} and L_{iG}, the government's maximization problem can most easily be expressed in terms of the controls $(L_i, L_{iX}, L_{iG}, X_i)$, $i = 1, 2$. Therefore, the optimum with income taxation is defined by means of

$$\left(\{X_i^S, L_i^S, L_{iX}^S, L_{iG}^S\}_{i=1}^2\right) := \underset{X_i, L_i, L_{iX}, L_{iG}}{\operatorname{argmax}} \{U(X_2, L_2) \mid U(X_1, L_1) - \bar{U}^1 \geq 0, \quad (1), (2), (3), (4)\}. \tag{5}$$

In the following, this allocation will be analyzed in order to investigate whether production efficiency is desirable in second best.

3 Inefficiency of production in second best

This section contrasts the production efficiency rule $F_1/F_2 = H_1/H_2$ with the optimal production rule which follows from the first-order conditions of the second-best problem (5). Substituting the eq. (3) into (4) the corresponding Lagrangian can be written in the form

$$\begin{aligned}\mathcal{L} \;=\; & U(X_2, L_2) + \delta[U(X_1, L_1) - \bar{U}^1] + \lambda_1[U(X_1, L_1) - U(X_2, L_2/f(\cdot))] \\ & + \lambda_2[U(X_2, L_2) - U(X_1, L_1 f(\cdot))] + \lambda_F[F(L_{1X}, L_{2X}) - X_1 - X_2] \\ & + \lambda_G[H(L_{1G}, L_{2G}) - \bar{G}] + \gamma_1[L_1 - L_{1X} - L_{1G}] + \gamma_2[L_2 - L_{2X} - L_{2G}].\end{aligned}$$

Using the definitions $U_X^i := \partial U(X_i, L_i)/\partial X_i$, $U_L^i := \partial U(X_i, L_i)/\partial L_i$, $\hat{L}_1 := L_2/\Omega$, $\hat{L}_2 := L_1\Omega$, $\hat{U}_X^i := \partial U(X_j, \hat{L}_i)/\partial X_j$, $\hat{U}_L^i := \partial U(X_j, \hat{L}_i)/\partial \hat{L}_i$, and $f_i := \partial f(L_{1X}, L_{2X})/\partial L_{iX}$, we get the first-order conditions

$$X_1: \quad \delta U_X^1 + \lambda_1 U_X^1 - \lambda_2 \hat{U}_X^2 - \lambda_F = 0 \tag{6}$$

$$X_2: \quad U_X^2 - \lambda_1 \hat{U}_X^1 + \lambda_2 U_X^2 - \lambda_F = 0 \tag{7}$$

$$L_1: \quad \delta U_L^1 + \lambda_1 U_L^1 - \lambda_2 \hat{U}_L^2 \Omega + \gamma_1 = 0 \tag{8}$$

$$L_2: \quad U_L^2 - \lambda_1 \hat{U}_L^1 \frac{1}{\Omega} + \lambda_2 U_L^2 + \gamma_2 = 0 \tag{9}$$

$$L_{iX}: \quad \lambda_F F_i - \gamma_i + \mathcal{L}_\Omega f_i = 0, \quad i = 1, 2 \tag{10}$$

$$L_{iG}: \quad \lambda_G H_i - \gamma_i = 0, \quad i = 1,2 \tag{11}$$

where the term $\mathcal{L}_\Omega := \lambda_2 \left(-\hat{U}_L^2\right) L_1 - \lambda_1 \left(-\hat{U}_L^1\right) (L_2/\Omega^2)$ shows how an increase of the wage ratio Ω affects welfare by influencing the self-selection constraints (4).

In order to investigate the optimal structure of production, consider the equations (10) and (11). They imply

$$\frac{F_1}{F_2} = \frac{\gamma_1 - f_1 \mathcal{L}_\Omega}{\gamma_2 - f_2 \mathcal{L}_\Omega} \qquad \text{and} \qquad \frac{H_1}{H_2} = \frac{\gamma_1}{\gamma_2}. \tag{12}$$

As shown in the Appendix, we have $f_2 > 0 > f_1$. Hence, the equation (12) is consistent with the rule $F_1/F_2 = H_1/H_2$ if and only if $\mathcal{L}_\Omega = 0$.[6] In order to work out the circumstances under which production efficiency holds in second best, we thus have to analyze the sign of the partial derivative $\mathcal{L}_\Omega$. Following the conventional taxonomy, the potential solutions of (6) - (11) can be classified by considering the set of binding self-selection constraints. This approach leads to four potential regimes: (i) $\lambda_2 = 0$, $\lambda_1 = 0$; (ii) $\lambda_2 > 0$, $\lambda_1 = 0$; (iii) $\lambda_2 = 0$, $\lambda_1 > 0$; and (iv) $\lambda_2 > 0$, $\lambda_1 > 0$. Since factor prices are endogenous, one has also to distinguish between the cases $\omega_1 \neq \omega_2$ and $\omega_1 = \omega_2$. Consider first the situation $\omega_1 \neq \omega_2$ and assume without loss of generality $\omega_1 < \omega_2$. For this case, the regimes (i) - (iv) are illustrated in parts (a) - (d) of Figure 1.[7] Part (a) of Figure 1 describes a situation where $\lambda_1 = \lambda_2 = 0$, i.e. where none of the two self-selection constraints is binding. Hence, second best and first best coincide. Consequently, we also have $\mathcal{L}_\Omega = 0$, i.e. production efficiency in the optimum. The illustrations (b) and (c) describe those cases which are mostly discussed in the literature and which have also been analyzed in Naito (1999). In situation (b) only the self-selection constraint of the 'skilled' individuals[8] is binding. Since this situation reflects the desirability of income redistribution in favor of the unskilled individuals, it is commonly termed the *redistributive case*. Conversely, situation (c) describes the *regressive case* where just the self-selection constraint of the unskilled individuals is binding. Hence, in contrast to first best, we have $\mathcal{L}_\Omega \neq 0$ in the cases (b) and (c) which in turn implies that production efficiency is violated.

- Figure 1 about here -

[6]Note that the first-order conditions (6) - (11) imply $\lambda_F > 0$, $\lambda_G > 0$, and $\gamma_i > 0$, $i = 1,2$. This can easily be shown by going through all the potential cases (a) - (f) discussed below.

[7]These illustrations are similar to those of Stiglitz (1982).

[8]In the following, an individual of type $i \in \{1,2\}$ will be labeled 'skilled' ('unskilled') if $\omega_i > (<)\, \omega_j$. Since factor prices are endogenous, these terms are not used in a global sense, but only to differentiate between the individuals in a specific allocation with $\omega_1 \neq \omega_2$.

Consider now case (d) which describes a second-best optimum with *bunching*. Here, both self-selection constraints are binding which means that production efficiency (i.e. $\mathcal{L}_\Omega = 0$) cannot be ruled out a priori. The same problem occurs if we have $\omega_1 = \omega_2$ in the optimum, i.e. the cases (e) or (f).[9] Again, both self-selection constraints hold with equality which may lead to $\mathcal{L}_\Omega = 0$. Hence, the potential optima (d) - (f) raise doubts that deviations from production efficiency are generally desirable in second best. Since it is not clear whether equilibria of the type (b) or (c) Pareto dominate the situations (d) - (f), the intuition derived from (b) and (c) might not be valid in the true optimum with income taxation (5).

Due to these concerns it is of interest to analyze the situations (d) - (f) as well and to identify those types of equilibria which lead to $\mathcal{L}_\Omega = 0$. The following result shows that this condition holds only if second best and first best coincide. This means that we have $\mathcal{L}_\Omega \neq 0$ not only in the cases (b) and (c), but in any second-best optimum provided that the self-selection constraints (4) do play any role at all.

Proposition 1: *If the optimum with income taxation (5) is not first best, then*
(a) the production efficiency rule $F_1/F_2 = H_1/H_2$ *is violated and*
(b) the two types $i = 1, 2$ *differ either with respect to* ω_i *or* Y_i.

Proof: See the Appendix.

Part (a) of Proposition 1 is the main result of this paper: Provided that the optimal income tax distorts the consumption decision of at least one type of individuals, the production efficiency rule $F_1/F_2 = H_1/H_2$ should also be violated. Hence, while the Diamond-Mirrlees (1971) framework leads to a dichotomy between inefficiencies in consumption and production, these distortions are intertwined in the income tax model not only in specific circumstances, but in all potential regimes of the second-best optimum (5). The intuition behind this finding is straightforward. Since a deviation from production efficiency has an impact on the market wage ratio Ω, it can be used to weaken the self-selection constraints (4). Hence, inefficient production is desirable because it affects the two types in opposite ways which in turn allows to mitigate the model's basic trade-off between distributional objectives and Pareto efficiency.[10]

[9] These cases are ruled out in the basic model with constant factor prices ω_i. Note also that the self-selection constraints (4) imply $U(X_1, L_1) = U(X_2, L_2)$ for equilibria with $\omega_1 = \omega_2$. Hence, the two indifference curves coincide and both constraints (4) must be satisfied with equality. Therefore, the two multipliers λ_i can be strictly positive.

[10] This interaction between inefficiencies in production and the self-selection constraints (4) becomes apparent not only by means of the first-order conditions (6) - (11). In fact, it is possible to construct examples where no productively efficient allocation satisfies the constraints contained in eq. (5) while allocations which violate production efficiency do.

Note that the second best allocation is first best only if the individuals' marginal rates of substitution between income Y_i and consumption X_i are equal to unity.[11] Since this cannot hold for both types in the situations (d) and (e), these cases imply $\mathcal{L}_\Omega \neq 0$. In contrast, an optimum of type (f) must be first best (see part (b) of Proposition 1). However, this situation requires 'identical' individuals in the sense that both types achieve the same wage rate when providing the same quantity of labor. Hence, if one imposes a minimum degree of asymmetry between the factors L_1 and L_2,[12] first-best allocations can only be of type (a). Here, however, the more productive individual is better off which means $U(X_1, L_1) \neq U(X_2, L_2)$. Whether such a solution is indeed second best depends on distributional considerations alone.[13]

Proposition 1 leaves open in which way the optimal shadow price ratio $\tilde{\Omega} = H_1/H_2$ deviates from the market price ratio $\Omega = F_1/F_2$. Still, using the eq. (12), it is clear that the inequality $\tilde{\Omega} < (>)\, \Omega$ holds in second best if and only if $\mathcal{L}_\Omega > (<)\, 0$. For the intuition behind this relationship consider the case $\mathcal{L}_\Omega > 0$. This condition means that an increase of Ω increases welfare by relaxing at least one of the self-selection constraints (4). Since Ω increases if labor of type 1 (i.e. the factor L_{1X}) becomes scarcer in the private sector, the constraints (4) provide an incentive to increase L_{1G} relative to L_{2G} which means that the shadow price $\tilde{\Omega}$ falls below of Ω.

Hence, the comparison between shadow prices and market prices depends on the sign of $\mathcal{L}_\Omega$. Provided that just one of the self-selection constraints is binding, this sign can directly be determined. While (b) and (c) are probably the most relevant cases, it should be noted that the arguments used in the proof of Proposition 1 allow for analogous results with respect to the cases (d) and (e) as well. The following Corollary thus compares $\tilde{\Omega}$ and Ω for all potential regimes of the second-best optimum (5).

Corollary: *Assume that the optimum with income taxation (5) is not first best.* *(a) Consider a situation where $\omega_1 \neq \omega_2$ and assume without loss of generality $\omega_2 > \omega_1$. (i) In the redistributive (regressive) case, the shadow wage ratio $\tilde{\Omega}$ between the unskilled type and the skilled type is below (above) the corresponding wage ratio Ω in the private sector. (ii) In an optimum with bunching, the comparison between $\tilde{\Omega}$*

[11] This condition is equivalent to $(-U^i_L)/U^i_X = \omega_i$, $i = 1, 2$.

[12] With endogenous factor prices, the second type (or equivalently, the first type) can be labeled as more 'productive' if both conditions $f(L_{1X}, L_{2X}) \mid_{L_{1X}=L_{2X}} \leq 1$ and $h(L_{1G}, L_{2G}) \mid_{L_{1G}=L_{2G}} \leq 1$ are satisfied where at least one of the inequalities is strict. This assumption rules out second-best optima of type (f).

[13] Note that case (a) cannot be second best if the government has sufficiently strong egalitarian preferences. This becomes clear by employing the Rawlsian welfare function $\text{Min}\{U(X_1, L_1), U(X_2, L_2)\}$ instead of the Pareto-programme used above.

and Ω follows the result obtained for the regressive case.
(b) Consider a situation where $\omega_1 = \omega_2$ and assume without loss of generality $Y_2 > Y_1$. Then the shadow wage ratio $\tilde{\Omega}$ between the low-income type and the high-income type is below the corresponding wage ratio Ω in the private sector.

Proof: See the Appendix.

The redistributive case reflects a situation where the government - in the absence of informational constraints - would like to redistribute income from the skilled individuals towards the unskilled individuals. This is the 'normal' case of Stiglitz (1982) and has attained most interest in the literature. From this perspective, part (a)-(i) of the Corollary provides the main result following from the production rule (12): Provided that income redistribution in favor of the unskilled individuals is desirable, the public sector should not minimize costs with respect to market prices, but should produce with an inefficiently high ratio between unskilled and skilled labor (see also Naito 1999). Note, however, that the comparison between $\tilde{\Omega}$ and Ω in the equilibrium with bunching follows the result obtained for the regressive case. While bunching might be dismissed from an empirical point of view, this regime cannot be ruled out a-priori even if the government has strong egalitarian preferences.[14] Hence, it is not clear whether strong incentives for redistribution generally lead to an optimum which has the same qualitative properties as the redistributive case.

4 Linear Income Taxation

While most of the income tax literature deals with the non-linear case discussed above, the special case of linear income taxation has gained attention in the literature as well.[15] Since a linear income tax is not only easier to analyze, but also less demanding with respect to implementation, it is interesting to know whether the finding $F_1/F_2 \neq H_1/H_2$ prevails in such a setting.

Proposition 1 shows that production efficiency is violated if at least one of the self-selection constraints (4) is binding in second best. Note, however, that these constraints are slack with linear income taxation. Therefore, the arguments presented above do not apply. Still, the results of Diamond and Mirrlees (1971), Stiglitz and Dasgupta (1971), and Munk (1980) show that the production efficiency theorem is

[14]For an illustration of this point, consider again the Rawlsian criterion $\text{Min}\{U(X_1, L_1), U(X_2, L_2)\}$. It is easy to see from the corresponding first-order conditions that this welfare function rules out the regressive case, but not an optimum with bunching.

[15]For a brief overview, see Myles (1995), Chap. 5. While most of this literature assumes exogenous factor prices, Allen (1982) analyzes a model which - except for his assumption of a single production sector - is quite similar to the one presented here.

based on the assumption that "consumer prices (or equivalently tax rates) can be chosen independently of producer prices." (Mirrlees 1986, p. 1220). A linear income tax, however, does not allow to vary the ratio of net wages (consumer prices) independently of the ratio of gross wages (producer prices). The subsequent analysis thus complements the earlier literature in showing that statement (a) of Proposition 1 holds with linear income taxation as well.

Consider a linear income tax $T(Y_i) = tY_i - I$, where $t < 1$ is the marginal tax rate[16] and $I \in \mathbb{R}$ is a uniform lump-sum transfer if positive and a tax if negative. Normalizing the price of the consumption commodity to unity, an individual of type i thus encounters the budget constraint $L_i\omega_i(1-t) + I - X_i \geq 0$. Maximizing utility $U(X_i, L_i)$ subject to this constraint, we get the individual's demand $X_i(\bar{\omega}_i, I)$, supply $L_i(\bar{\omega}_i, I)$, and indirect utility $W_i(\bar{\omega}_i, I)$ in terms of the transfer I and the net wage $\bar{\omega}_i := (1-t)\omega_i$. As in the model of section 2, profit maximization implies $\omega_i = F_i(L_{1X}, L_{2X})$, $i = 1, 2$. In a competitive equilibrium, we thus have

$$X_i = X_i(F_i(1-t), I), \quad L_i = L_i(F_i(1-t), I), \quad W_i = W_i(F_i(1-t), I). \tag{13}$$

Following the literature on linear income taxation, it is assumed that the government maximizes a welfare function $\Psi(W_1, W_2)$ where the partial derivatives $\Psi_i := \partial\Psi/\partial W_i$ are strictly positive. Hence, the second-best optimum is defined by means of

$$\left(t^*, I^*, \{L_{iX}^*, L_{iG}^*\}_{i=1}^2\right) := \underset{t,I,L_{iX},L_{iG}}{\operatorname{argmax}} \{\Psi(W_1, W_2) \mid (1), (2), (13)\}. \tag{14}$$

Substituting the constraints (13) into (1) and (2), the Lagrangian of the government's maximization problem can be written as follows:

$$\begin{aligned}\mathcal{L} = {} & \Psi(W_1(\cdot), W_2(\cdot)) + \lambda_F[F(L_{1X}, L_{2X}) - X_1(\cdot) - X_2(\cdot)] \\ & +\lambda_G[H(L_{1G}, L_{2G}) - \bar{G}] + \gamma_1[L_1(\cdot) - L_{1X} - L_{1G}] \\ & +\gamma_2[L_2(\cdot) - L_{2X} - L_{2G}].\end{aligned}$$

Let denote W_I^i, X_I^i, and L_I^i the partial derivatives of the functions $W_i(\cdot), X_i(\cdot)$, and $L_i(\cdot)$ with respect to income I. Using these definitions, we get the first-order conditions

$$t: \qquad \frac{\partial \mathcal{L}}{\partial \bar{\omega}_1}\omega_1 + \frac{\partial \mathcal{L}}{\partial \bar{\omega}_2}\omega_2 = 0 \tag{15}$$

$$I: \qquad \left(\Psi_1 W_I^1 - \lambda_F X_I^1 + \gamma_1 L_I^1\right) + \left(\Psi_2 W_I^2 - \lambda_F X_I^2 + \gamma_2 L_I^2\right) = 0 \tag{16}$$

$$L_{iX}: \quad \lambda_F F_i - \gamma_i + \frac{\partial \mathcal{L}}{\partial \bar{\omega}_1}(1-t)F_{1i} + \frac{\partial \mathcal{L}}{\partial \bar{\omega}_2}(1-t)F_{2i} = 0, \quad i = 1, 2 \tag{17}$$

[16]Note that the individuals do not supply a strictly positive amount L_i if $t \geq 1$.

$$L_{iG}: \quad \lambda_G H_i - \gamma_i = 0, \quad i = 1,2 \tag{18}$$

where the terms $\partial\mathcal{L}/\partial\bar{\omega}_i$ are the partial derivatives of the Lagrangian with respect to $\bar{\omega}_i$. Again, the potential deviations from production efficiency can be described by analyzing the first-order conditions with respect to L_{iX} and L_{iG}. As shown in the Appendix, they imply

$$\frac{F_1}{F_2} >,=,< \frac{H_1}{H_2} \quad \Leftrightarrow \quad \frac{\partial\mathcal{L}}{\partial\bar{\omega}_1} >,=,< \frac{\partial\mathcal{L}}{\partial\bar{\omega}_2}. \tag{19}$$

Hence, the production efficiency rule $F_1/F_2 = H_1/H_2$ holds in second best if and only if $\partial\mathcal{L}/\partial\bar{\omega}_1 = \partial\mathcal{L}/\partial\bar{\omega}_2 = 0$. The following result shows that this condition can only be satisfied if the marginal tax rate t^* is zero which means that the optimum with linear income taxation (14) is first best as well.

Proposition 2: *If the optimum with linear income taxation (14) is not first best, then the production efficiency rule $F_1/F_2 = H_1/H_2$ is violated.*

Proof: See the Appendix.

Proposition 2 demonstrates that a linear income tax also implies the desirability to complement the distortion between consumer and producer prices by means of a corresponding distortion in production. Note that Proposition 2 is close to the findings of Stiglitz and Dasgupta (1971) and Munk (1980) who have shown that production efficiency does not hold within the Diamond-Mirrlees framework provided that some factors or commodities cannot be taxed at all.[17] Here, the result is not based on the choice of some untaxed commodities, but on the implicit restriction to tax both factors at the same rate. Of course, such exogenous restrictions on the powers to tax can be dismissed for being ad hoc (see Mirrlees 1986, p. 1220). In contrast to the case of untaxed commodities, however, the linearity-constraint underlying Proposition 2 does only change the formal argument, but not the conclusion which is similar to that of Proposition 1.

The eq. (19) does also make clear how the shadow wage ratio $\tilde{\Omega} = H_1/H_2$ and the market wage ratio $\Omega = F_1/F_2$ differ from each other. It shows that the situations (i) $\partial\mathcal{L}/\partial\bar{\omega}_1 > \partial\mathcal{L}/\partial\bar{\omega}_2$ and (ii) $\partial\mathcal{L}/\partial\bar{\omega}_1 < \partial\mathcal{L}/\partial\bar{\omega}_2$ correspond to $\Omega > \tilde{\Omega}$ and $\Omega < \tilde{\Omega}$ respectively. For an intuition behind this relationship, consider the case (i). Since the indirect utility functions $W_i(\cdot)$ are strictly increasing in ω_i, this situation means that redistribution in favor of the first type of individuals is desirable.[18] Accordingly,

[17]These papers deal also with question whether profits can be taxed at the full rate. Since profits are zero in the model presented here, Propositions 1 and 2 are not related to that issue.

[18]Using the same techniques as in the proof of Proposition 2 it can be shown the case (i) [case (ii)] implies a positive sign of the first [second] term on the left hand side of eq. (16). The individuals of type 1 thus have a higher [lower] social marginal utility of income than those of type 2.

welfare increases if the factor L_{1X} becomes scarcer in the private sector. Since this can be accomplished by reducing the shadow wage ratio $\tilde{\Omega}$, the result $\tilde{\Omega} < \Omega$ follows immediately.

5 Conclusion

Using a two-type two-sector model with endogenous wage rates, this paper demonstrates that production efficiency is violated in the optimum with (i) non-linear and (ii) linear income taxation provided that these allocations are not first best as well. These results complement the analysis of Naito (1999) and indicate that distortive income taxation generally provides an incentive to influence relative factor prices by means of inefficient production.

The findings are derived within the simplest model of income taxation given the requirements that factor prices are endogenous and that at least two sectors of production are available. This framework allows to analyze the interaction between inefficient production and the equilibrium price vector in terms of a one-dimensional variable, i.e. the wage ratio between the two types of individuals. It is clear that generalizations of the model with respect to an arbitrary number of types and consumption commodities would intricate the analysis considerably. This holds in particular for the model with non-linear income taxation which becomes far more complex if a multidimensional vector of relative factor and consumption prices is taken into account. Hence, the question whether the rather strong relationship between deviations from first best and deviations from production efficiency survives in a more general setting is left open for further investigation.

Appendix

A1: Proof of $\mathbf{f_2 > 0 > f_1}$

The function $F(L_{1X}, L_{2X})$ is assumed to be homogeneous of degree one in (L_{1X}, L_{2X}). Therefore, the functions $F_1(\cdot)$, $F_2(\cdot)$, and $f(\cdot)$ are homogeneous of degree zero in (L_{1X}, L_{2X}). This implies $f_1 L_{1X} + f_2 L_{2X} = 0$. Since $F(L_{1X}, L_{2X})$ is strictly quasiconcave, we also have

$$\begin{aligned} F_2 f_1 - F_1 f_2 &= F_2 \left(\frac{F_{11}F_2 - F_{21}F_1}{(F_2)^2} \right) - F_1 \left(\frac{F_{12}F_2 - F_{22}F_1}{(F_2)^2} \right) \\ &= \left(\frac{1}{F_2} \right)^2 \left(F_{11}(F_2)^2 + F_{22}(F_1)^2 - 2F_{12}F_1F_2 \right) < 0. \end{aligned} \tag{20}$$

It is immediate to see that the restrictions $f_1 L_{1X} + f_2 L_{2X} = 0$ and (20) do not allow for $f_1 = 0$ or $f_1 > 0$. Hence, $f_1 < 0$ which in turn implies $f_2 > 0$.

A2: Proof of Proposition 1

(a) Because of the eq. (12), the claim $F_1/F_2 \neq H_1/H_2$ is correct if and only if $\mathcal{L}_\Omega \neq 0$. Therefore, it has to be shown that the inequality $\mathcal{L}_\Omega \neq 0$ holds in the optimum (5) provided that this allocation is not first best. Note that the allocation (5) is first best if $\lambda_1 = \lambda_2 = 0$ and that the situations $\lambda_1 > \lambda_2 = 0$, $\lambda_2 > \lambda_1 = 0$ obviously imply $\mathcal{L}_\Omega \neq 0$. We can thus restrict attention to the cases (d) - (f), where it is possible that both multipliers λ_i are strictly positive. In the first step, it is shown that case (d) [i.e. $\omega_1 \neq \omega_2$, $\lambda_1 > 0$, $\lambda_2 > 0$] rules out $\mathcal{L}_\Omega = 0$. The second step shows that an optimum with $\omega_1 = \omega_2$ [i.e. of type (e) or (f)] is only consistent with $\mathcal{L}_\Omega = 0$ as long as it is first best as well.

(Step 1) Assume $\omega_1 \neq \omega_2$, $\lambda_1 > 0$, and $\lambda_2 > 0$ in second best. Due to the agent monotonicity condition, we thus have bunching, i.e. $Y_1 = Y_2$ and $X_1 = X_2$. Because of $X_1 = X_2$, the self-selection constraints (4) imply $\hat{L}_1 = L_1$ and $\hat{L}_2 = L_2$. Accordingly, $\hat{U}_L^1 = U_L^1$, $\hat{U}_X^1 = U_X^1$, $\hat{U}_L^2 = U_L^2$, and $\hat{U}_X^2 = U_X^2$. Hence, after substituting the eqs. (10) into (8) and (9) the first-order conditions (6) - (9) can be written in the form

$$X_1: \quad \delta U_X^1 - \left(\lambda_2 U_X^2 - \lambda_1 U_X^1\right) = \lambda_F \tag{21}$$

$$X_2: \quad U_X^2 + \left(\lambda_2 U_X^2 - \lambda_1 U_X^1\right) = \lambda_F \tag{22}$$

$$L_1: \quad \delta U_L^1 + \left(\lambda_2 \left(-U_L^2\right) \Omega - \lambda_1 \left(-U_L^1\right)\right) + \lambda_F F_1 + \mathcal{L}_\Omega f_1 = 0 \tag{23}$$

$$L_2: \quad U_L^2 - \frac{1}{\Omega} \left(\lambda_2 \left(-U_L^2\right) \Omega - \lambda_1 \left(-U_L^1\right)\right) + \lambda_F F_2 + \mathcal{L}_\Omega f_2 = 0 \tag{24}$$

where $\mathcal{L}_\Omega = \lambda_2 \left(-U_L^2\right) L_1 - \lambda_1 \left(-U_L^1\right) \left(L_2/\Omega^2\right)$. Note that $Y_1 = Y_2$ (i.e. $\omega_1 L_1 = \omega_2 L_2$) implies $\Omega = L_2/L_1$. Hence,

$$\mathcal{L}_\Omega = \frac{L_2}{\Omega^2} \left(\lambda_2 \left(-U_L^2\right) \Omega - \lambda_1 \left(-U_L^1\right)\right) = \frac{1}{\Omega} \left(\lambda_2 \left(-U_L^2\right) L_2 - \lambda_1 \left(-U_L^1\right) L_1\right). \tag{25}$$

Consider now the eqs. (23) and (24) and note that profit maximization leads to $F_i = \omega_i$, $i = 1, 2$. We thus have $F_i L_i = Y_i$, $i = 1, 2$. Hence,

$$L_1: \quad -\delta U_L^1 L_1 = L_1 \left(\lambda_2 \left(-U_L^2\right) \Omega - \lambda_1 \left(-U_L^1\right)\right) + \lambda_F Y_1 + \mathcal{L}_\Omega f_1 L_1$$

$$L_2: \quad -U_L^2 L_2 = -\frac{L_2}{\Omega} \left(\lambda_2 \left(-U_L^2\right) \Omega - \lambda_1 \left(-U_L^1\right)\right) + \lambda_F Y_2 + \mathcal{L}_\Omega f_2 L_2.$$

Using the first part of (25), these equations can be written in the form

$$L_1: \quad -\delta U_L^1 L_1 = \lambda_F Y_1 + \mathcal{L}_\Omega(\Omega + f_1 L_1) \tag{26}$$

$$L_2: \quad -U_L^2 L_2 = \lambda_F Y_2 + \mathcal{L}_\Omega(-\Omega + f_2 L_2). \tag{27}$$

Substituting these equations into (21) and (22), we get

$$\alpha_1 := \frac{(-U_L^1)L_1}{U_X^1} = \frac{\lambda_F Y_1 + \mathcal{L}_\Omega(\Omega + f_1 L_1)}{\lambda_F + \kappa} \tag{28}$$

$$\alpha_2 := \frac{(-U_L^2)L_2}{U_X^2} = \frac{\lambda_F Y_2 + \mathcal{L}_\Omega(-\Omega + f_2 L_2)}{\lambda_F - \kappa} \tag{29}$$

where $\kappa := (\lambda_2 U_X^2 - \lambda_1 U_X^1)$. Consider again the derivative $\mathcal{L}_\Omega$ and note that the second part of (25) implies

$$\frac{\lambda_1}{\lambda_2} = \frac{(-U_L^2)\,L_2}{(-U_L^1)\,L_1}\left(1 - \frac{\mathcal{L}_\Omega \Omega}{\lambda_2\,(-U_L^2)\,L_2}\right).$$

Substituting this relationship into the definition of the variable κ, we get

$$\begin{aligned}
\kappa &= \lambda_2 \left(U_X^2 - \frac{(-U_L^2)\,L_2}{(-U_L^1)\,L_1}\left(1 - \frac{\mathcal{L}_\Omega \Omega}{\lambda_2\,(-U_L^2)\,L_2}\right) U_X^1\right) \\
&= \lambda_2 \left(-U_L^2\right) L_2 \left(\frac{U_X^2}{(-U_L^2)\,L_2} - \left(1 - \frac{\mathcal{L}_\Omega \Omega}{\lambda_2\,(-U_L^2)\,L_2}\right)\frac{U_X^1}{(-U_L^1)\,L_1}\right) \\
&= \lambda_2 \left(-U_L^2\right) L_2 \left(\frac{1}{\alpha_2} - \frac{1}{\alpha_1}\right) + \frac{\mathcal{L}_\Omega \Omega}{\alpha_1}.
\end{aligned} \tag{30}$$

Now assume $\mathcal{L}_\Omega = 0$. Using the eqs. (28) - (30), it is easy to see that this assumption rules out the possibilities $\kappa > 0$ and $\kappa < 0$. Hence, we have $\kappa = 0$ which in turn leads to $(-U_L^1 L_1)/U_X^1 = (-U_L^2 L_2)/U_X^2$. However, since $L_1 \neq L_2$ and $X_1 = X_2$, this contradicts the agent monotonicity condition.

(Step 2) Consider now a second-best allocation with $\omega_1 = \omega_2$. Because of the self-selection constraints (4) we thus have $U(X_1, L_1) = U(X_2, L_2)$ as well as $\hat{L}_1 = L_2$ and $\hat{L}_2 = L_1$. This implies $\hat{U}_L^1 = U_L^2$, $\hat{U}_X^1 = U_X^2$, $\hat{U}_L^2 = U_L^1$, and $\hat{U}_X^2 = U_X^1$. Hence, after substituting the eqs. (10) into (8) and (9) the first-order conditions (6) - (9) reduce to

$$X_1: \quad U_X^1(\delta + (\lambda_1 - \lambda_2)) = \lambda_F$$

$$X_2: \quad U_X^2(1 + (\lambda_2 - \lambda_1)) = \lambda_F$$

$$L_1: \quad \left(-U_L^1\right)(\delta + (\lambda_1 - \lambda_2)) = \lambda_F F_1 + \mathcal{L}_\Omega f_1$$

$$L_2: \quad \left(-U_L^2\right)(1+(\lambda_2-\lambda_1)) = \lambda_F F_2 + \mathcal{L}_\Omega f_2,$$

where $\mathcal{L}_\Omega = \lambda_2 \left(-U_L^1\right) L_1 - \lambda_1 \left(-U_L^2\right) L_2$. Note that profit maximization leads to $F_i = \omega_i$, $i = 1, 2$. We thus have

$$\frac{-U_L^1}{U_X^1} = \omega_1 + \frac{\mathcal{L}_\Omega f_1}{\lambda_F} \quad \text{and} \quad \frac{-U_L^2}{U_X^2} = \omega_2 + \frac{\mathcal{L}_\Omega f_2}{\lambda_F}. \tag{31}$$

Now assume $\mathcal{L}_\Omega = 0$. Because of (31), this implies not only production efficiency, but also $(-U_L^i/U_X^i) = \omega_i$, $i = 1, 2$. Hence, the marginal tax rates of both types are zero which means that the second-best allocation would be optimal also in absence of the self-selection constraints (4). This, however, contradicts the presumption that second best and first best do not coincide.

(b) It has to be shown that at least one of the inequalities $\omega_1 \neq \omega_2$, $Y_1 \neq Y_2$ holds in second best. For a proof of this claim, consider a situation with $\omega_1 = \omega_2$ and assume that the corresponding allocation is not first best. Because of $\mathcal{L}_\Omega \neq 0$ and $f_1 < 0 < f_2$, the eq. (31) implies $\mathcal{L}_\Omega f_1 \neq \mathcal{L}_\Omega f_2$. Hence, we have $-U_L^1/U_X^1 \neq -U_L^2/U_X^2$. Since $U(X_1, L_1) = U(X_2, L_2)$ in the second-best optimum with $\omega_1 = \omega_2$, this means $L_1 \neq L_2$ which in turn implies $Y_1 \neq Y_2$.

A3: Proof of Corollary

Since the relation $\tilde{\Omega} > (<)\ \Omega$ holds in second best if and only if $\mathcal{L}_\Omega < (>)\ 0$, the proof of the statement (a)-(i) is obvious. With respect to the statements (a)-(ii) and (b), it has to be shown that bunching implies $\mathcal{L}_\Omega < 0$ for the case $\omega_2 > \omega_1$ (Step 1), and that $\omega_2 = \omega_1$ implies $\mathcal{L}_\Omega > 0$ for the case $Y_2 > Y_1$ (Step 2).

(Step 1) Consider an optimum with bunching where $\omega_2 > \omega_1$. Since bunching implies $Y_1 = Y_2$, we have $L_1 > L_2$. Using the agent monotonicity condition, this means $\alpha_1 = (-U_L^1 L_1)/U_X^1 > (-U_L^2 L_2)/U_X^2 = \alpha_2$. Because of $\alpha_1 > \alpha_2$ and $\lambda_2 \left(-U_L^2\right) L_2 = \mathcal{L}_\Omega \Omega + \lambda_1 \left(-U_L^1\right) L_1 > \mathcal{L}_\Omega \Omega$, the eq. (30) implies

$$\kappa > \mathcal{L}_\Omega \Omega \left(\frac{1}{\alpha_2} - \frac{1}{\alpha_1}\right) + \frac{\mathcal{L}_\Omega \Omega}{\alpha_1} = \frac{\mathcal{L}_\Omega \Omega}{\alpha_2}. \tag{32}$$

Subtracting (22) from (21) and (27) from (26), we get $\delta U_X^1 - U_X^2 - 2\kappa = 0$ and $\delta(-U_L^1)L_1 - (-U_L^2)L_2 - \mathcal{L}_\Omega(2\Omega + f_1 L_1 - f_2 L_2) = 0$. This leads to

$$\frac{\delta U_X^1}{U_X^2} - 1 - \frac{2\kappa}{U_X^2} = 0, \quad \frac{\delta(-U_L^1)L_1}{(-U_L^2)L_2} - 1 - \frac{\mathcal{L}_\Omega(2\Omega + f_1 L_1 - f_2 L_2)}{(-U_L^2)L_2} = 0. \tag{33}$$

Note that the inequality $\alpha_1 > \alpha_2$ implies $(\delta(-U_L^1)L_1)/((-U_L^2)L_2) \geq (\delta U_X^1)/U_X^2$. Because of the equations in (33), we thus have

$$\mathcal{L}_\Omega(2\Omega + f_1 L_1 - f_2 L_2) \geq 2\kappa \frac{(-U_L^2)L_2}{U_X^2} = 2\kappa\alpha_2 > 2\mathcal{L}_\Omega \Omega,$$

where the last inequality follows from (32). This implies $\mathcal{L}_\Omega(f_1 L_1 - f_2 L_2) > 0$. Since $f_1 < 0 < f_2$, we thus have $\mathcal{L}_\Omega < 0$.

(Step 2) Consider now a situation with $\omega_1 = \omega_2$ and $Y_2 > Y_1$. This implies $L_2 > L_1$. Since $U(X_1, L_1) = U(X_2, L_2)$ and since the utility function $U(X_i, L_i)$ is strictly quasiconcave, we thus have $-U_L^2/U_X^2 > -U_L^1/U_X^1$. Using the eq. (31) as well as $f_1 < 0 < f_2$, this implies $\mathcal{L}_\Omega > 0$.

A4: Derivation of eq. (19)

Because of $\omega_1/\omega_2 = F_1/F_2$, the eq. (15) can be written in the form $\partial\mathcal{L}/\partial\bar{\omega}_2 = -(F_1/F_2)(\partial\mathcal{L}/\partial\bar{\omega}_1)$. Substituting this relationship into (17) and combining the eqs. in (17) and (18), we get

$$\frac{F_1}{F_2} = \frac{\gamma_1 - (\partial\mathcal{L}/\partial\bar{\omega}_1)(1-t)A_1}{\gamma_2 - (\partial\mathcal{L}/\partial\bar{\omega}_1)(1-t)A_2} \qquad \text{and} \qquad \frac{H_1}{H_2} = \frac{\gamma_1}{\gamma_2} \tag{34}$$

where $A_i := (F_{1i} - (F_1/F_2)F_{2i})$, $i = 1, 2$. Note that $A_i = F_2(F_{1i}F_2 - F_1F_{2i})(1/F_2^2) = F_2 f_i$ and that $t < 1$. Hence, the finding $f_1 < 0 < f_2$ implies $(1-t)A_1 < 0 < (1-t)A_2$. Using this result the eq. (19) follows immediately from (34).

A5: Proof of Proposition 2

Because of the eqs. (19) and (15), the claim $F_1/F_2 \neq H_1/H_2$ is correct if and only if the partial derivatives

$$\frac{\partial\mathcal{L}}{\partial\bar{\omega}_i} = \Psi_i\frac{\partial W_i}{\partial\bar{\omega}_i} - \lambda_F\frac{\partial X^i}{\partial\bar{\omega}_i} + \gamma_i\frac{\partial L^i}{\partial\bar{\omega}_i}, \quad i = 1, 2$$

are not equal to zero. Therefore, it has to be shown that the inequality $\partial\mathcal{L}/\partial\bar{\omega}_1 \neq 0$ holds in the optimum (14) provided that this allocation is not first best as well. For a proof of this claim, assume $\partial\mathcal{L}/\partial\bar{\omega}_1 = \partial\mathcal{L}/\partial\bar{\omega}_2 = 0$. Using the eqs. (17) and the relationship $F_i = \omega_i$, $i = 1, 2$, this implies $\lambda_F = \gamma_i/\omega_i$, $i = 1, 2$. Hence,

$$\Psi_i\frac{\partial W_i}{\partial\bar{\omega}_i} - \frac{\gamma_i}{\omega_i}\left(\frac{\partial X^i}{\partial\bar{\omega}_i} - \omega_i\frac{\partial L^i}{\partial\bar{\omega}_i}\right) = 0, \quad i = 1, 2.$$

Note that the individuals' budget constraints imply $\partial X^i/\partial\bar{\omega}_i = L_i + \bar{\omega}_i(\partial L^i/\partial\bar{\omega}_i) = L_i + (1-t)\omega_i(\partial L^i/\partial\bar{\omega}_i)$. We thus have

$$\Psi_i\frac{\partial W_i}{\partial\bar{\omega}_i} = \frac{\gamma_i}{\omega_i}\left(L_i - t\omega_i\frac{\partial L^i}{\partial\bar{\omega}_i}\right), \quad i = 1, 2. \tag{35}$$

Since labor supply L_i is the negative of the excess demand for leisure, Roy's identity and the Slutzky equation lead to $\partial W_i/\partial\bar{\omega}_i = W_I^i L_i$ and $\partial L^i/\partial\bar{\omega}_i = \partial S^i/\partial\bar{\omega}_i + L_I^i L_i$

where S^i is the Hicksian labor supply (because of strict quasiconcavity, we thus have $\partial S^i/\partial\bar{\omega}_i > 0$). Substituting these relationships into (35), we get

$$\Psi_i W_I^i = \frac{\gamma_i}{\omega_i}\left(1 - \frac{t\omega_i}{L_i}\frac{\partial S^i}{\partial\bar{\omega}_i} - t\omega_i L_I^i\right), \quad i = 1, 2. \tag{36}$$

Consider now the eq. (16). Because of $\lambda_F = \gamma_i/\omega_i$, $i = 1, 2$ it leads to

$$\left(\Psi_1 W_I^1 - \frac{\gamma_1}{\omega_1}\left(X_I^1 - \omega_1 L_I^1\right)\right) + \left(\Psi_2 W_I^2 - \frac{\gamma_2}{\omega_2}\left(X_I^2 - \omega_2 L_I^2\right)\right) = 0.$$

Note that the individuals' budget constraints imply $X_I^i = (1-t)\omega_i L_I^i + 1$. Hence,

$$\left(\Psi_1 W_I^1 - \frac{\gamma_1}{\omega_1}\left(1 - t\omega_1 L_I^1\right)\right) + \left(\Psi_2 W_I^2 - \frac{\gamma_2}{\omega_2}\left(1 - t\omega_2 L_I^2\right)\right) = 0. \tag{37}$$

Substituting the eqs. (36) into (37), we get

$$-\frac{\gamma_1}{\omega_1}\left(\frac{t\omega_1}{L_1}\frac{\partial S^1}{\partial\bar{\omega}_1}\right) - \frac{\gamma_2}{\omega_2}\left(\frac{t\omega_2}{L_2}\frac{\partial S^2}{\partial\bar{\omega}_2}\right) = -t\left[\frac{\gamma_1}{L_1}\frac{\partial S^1}{\partial\bar{\omega}_1} + \frac{\gamma_2}{L_2}\frac{\partial S^2}{\partial\bar{\omega}_2}\right] = 0.$$

Since the eqs. (36) imply $\gamma_1 > 0$ and $\gamma_2 > 0$, we thus have $t = 0$. This, however, contradicts the presumption that the allocation (14) is not first best as well.

References

Allen, F. (1982), Optimal linear income taxation with general equilibrium effects on wages, *Journal of Public Economics* 17, 135-143.

Diamond, P. A. and J. A. Mirrlees (1971), Optimal Taxation and Public Production I: Production Efficiency, *American Economic Review* 61, 8-27.

Drèze, J. and N. Stern (1990), Policy Reform, shadow prices, and market prices, *Journal of Public Economics* 42, 1-45.

Eggert, W. and A. Haufler (1999), Capital taxation and production efficiency in an open economy, *Economics Letters* 62, 85-90.

Feehan, J. P. and M. Matsumoto (1999), Optimal Spending on Public Productivity-increasing activities, mimeo, University of Newfoundland.

Guesnerie, R. and J. Seade (1982), Nonlinear Pricing in a Finite Economy, *Journal of Public Economics* 17, 157-179.

Keen, M. and D. Wildasin (2000), Pareto efficiency in international taxation, mimeo, International Monetary Fund.

Mirrlees, J. A. (1986), The Theory of Optimal Taxation, in: Arrow, K.J. and M. D. Intriligator (eds.), *Handbook of Mathematical Economics Vol. III*, Amsterdam: North-Holland, 1197-1249.

Munk, K. J. (1980), Optimal taxation with some non-taxable Commodities, *Review of Economic Studies* 47, 755-765.

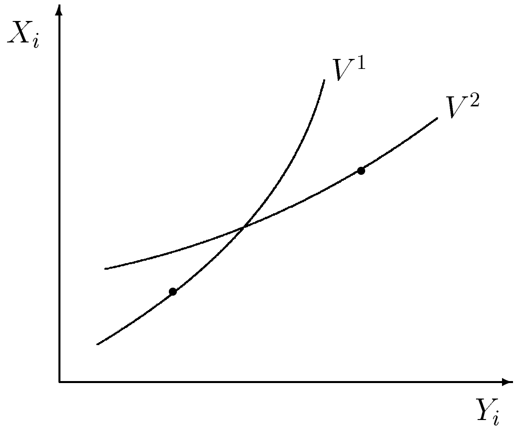
Myles, G. D. (1995) *Public Economics*, Cambridge: Cambridge University Press.

Naito, H. (1999), Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency, *Journal of Public Economics* 71, 165-188.

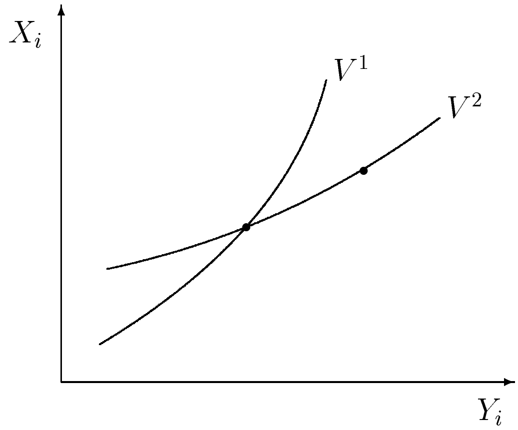
Stern, N. (1982), Optimum Taxation with Errors in Administration, *Journal of Public Economics* 17, 181-211.

Stiglitz, J. E. (1982), Self-Selection and Pareto Efficient Taxation, *Journal of Public Economics* 17, 213-240.

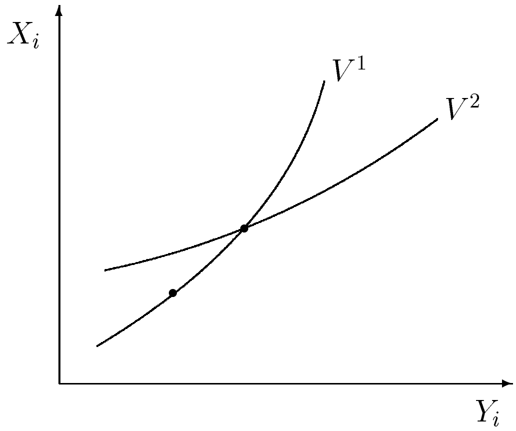
Stiglitz, J.E. and P. Dasgupta (1971), Differential Taxation, Public Goods, and Economic Efficiency, *Review of Economic Studies* 38, 151-174.



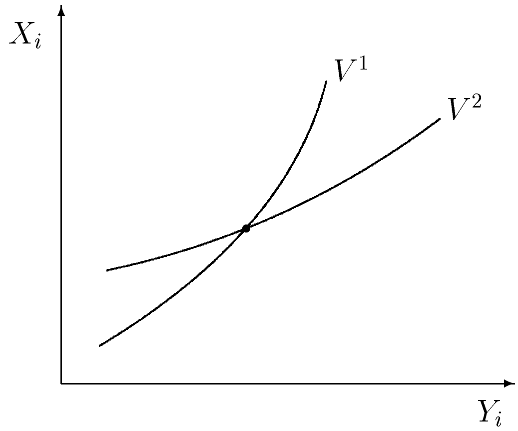
(a) $\omega_2 > \omega_1$, $\lambda_1 = 0$, $\lambda_2 = 0$



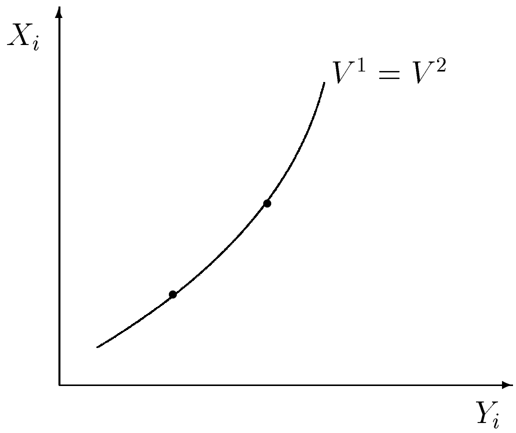
(b) $\omega_2 > \omega_1$, $\lambda_1 = 0$, $\lambda_2 > 0$



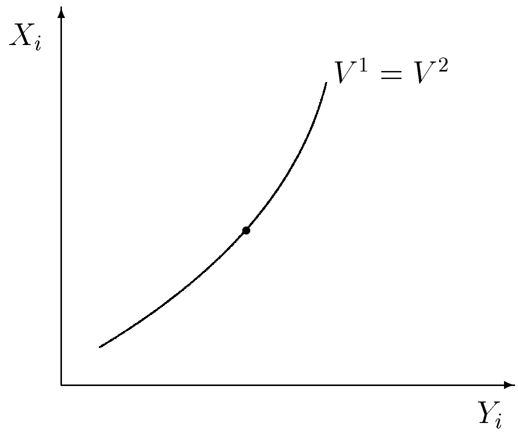
(c) $\omega_2 > \omega_1$, $\lambda_1 > 0$, $\lambda_2 = 0$



(d) $\omega_2 > \omega_1$, $\lambda_1 > 0$, $\lambda_2 > 0$



(e) $\omega_2 = \omega_1$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$



(f) $\omega_2 = \omega_1$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$

Fig. 1