

Direct versus Indirect Taxation: The Design of the Tax Structure Revisited*

Helmuth Cremer

IDEI and GREMAQ, University of Toulouse
and Institut Universitaire de France

Pierre Pestieau

CREPP, University of Liège, CORE and DELTA

Jean-Charles Rochet

IDEI and GREMAQ, University of Toulouse
and Institut Universitaire de France

April 1997, revised February 1999

*We would like to thank Dieter Bös, Christophe Chamley, Jacques Crémer, Jonathan Hamilton and, particularly, Alain Joustien for their comments

Abstract

This paper studies the optimal direct/indirect tax mix in a setting where individuals differ in several unobservable characteristics (productivity and endowments). Tax instruments (income and commodity taxes) are constrained solely by the information structure. It presents general expressions for the optimal commodity tax rates and proves that contrary to Atkinson and Stiglitz's (1976) result, differential commodity taxation remains a useful instrument of optimal tax policy *even if* preferences are separable between labor and produced goods. The following, more specific, results are also derived. First, when cross substitution effects are small the expressions resemble traditional (many households) Ramsey rules. Second, if differences in endowments are confined to some of the goods, the tax rate on goods in which endowments are zero is positively related to income elasticity. Finally, in a Cobb-Douglas illustration with two goods, where endowments differ only in good 1 (and are interpreted as "wealth"), the tax on good 2 provides an indirect way to tax the unobservable wealth. It is higher *(i)* the more significant are the wealth differentials, *(ii)* the stronger is the correlation between wealth and earning ability and *(iii)* the larger are the (political) weights attached to low wealth individuals.

1 Introduction

The choice between direct and indirect taxation is one of the “oldest issues of taxation policy” (Atkinson (1977)). It raises challenging theoretical questions and is of significant policy relevance. The theoretical controversy is mirrored by the diverging solutions which have been adopted in various countries. Cross-country comparisons indeed reveal striking differences in the tax mix. Indirect taxes (VAT in particular) represent an important part of tax revenues in European countries, but they play only an insignificant role in the US (at least at the federal level). Even within the European Union, the degree of reliance on indirect taxes differs notably between member states.¹

Much of the earlier debate was marked by an ambiguity in the very definition of direct and indirect taxation. Originally, the distinction was often based on the notion of “shifting”, with taxes believed to be shifted called *indirect* taxes (Due and Friedlander (1973, p. 229)). A more satisfactory approach has been suggested by Atkinson (1977) (see also Atkinson and Stiglitz (1980)) who argues that “...the essential aspect of the distinction [is] the fact that direct taxes may be adjusted to the individual characteristics of the taxpayer, whereas indirect taxes are levied on transactions irrespective of the circumstance of buyer and seller.” Our modelling of the different tax instruments is closely inspired by Atkinson’s information based distinction. Income taxation is viewed as direct because it can be made progressive, whereas commodity taxation is indirect, in the sense that it is based on anonymous transactions and can only be proportional.

The role of indirect taxes as instruments of optimal tax policy has been severely undermined by the publication of Atkinson and Stiglitz’s (1976) influential paper. They have shown that when preferences are weakly separable in labor supply and produced goods, nonlinear income taxation does not need to be supplemented by commodity

¹According to OECD figures (for 1995), indirect taxes represent on average 31.4% of total tax revenues in the EU (versus 17.9% in the US); their share ranges from about 26% (Belgium, The Netherlands and Sweden) to 44.6% (Portugal).

taxation: any second best optimum can be achieved by income taxation alone.² The practical implications of Atkinson and Stiglitz's result certainly hinge on the empirical validity of their separability assumption, which has been questioned (see e.g., Browning and Meghir (1991)). However, the spirit of their result goes through even under non-separability as long as labor supply has no "significant" effect on individuals' marginal rates of substitution (between goods). A strong case for commodity taxation must thus rest on a substantial impact of labor supply on the willingness to pay for the produced goods and to the best of our knowledge, such a case has not yet been made. Some alternative lines of attack have been explored (e.g., tax evasion in Boadway *et al.* (1994)), and uncertainty in Cremer and Gahvari (1995)), but at this point, neither of them appears to have inflicted a fatal blow to the argument that commodity taxes have no (or at best a minor) role.

The most fundamental shortcoming of Atkinson and Stiglitz's analysis appears, however, to be of a completely different nature. They assume (like most of the literature on non linear income taxation) that individuals differ in one single characteristic, namely wage (earning ability). There are no taste differences nor any other sources of heterogeneity (like wealth differentials). This assumption appears to be motivated by technical considerations rather than by economical or empirical arguments. There is certainly no reason to believe that individuals are alike in all respects but their earning ability; however, so far, this assumption has been the price to pay to solve the optimal taxation problem. Multidimensional heterogeneity (adverse selection) would have made the problem quite untractable. As a matter of fact, the difficulty is not so much to show that Atkinson and Stiglitz's result does not in general hold in such a setting. This can easily be shown without solving the overall problem and Mirrlees (1976) has already made this point (for the case of taste differences). The challenging problem is to provide

²Stiglitz (1988) forcefully points out the policy implications of this result and writes: "It can be shown that if one has a well-designed income tax, adding differential commodity tax is likely to add little, if anything, to the ability to redistribute income" (p. 494).

a characterization of optimal commodity taxes.

In this paper we consider a setting where individuals differ not only in earning abilities but also in endowments. Intuitively, such a difference can be thought of as reflecting wealth inequality. In the general model, endowments are allowed to differ for all goods. However, the analogy with wealth is most compelling in the special case when endowments are positive and different for only one good.³ Consequently, this case is given special emphasis. This interpretation in terms of wealth inequality is clearly only a static shortcut for a problem which is in reality dynamic. Furthermore, wealth differentials are, at least in part, endogenous and result from individual accumulation decisions. On the other hand, the evidence shows that a significant fraction of wealth inequalities are explained by differences in inherited wealth (see e.g., Arrondel *et al.*(1997)). As far as this part of individuals' wealth is concerned, the analogy with endowments appears quite compelling.

Following Atkinson and Stiglitz's approach, we do not impose any *ad hoc* restrictions on the class of available tax instruments. These are constrained solely by the information structure. We assume that individual types (characterized by earning abilities and endowments) are not publicly observable. This rules out first-best taxation of types. Before-tax (labor) incomes on the other hand are observable so that a non-linear income tax can be imposed. As for the produced goods, neither *personal* consumption levels, nor *personal* net transactions (consumption minus endowment) are publicly observable. The tax administration has, however, information on anonymous transactions. Consequently, non-linear commodity taxes are not feasible, while linear (proportional) commodity taxes are available.⁴ Throughout the paper we shall concentrate on the case

³This is of course in addition to the individuals' endowments in time.

⁴Atkinson and Stiglitz do allow for non-linear commodity taxes (they assume that individual consumption levels of all goods are observable). Though not realistic, this assumption is perfectly appropriate from their perspective. It can only strengthen their case: when non-linear taxes are not needed, linear taxes certainly cannot do any better. Our purpose is, however, exactly the opposite of theirs: we want to document the usefulness of commodity taxes. This case is strengthened by imposing the most stringent restrictions on commodity taxes. Furthermore, we want to provide a characterization of

of separable preferences, for this allows us to keep in line with the spirit of Atkinson and Stiglitz's result.⁵

Within this setting we characterize the optimal tax mix and study under which conditions it includes (non-uniform) commodity taxes. We show that commodity taxes are needed, except when all pairs of types which are linked by a binding incentive constraint have identical endowments. This condition is satisfied in a trivial way when *all* have identical endowments, but then one essentially returns to Atkinson and Stiglitz's setting. Otherwise, the condition is quite stringent and not likely to be satisfied in economically meaningful cases.

Next, we present general expressions for the optimal commodity tax rates. These expressions include efficiency terms (compensated demand elasticities) reflecting dead-weight losses and incentive terms reflecting the impact of commodity taxes on self-selection constraints (and thus on the feasible extent of redistribution). Incentive effects are measured by weighted sums of incremental *net* demands (consumption minus endowment) of the mimicker (compared to the mimicked individual). Interestingly, it thus appears that the benefits of commodity taxes are of redistributive nature which contradicts the traditional view that commodity taxes tend to be regressive and can be justified (if at all) only by efficiency considerations.

When cross substitution effects are small (so that the "reduced" Slutsky matrix is almost diagonal), the expressions are considerably simplified. Optimal tax rates are then (roughly speaking) proportional to inverse compensated price elasticities and to (the weighted average of) incremental net consumption levels of mimickers (for the considered good). These expressions have a familiar flavor and resemble traditional (many households) Ramsey rules. However, redistributive benefits are no longer measured by

optimal commodity taxes and such an exercise is meaningful only if it rests on a realistic information structure. Put differently, there is no point in characterizing hypothetical optimal non-linear taxes when the information required to impose them is not available.

⁵To show the usefulness of commodity taxation this assumption does not involve any loss of generality. It could be restrictive, though, as far as the characterization of optimal commodity taxes is concerned.

differences in effective consumption levels but depend on the hypothetical consumption levels of mimickers. This is because there is also a (non-linear) income tax, which implies that commodity taxes enhance redistribution only if they relax otherwise binding self-selection constraints.

If differences in endowments are confined to some of the goods, our results imply particularly simple tax rules for the goods in which endowments are zero (or identical). For these goods, the tax rate is positively related to income elasticity. In particular, normal goods ought to be taxed, while inferior goods should be subsidized.

Finally, we consider a Cobb-Douglas illustration with two goods, where endowments differ only in one of the goods, say good 1 (with endowments interpreted as “wealth” or non-labor income). In this case, the redistributive role of commodity taxation becomes even more evident. A tax on the produced good 2 now provides an indirect way to tax the unobservable wealth. This point is reinforced by the numerical calculations which show that the tax (on good 2) is higher *(i)* the more significant are the wealth differentials, *(ii)* the stronger is the correlation between wealth and earning ability and *(iii)* the larger are the (political) weights attached to low wealth individuals.

2 The model

It is a simple extension of the standard income taxation model (Mirrlees (1971), Atkinson and Stiglitz (1976)). There are N types of individuals ($i = 1, \dots, N$) who differ in their labor productivities n^i and their initial endowments ω^i in m consumption goods. The proportion of types i in the population is denoted π^i . All individuals have the same strictly quasi-concave utility function: $u(C^i) - v(L^i)$, where C^i is the consumption vector and L^i denotes labor supply. Separability is assumed for the sake of simplicity but also to keep in line with the spirit of the Atkinson and Stiglitz result. Technologies are linear: with one unit of time, individual i produces n^i units of “elementary” labor; one unit of “elementary” labor, in turn, produces one unit of any of the m commodities.

Normalize the producer price of commodity 1 at one. All markets being competitive, the wage rate of individual i equals n^i and the producer prices of all commodities equal one.

In the tradition of the optimal income taxation literature, we assume that an individual's type (characterized by n^i and ω^i) and labor supply are not observable by the tax administration. His before-tax (labor) income, $I^i = n^i L^i$, on the other hand, is. This rules out first-best taxation of types, while allowing non-linear taxation of incomes. Furthermore, neither personal consumption levels, C^i , nor personal net transactions $Z^i = C^i - \omega^i$ are publicly observable. The tax administration has, however, information on anonymous transactions. Under this circumstance, non-linear commodity taxes are not feasible, while linear commodity taxes are available. This is the standard (though often only implicit) assumption in the literature, so much that it has been used as part of the very definition of indirect taxes; see Section 1.

To sum up, the tax policy consists of a non-linear tax $T(\cdot)$ on labor income and of t , the m -vector of per-unit commodity taxes which determine $p = 1 + t$, the m -vector of consumer prices. The problem of individual i is then given by:

$$\begin{aligned} \max \quad & u(C^i) - v(L^i) \\ \text{s.t.} \quad & p \cdot (C^i - w^i) = I^i - T(I^i) \end{aligned} \tag{1}$$

$$I^i = n^i L^i. \tag{2}$$

Note that all vectors are column matrices. Transposed vectors are denoted by the superscript T . The symbol “ \cdot ” is used for the scalar product; for instance $p \cdot C^i = p^T C^i$. Because of the separability of preferences, the problem can be also written as

$$\begin{aligned} \max \quad & V(p, R^i) - v\left(\frac{I^i}{n^i}\right) \\ \text{s.t.} \quad & R^i = I^i - T(I^i) + p \cdot w^i, \end{aligned} \tag{3}$$

where R^i denotes the disposable income of individual i (after tax labor income plus

market value of endowments), and $V(p, R)$ is the indirect utility function associated with $u(\cdot)$:

$$V(p, R) = \max \{u(C), p \cdot C = R\}. \quad (4)$$

The m vector of demand functions (solutions of (4)) is denoted $C(p, R)$, while $\tilde{C}(p, u)$ denotes the vector of compensated demand functions.⁶ For future reference note that the traditional Slutsky decomposition, as well as Roy's identity apply (for all $k = 1, \dots, m$):

$$\frac{\partial \tilde{C}}{\partial p_k}(p, V(p, R)) = \frac{\partial C}{\partial p_k}(p, R) + \frac{\partial C}{\partial R}(p, R)C_k(p, R), \quad (5)$$

$$\frac{\partial V}{\partial p_k}(p, R) + C_k(p, R)\frac{\partial V}{\partial R}(p, R) = 0. \quad (6)$$

Since the distribution of types is discrete ($i = 1, \dots, N$) the income tax schedule can be summarized by the N points (I^i, T^i) representing the individuals' respective choices. Because types are private information, incentive compatibility constraints require that the utility obtained by each individual i when he consumes (I^i, T^i) :

$$U^i = V(p, R^i) - v\left(\frac{I^i}{n^i}\right), \quad (7)$$

be at least equal to the utility level he would obtain by choosing (I^h, T^h) ("mimicking" individual h), that is:

$$U^{ih} = V(p, R^{ih}) - v\left(\frac{I^h}{n^i}\right), \quad (8)$$

where

$$R^{ih} = R^h + p(\omega^i - \omega^h).$$

We can now move on to the government's problem. It aims at maximizing a weighted sum of utilities, subject to a revenue constraint and to the incentive compatibility constraints. The weight of types i is denoted $\alpha^i \pi^i$ (with $\alpha^i \geq 0$, and $\sum_j \alpha^j = 1$)⁷. Formally, this problem can be stated as follows:

⁶Observe that these *functions* are the same for all individuals. They are independent of n^i because preferences are separable and ω^i only enters through R^i .

⁷This is useful for separating the effect of varying the distribution of types and varying the distribution of political weights. In particular the utilitarian solution corresponds to $\alpha^i \equiv 1/N$ (uniform weights).

Problem \mathcal{P}

$$\begin{aligned} \max_{(R^i, I^i)_i, p_2, \dots, p_m} \quad & \sum_{i=1}^N \alpha^i \pi^i U^i \\ \text{s.t.} \quad & \sum_{i=1}^N \pi^i \left(I^i - R^i + (p - \mathbb{I}) \cdot C(p, R^i) + \mathbb{I} \cdot \omega^i \right) \geq G \end{aligned} \quad (9)$$

$$U^i \geq U^{ih} \quad i, h = 1, \dots, N \quad (10)$$

where \mathbb{I} is the unit vector, G is the (exogenous) revenue requirement, while U^i is given by (7) and U^{ih} by (8). Observe that with one extra degree of freedom in setting commodity tax rates, t_1 is set equal to zero so that $p_1 = 1$.⁸ Finally, $(p - \mathbb{I}) = (0, p_2 - 1, \dots, p_m - 1)^T = t$ is the vector of per unit taxes.

3 Characterizing the optimal tax mix

We now derive the first order conditions of the government's problem. Denoting the multipliers of constraints (9) and (10) by λ and μ^{ih} respectively, one can write the Lagrangian expression associated with \mathcal{P} as follows:

$$\begin{aligned} \Lambda = & \sum_{i=1}^N \alpha^i \pi^i U^i + \lambda \left(\sum_{i=1}^N \pi^i \left(I^i - R^i + (p - \mathbb{I}) \cdot C(p, R^i) \right) + \mathbb{I} \cdot \omega^i - G \right) \\ & + \sum_{i,h=1}^N \mu^{ih} (U^i - U^{ih}), \end{aligned} \quad (11)$$

The first order conditions with respect to R^i ($i = 1, \dots, n$) and p_k ($k = 2, \dots, m$)

⁸To see this most easily, notice that the revenue constraint of the government could also be written as:

$$\sum_{i=1}^N \pi^i \left(I^i - \sum_{k=1}^m C_k(p, R^i) + \sum_{k=1}^m \omega_k^i \right) \geq G.$$

This expression is homogenous of degree zero in p and R^i 's and this property clearly also applies to the objective function and incentive constraints.

are given by:⁹

$$\begin{aligned}\frac{\partial \Lambda}{\partial R^i} &= \left(\alpha^i \pi^i + \sum_h \mu^{ih} \right) \frac{\partial V}{\partial R}(p, R^i) - \lambda \pi^i \left[1 - (p - \mathbb{I}) \frac{\partial C}{\partial R}(p, R^i) \right] \\ &\quad - \sum_h \mu^{hi} \frac{\partial V}{\partial R}(p, R^{hi}) = 0,\end{aligned}\tag{12}$$

$$\begin{aligned}\frac{\partial \Lambda}{\partial p_k} &= \sum_i \left(\alpha^i \pi^i + \sum_h \mu^{ih} \right) \frac{\partial V}{\partial p_k}(p, R^i) + \lambda \sum_i \pi^i \left(C_k(p, R^i) \right. \\ &\quad \left. + (p - \mathbb{I}) \cdot \frac{\partial C}{\partial p_k}(p, R^i) \right) - \sum_{h,i} \mu^{hi} \frac{\partial V}{\partial p_k}(p, R^{hi}) \\ &\quad - \sum_{h,i} \mu^{hi} \frac{\partial V}{\partial R}(p, R^{hi}) (\omega_k^h - \omega_k^i) = 0.\end{aligned}\tag{13}$$

This last expression can be simplified by introducing (for all $k = 2, \dots, m$) the compensated derivative of Λ with respect to p_k , defined by:

$$\frac{\partial \tilde{\Lambda}}{\partial p_k} \equiv \frac{\partial \Lambda}{\partial p_k} + \sum_i \frac{\partial \Lambda}{\partial R^i} \cdot C_k(p, R^i).\tag{14}$$

Substituting (12) and (13) into (14), using (5) and (6), and rearranging yields:

$$\begin{aligned}\frac{\partial \tilde{\Lambda}}{\partial p_k} &= (p - \mathbb{I}) \cdot \left(\lambda \sum_i \pi^i \frac{\partial \tilde{C}}{\partial p_k}(p, V(p, R^i)) \right) \\ &\quad + \sum_{h,i} \mu^{hi} \frac{\partial V}{\partial R}(p, R^{hi}) \left((C_k(p, R^{hi}) - \omega_k^h) - (C_k(p, R^i) - \omega_k^i) \right).\end{aligned}\tag{15}$$

At the optimum, we must have $\partial \tilde{\Lambda} / \partial p_k = 0$, for $k = 2, \dots, m$. (Recall that p_1 is conventionally set to 1). Consequently, we obtain a system of $(m - 1)$ equations, represented synthetically by:¹⁰

$$- \left(\lambda \sum_i \pi^i \frac{\partial \tilde{C}}{\partial p}(p, V(p, R^i)) \right) (p - \mathbb{I}) = \sum_{h,i} \mu^{hi} \frac{\partial V}{\partial R}(p, R^{hi}) (Z^{hi} - Z^i),\tag{16}$$

⁹The first-order conditions with respect to I^i are not directly relevant for our purposes.

¹⁰From now on, we redefine p as the $(m-1)$ vector $(p_2, \dots, p_m)^T$. Similarly, $\partial \tilde{C} / \partial p$ is a $(m-1) \times (m-1)$ matrix, namely the Slutsky matrix from which the first line and the first column are removed; see (17).

where, Z^i and Z^{hi} respectively denote the net consumption vectors (consumption minus endowment) of individual i and of individual h mimicking individual i respectively.¹¹

The left-hand side of system (16) represents the marginal dead-weight losses associated to the distortions of consumption prices. It vanishes when $p = \mathbb{I}$ (no commodity taxation) and is proportional to the image of the tax vector $(p - \mathbb{I})$ by the (aggregate) “reduced” Slutsky substitution matrix, a $(m - 1) \times (m - 1)$ matrix defined by:

$$S = (\sigma_{jk})_{j,k=2,\dots,m}$$

with

$$\sigma_{jk} = \sum_i \pi^i \frac{\partial \tilde{C}_j}{\partial p_k}(p, V(p, R^i)). \quad (17)$$

The right-hand side of (16) figures the marginal benefits (in terms of “rent extraction”) of distorting consumption prices.

Observe that with u strictly quasi concave, the matrix S is a convex combination of negative definite matrices. Therefore, it is itself negative definite, and thus invertible. Rearranging (16) then yields the following proposition.

Proposition 1 *The optimal commodity taxes are characterized by*

$$t = p - \mathbb{I} = -S^{-1} \left(\sum_{h,i} \frac{\mu^{hi}}{\lambda} \frac{\partial V}{\partial R}(p, R^{hi}) \cdot \Delta^{hi} Z \right), \quad (18)$$

where $\Delta^{hi} Z$ denotes $Z^{hi} - Z^i = [C(p, R^{hi}) - \omega^h] - [C(p, R^i) - \omega^i]$ i.e., the incremental net demands of the mimicker.

Let us first analyze the conditions under which the right hand side of (18) vanishes. In that case, one indeed returns to the Atkinson and Stiglitz result: optimal commodity taxes are zero and the tax policy solely relies on income taxation. This result emerges when, for all binding incentive compatibility constraints (represented by the couples

¹¹Formally, Z^i is the vector of Z_k^i , $k = 2, \dots, m$, where $Z_k^i \equiv C_k(p, R^i) - \omega_k^i$. Similarly, Z^{hi} is the vector of Z_k^{hi} , $k = 2, \dots, m$ with $Z_k^{hi} \equiv C_k(p, R^{hi}) - \omega_k^h$.

(h, i) for which $\mu^{hi} > 0$), the net consumption bundle of the “mimicker” Z^{hi} equals the net consumption bundle of the “mimicked” Z^i . Since $R^{hi} = R^i + p(w^h - w^i)$, this happens essentially only when mimicker and mimicked have the same initial endowments *for all goods*. Thus we obtain a simple extension of the Atkinson and Stiglitz result:

Proposition 2 *Suppose that, at the solution of (\mathcal{P}) , $\mu^{hi}(w_k^h - w_k^i) = 0$ for all $h, i = 1, \dots, N$ and $k = 1, \dots, m$. Then the vector of optimal commodity tax rates is zero: $p^* = \mathbb{I}$.*

The condition in proposition 2 is quite stringent. It holds in a trivial way, when *all* types have identical endowment vectors. This does not come as a surprise for with identical endowments one essentially returns to the original Atkinson and Stiglitz setting. However, the condition also holds in a slightly more general setting: it is sufficient that all pairs of types which are linked by a binding incentive constraint have identical endowments. For instance, if all types except for, say, j have identical endowments the optimum implies zero commodity taxes if no incentive constraint involving type j (either from j or towards j) is binding. Though not impossible, such an outcome appears to be more of a technical curiosity than an economically meaningful case.¹²

Except under these conditions, the compensated derivative of Λ with respect to p does not, in general, vanish at $p = \mathbb{I}$. Consequently, the Atkinson and Stiglitz result fails to apply (as a general property) when endowments differ between individuals.

Having shown that commodity taxation is (generally) needed, we shall now discuss its optimal structure. This calls for a more thorough analysis of (18).

When cross substitution effects are small (so that S^{-1} is almost diagonal), expression (18) implies that optimal tax rates are (roughly speaking) proportional to inverse (own compensated price) elasticities and to (the weighted average of) incremental net

¹²The only case with some economic significance we can think of is when the weights (in the objective function) are such that the solution coincides with, or is in the neighborhood of the competitive equilibrium (which, as is well known, is incentive compatible). In that case, however, there is no *distortionary* taxation whatsoever, neither on income nor on commodities.

consumption bundles of mimickers. The “inverse elasticity” part of this rule is rather standard. Everything else equal, the goods which ought to be taxed most heavily are those for which compensated demand is least elastic (for they imply a smaller distortion).

The “incremental net demand” term, on the other hand, reflects the welfare impact of commodity taxes via their effects on the incentive constraints. Its sign determines the sign of the optimal tax on the considered good. For instance, if every mimicker for whom the incentive constraint binds has a higher net demand for good k than the type whom he mimics (i.e., $Z_k^{hi} - Z_k^i > 0$ for each pair h, i such that $\mu^{hi} > 0$) then good k should face a positive tax. This is quite intuitive: a tax on k “hurts” the mimicker more than the mimicked and thus allows one to relax an otherwise binding incentive constraint.¹³ Not surprisingly the opposite condition (i.e., $Z_k^{hi} - Z_k^i < 0$ for each pair h, i such that $\mu^{hi} > 0$) yields a negative tax (i.e., a subsidy) for good k . Finally if the sign of $Z_k^{hi} - Z_k^i$ is not the same for all relevant pairs, the weighted average of these terms (with weights proportional to the shadow price of the corresponding incentive constraint) determines the sign of the expression and thus of the optimal tax.

This argument can be further refined for the case where individuals have identical (possibly zero) endowments in good k (while endowments differ in some other good(s)). Concentrating on the “normal case” where redistribution occurs from the rich (with higher consumption budgets) to the poor we then obtain a simple result according to which normal goods ought to be taxed while inferior goods should be subsidized.¹⁴ Put differently, the tax on good k tends to be higher, the higher its income elasticity. Recall that the compensated price elasticity also matters and acts in the opposite direction.

¹³Net demands can of course be negative, but this does not affect the argument. For instance if net demands are negative for both mimicker and mimicked, the above condition implies that the mimicker “benefits” less from the tax than the mimicked.

¹⁴With uniform endowments (for the considered good), differences in net demands equal differences in consumption levels. Now, if all the mimickers have higher disposable incomes than the mimicked, they will have larger consumption levels of the normal goods.

This is quite in line with the results obtained in standard optimal tax models (e.g., a many households Ramsey model combined with a linear income tax; see Atkinson and Stiglitz (1980, Section 14-2)).

Such a simple rule, however, cannot be obtained if endowments in k differ. Mimickers may continue to consume more of the normal goods than the mimicked; however, their *net* demand is not necessarily higher. In particular, they may be rich precisely because they have a large endowment in good k .

Recall that all this applies only if cross-substitution effects are “small”. For the general case, not much can be said without further restrictions. However, (16) can still be interpreted along the lines familiar in Ramsey type models. Because of the symmetry of the substitution matrix, the LHS is then essentially (a linear approximation of) the “reduction in compensated demand” induced by the tax system.¹⁵ From the interpretation of the RHS it then follows that this reduction ought to be larger for the goods bought more heavily by the mimickers. For the goods with zero (or identical) endowments, we obtain (in the “normal case”) that the reduction in compensated demand ought to be larger for normal than for inferior goods.

Before proceeding, we should point out that these interpretations have to be considered with great care. The link between income elasticity and taxation rests on the “normal case” assumption which corresponds to the intuitively appealing idea that incentive constraints bind along the direction of decreasing consumption budgets. This assumption is quite meaningful in the traditional one-dimensional characteristics case where it arises for a large class of welfare functions (weights).¹⁶ At first, one might be tempted to think that such an intuitive property ought to carry over to the two-dimensional case.¹⁷ Under closer scrutiny, however, such a conjecture turns out to be

¹⁵When the tax is negative, the “reduction” become an increase, but the argument does not change.

¹⁶Including a utilitarian objective. The opposite result arises only if the rich receive a higher (and sufficiently significant) weight than the poor.

¹⁷And admittedly, this was certainly our first guess.

misleading. The examples considered below will make it clear that what is “normal” in the one-dimensional case may be rather “special” in a multi-dimensional setting. Consequently, we cannot claim that the results discussed above have some general relevance (even if one is prepared to accept the underlying assumptions on the preference structure).

At this level of generality, any prospects for a more precise interpretation of (16) and/or (18) appear to be limited. Any results rest on the pattern of binding incentive constraints, and this cannot be predicted on the basis of the first-order conditions alone. Consequently, we shall now turn to the study of a more specific example. Though not meant to be realistic, these illustrations will allow us to reach a better understanding of the qualitative properties of the optimal tax structure.

The illustration we adopt is based on a Cobb-Douglas specification. This choice is motivated by the fact that we want to make a case for commodity taxation and Cobb-Douglas preferences certainly are the least likely to support such a claim. If types differ only in one dimension (productivity), Cobb-Douglas preferences consistently imply uniform commodity taxes. Furthermore, even a linear income tax is then sufficient to make commodity taxes a redundant instrument. This is essentially because, with this preference structure, all goods have the same income elasticity (which moreover equals one). Consequently, it is plain that whenever Cobb-Douglas preferences require (non-uniform) commodity taxes the adoption of a more general preference structure can only reinforce this property.

The simple specification we use has the following features. There are only two produced goods ($m = 2$) and initial endowments consist only of commodity one: $\omega_2^i \equiv 0$. Setting the consumer price of commodity one at one, we can therefore drop the index k and adopt the following simplified notation: $\omega^i = \omega_1^i$, $p_2 = p$. Observe that within this setting an individual’s endowment can be interpreted as his “wealth” (or non-labor

income). The utility function is given by: $u(C) = C_1^{1/4} C_2^{1/4}$, so that:¹⁸

$$C_2(p, R) = \frac{R}{2p} \quad \text{and} \quad \frac{\partial \tilde{C}_2}{\partial p} = -\frac{R}{4p^2}. \quad (19)$$

Expression (18) then becomes:

$$\frac{p-1}{p} = \frac{2}{\lambda} \cdot \frac{\sum_{h,i} \mu^{hi} \frac{\partial V}{\partial R}(p, R^{hi})(\omega^h - \omega^i)}{(\sum_i \pi^i R^i)}. \quad (20)$$

In words, the optimal tax rate (on good 2) has the same sign as a weighted sum of the difference between the mimickers' endowment ("wealth"), ω^h , and the mimicked's endowment ("wealth"), ω^i . Intuitively this makes a lot of sense. Individuals have no endowments in good 2, which has an income elasticity of one. Consequently, wealthier individuals necessarily consume (and buy) a larger quantity of that good. Taxing this good is thus desirable if the incentive constraints (or at least those with the highest shadow prices) bind from high wealth to low wealth individuals. In that case, the commodity tax can be seen as an indirect way to tax the unobservable endowment. Unlike a direct taxation of the endowments, it has of course the drawback that it creates distortions. Put differently, commodity taxation appears as an imperfect (second-best) substitute for a (first-best) lump sum tax based on endowments. Notice that if incentive constraints bind from low to high wealth individuals, one gets exactly the opposite argument. In both cases, however, the tax on good 2 will not in general be equal to zero.

The crucial question which remains open is to know whether incentive constraints effectively bind from high wealth to low wealth individuals. Clearly, this will not necessarily be the case. One can expect this to depend on the relative inequality of endowments (as compared to productivity differentials) as well as on the correlation between endowments and productivity. In addition, the weights in the welfare function may be

¹⁸Equal weights are used to avoid additional *a priori* asymmetries between the goods. This gives uniform commodity taxes the best chances to emerge.

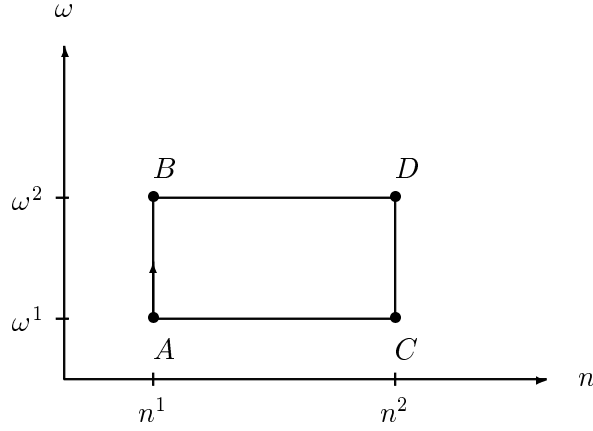


Figure 1: The distribution of types

of crucial importance. These issues will be addressed in the simulations to which we now turn.

4 The determinants of the optimal tax mix: some illustrations

Our simulations are based on the following specification:

- 2 commodities ($m = 2$), endowments only in good 1: $(w^i, 0)$
- 4 types (denoted by $i = A, B, C, D$) distributed on a rectangle (as depicted on Figure 1) with $n^1 = 5$, $n^2 = 10$ and $\omega^1 \leq \omega^2$. Binding incentive constraints will be represented by arrows (e.g., AB on Figure 1)
- Cobb-Douglas utility for consumption:

$$u(C_1, C_2) = C_1^{1/4} C_2^{1/4}$$

- quadratic disutility of labor:

$$v(L) = L^2$$

Types i	A	B	C	D
Incomes I^i	7.37	6.70	23.73	20.47
Taxes T^i	1.92	1.99	7.10	6.87
Disposable income R^i	7.45	14.71	18.63	23.57
MRS ^{i}	0.82	1.05	1.04	1.01
Utilities U^i	2.14	3.32	2.84	3.73

Table 1: Central variant: results

- revenue requirement:

$$G = 5.$$

For the remaining parameters we consider several alternative values. We start with the following “central” specification:

- endowments:

$$\omega^1 = 2 \quad \omega^2 = 10$$

- uniform distribution of types:

$$\pi^i \equiv \frac{1}{4}$$

- uniform weights:

$$\alpha^i \equiv \frac{1}{4}.$$

The numerical solution of this central specification is given in the next subsection.

4.1 The Solution of Central Variant

The optimal consumer price (of commodity 2) is given by $p = 1.07$; this corresponds to a tax rate of 7%. The values of the remaining relevant variables are summarized in Table 1. In addition to labor incomes I^i , taxes T^i and disposable incomes R^i (defined by (3)) we have also reported utility levels U^i as well as marginal rates of substitution

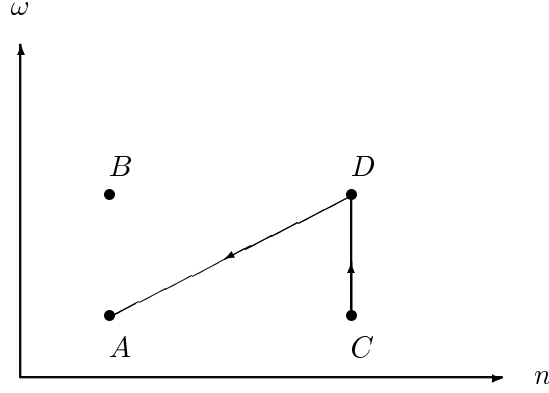


Figure 2: Central variant: binding incentive constraints

(MRS^i) between disposable income and labor income.¹⁹ This variable characterizes the marginal tax rate implied by the implementing income tax schedule. Specifically, one has $MRS^i = 1 - T'(I^i)$, where $T'(I^i)$ is the marginal income tax rate faced by type i . Observe that $MRS^i < 1$ (resp. > 1) corresponds to a positive (resp. negative) marginal income tax rate on type i .²⁰ Only two incentive compatibility constraints are binding, namely CD and DA ; see Figure 2:

This solution presents a number of interesting features. Let us first consider the pattern of binding incentive constraints. Based on the intuitions obtained in the one dimensional case, one would expect these constraints to bind (in a more or less monotonic way) along the direction of decreasing disposable incomes.²¹ The solution shows that

¹⁹Formally it is defined as follows:

$$MRS^i = \frac{\frac{dv}{dL}(\frac{I^i}{n^i})}{n^i \frac{\partial V}{\partial R}(p, R^i)}.$$

²⁰See Stiglitz (1987) for a detailed discussion of the implementation in a discrete types setting.

²¹With uniform weights (and separability) first best optimality calls for an equalization of disposable incomes. Consequently, it is tempting to think that under asymmetric information one attempts to reduce inequalities in disposable income as much as possible given the incentive constraints.

this is not the case. The constraint CD binds, even though D has higher disposable income than C . The other binding constraint, DA , goes in the expected direction, but it is not “local”, in the sense that it links the two extreme types; recall that D has both high productivity and high wealth, while A has a low productivity and a low wealth. Furthermore, A and B differ in disposable incomes but neither BA nor AB binds.

To explain this surprising pattern, one has to realize that as far as pairs of types are concerned, the socially most beneficial redistribution is precisely from D to A (most significant difference in marginal utilities of income). Accordingly, differences in disposable incomes between these two types are reduced until the incentive constraint becomes binding. Once this constraint is binding, the opportunities of redistribution between other pairs of types become limited. For instance, an attempt to redistribute from B to A would interfere with this constraint (making A ’s consumption bundle more attractive for D). Similarly, while it is in principle desirable to transfer money from C to, say, B , this is also limited by the fact that D has to be granted some informational rent; if C were taxed “too heavily” he might then want to mimic type D .

Let us now turn to the optimal commodity tax (good 2) which is positive in this case. Recall the discussion of (20): what matters is the weighted sum of endowment differentials between mimicker and mimicked individual. Here, this sum has two terms, one associated with each binding incentive constraint. In the case of CD , the mimicker has a *lower* wealth than the mimicked type, while the opposite is true for DA . Consequently, the two terms are of opposite sign and the sign of the optimal tax depends on their relative weights. Now, the positive tax indicates that DA receives a higher weight, which considering the above discussion does not come as a surprise.

All this is of course heavily dependent on the distribution of types (support and correlation) as well as on the weights. In the following subsections, we shall study the impact of each of these factors separately.

ω^1	ω^2	Binding IC constraints	p
0	12	BA, CD, DA	1.14
2	10	CD, DA	1.07
3	9	AB, CD, DA	1.05
4	8	AB, BA, CD, DA, DB	1.02
5	7	AB, BA, CD, DA, DB	1.01
6	6	CA	1.00

Table 2: Impact of wealth inequality

4.2 The Impact of Wealth Inequality

We shall now study the impact of varying the difference in endowments ($\omega^2 - \omega^1$), while keeping total endowments $\omega^2 + \omega^1$ constant. All other parameters are set at their central variant values. The results are summarized in Table 2. To keep the presentation intelligible, we concentrate on the most relevant results for each case, namely the pattern of binding incentive constraints and the optimal consumer price of good 2.²²

It thus appears that for a given level of total endowments, the optimal commodity tax on good 2 increases with wealth (endowment) inequality. This result is not surprising and it reinforces the point that the tax on good 2 is used as an indirect way to tax the unobservable endowment in good 1 (see Section 3). As expected, the optimal tax is zero when all types have identical endowments (Proposition 2).²³ When wealth differentials are small, the tax is positive, but “small”. Types who share the same productivity level are then “close” and the redistributive potential of the commodity tax is limited. Technically, this is reflected by the fact that there is bunching between A and B .²⁴ At the other extreme, when wealth differentials are larger than in the central variant case (line 1), an additional incentive constraint from high to low wealth becomes binding

²²More detailed results are available from the authors upon request.

²³There are then effectively only two types $A = B$ and $C = D$.

²⁴The two types are treated alike: same labor supply (labor income) and same tax. In terms of incentive constraints, this translates into both AB and BA being binding.

π	Correlation	Binding IC constraints	p
0.5, 0, 0, 0.5	1	DA	1.17
0.4, 0.1, 0.1, 0.4	Positive	BA, CD, DA	1.13
0.25, 0.25, 0.25, 0.25	0	CD, DA	1.07
0.1, 0.4, 0.4, 0.1	Negative	AB, CD, DA	1.02
0, 0.5, 0.5, 0	-1	CB	0.98

Table 3: Impact of correlation between productivity and endowments

and the benefits of commodity taxation are increased.²⁵

4.3 The Impact of the Correlation Between Productivity and Initial Wealth

We now return to the central variant endowment levels (namely $w_1 = 2, w_2 = 10$), but consider alternative distribution of types. In the central variant specification (uniform distribution) the correlation between productivity and endowment is zero. When the proportion of types A and D is increased, the distribution reflects a positive correlation between productivities and endowments. Similarly, higher levels of π^B and π^C yield a negative correlation. Once again, we restrict our attention to the most significant variables which are presented in Table 3.

Not surprisingly, the commodity tax tends to be higher the higher is the correlation between productivity and wealth. Recall that the commodity tax is an indirect way to tax endowments. Now, taxing endowments (albeit indirectly) is more beneficial when a higher endowment tends to go hand in hand with a higher productivity. Put differently, under positive correlation, a commodity tax is a substitute not only for the taxation of (unobservable) endowments but also for the taxation of (unobservable) productivities. Technically, this is reflected by the direction of the binding incentive

²⁵Contrast this with the arguments presented in Section 4.1. In the central variant case, it does not pay to redistribute from B to A ; the negative impact on DA , outweighs the immediate benefits. When $\omega^1 = 0, \omega^2 = 12$, on the other hand, such redistribution becomes beneficial and is carried out up to the point where BA binds.

α	Binding IC constraints	p
1,0,0,0 (Rawls)	AB, BA, CD, DA, DB	1.76
0.5, 0, 0.5, 0 (left)	BA, DA, DC	1.51
0.25, 0.25, 0.25, 0.25 (central variant)	CD, DA	1.07
0.1, 0.4, 0.1, 0.4 (right)	AB, CD, DA	0.67

Table 4: Impact of political weights

constraints. For instance, under perfect positive correlation only DA is binding, so that (20) unambiguously calls for a positive commodity tax. Similarly, under perfect negative correlation, only CB bind which calls for a negative commodity tax.²⁶

4.4 The impact of the government’s political preferences

So far, we have restricted our attention to the case of uniform weights (utilitarian objective). To conclude our simulation exercise, we shall now consider a few alternative specifications. We shall concentrate on three extreme cases, which reflect in a highly stylized way different types of political preferences. First, we consider the Rawlsian case $\alpha = (1, 0, 0, 0)$ where all the weight is put on the poorest individual.²⁷ This is a quite traditional case in the optimal tax literature. The other two cases are more unusual. On the one hand, we consider the case of a “liberal” government which puts all the weight on the low-wealth individual, irrespective of their productivity ($\alpha = (0.5, 0, 0.5, 0)$). On the other hand, we study the case where wealthy individuals receive most of the weight—a stylized (and certainly debatable) way to account for a right-wing government. The results are summarized in Table 4.

Both the Rawlsian and the “liberal” case yield a positive and rather significant

²⁶Observe, however, that even a perfect negative correlation does not *necessarily* imply a negative commodity tax. The fact that CB binds in our case means that productivity differential constitute the “dominant” source of inequality. If, on the other hand, productivity differentials were small compared to wealth differentials, a positive tax could arise even under negative correlation.

²⁷Because of the incentive constraints, type A is necessarily the worst off in our setting.

commodity tax (up to 76% !).²⁸ At the other extreme, when high-wealth individuals receive higher weights, the optimal commodity tax is negative (the consumption of good 2 is subsidized). This may at first appear surprising for commodity taxation is often viewed as “regressive”. However, it is perfectly in line with the role played by commodity taxation in this specific setting. Rawlsian and “liberal” governments clearly have a strong incentive to tax wealth and, once again, the commodity tax is used as an indirect way to screen for the unobservable endowment.²⁹ Similarly, a subsidy on good 2 favors wealthier individuals and is optimal if they receive a high weight.

5 Concluding comments

This paper has re-examined the optimal direct/indirect tax mix in a framework where individuals differ in several unobservable characteristics. It has shown that contrary to Atkinson and Stiglitz’s (1976) result, differential commodity taxes *do* have a role to play as instruments of optimal tax policy—an optimal (general) income tax will not suffice. To make this point, we have used a model which differs from Atkinson and Stiglitz’s framework in essentially only one respect: we have considered more than one dimension of (unobservable) heterogeneity. This single variation has a drastic impact on the results, though. While Atkinson and Stiglitz’s result remains (of course) valid within their framework, it does not appear to be robust, and cannot be generalized in any significant way. From that perspective it can be considered as a technical curiosity which arises only in a special (and not particularly compelling) case. Though insightful and interesting in itself, their result thus has to be qualified accordingly. In particular, it may not be the right basis to justify strong policy recommendations.

We have also derived general expressions for the optimal commodity tax rates and

²⁸Recall that the simulation is based on purely hypothetical parameters, with no claim for realism. Consequently, only the ranking between scenarios matters; actual levels of taxation have no specific meaning here.

²⁹Formally, these two specifications translate into incentive constraints which bind from high-wealth to low-wealth individuals, and the positive tax then follows immediately from (20).

discussed their interpretation. Commodity taxes impose an efficiency cost through the distortions they create. However, they are also generally effective in relaxing otherwise binding incentive constraints. Put differently, they allow to better separate the individuals and to improve screening for unobservable characteristics. The optimal tax structure strikes a balance between these conflicting effects, and from that perspective our expressions resemble traditional optimal tax formulas.

Finally, we have attempted to go beyond these general interpretations and to provide more specific results. On these grounds, our results are admittedly preliminary, and more research is certainly needed. The major difficulty with the optimal tax formulas is that the implied structure of commodity taxes crucially depends on the pattern of binding incentive constraints. In the single dimensional case it is generally possible to make “plausible” assumptions about this pattern. The normal assumption is that self selection constraints bind from the more able to the less able. Under multi-dimensional adverse selection, however, there is no longer such a compelling “normal” case. For instance, it is not clear which (if any) of the incentive constraints linking the wealthy but unproductive and the poor but productive can be expected to bind (see Section 4).

To overcome this difficulty we have adopted two approaches. First, we have considered general analytical restrictions pertaining to the correlation between characteristics. When combined with additional assumptions on demands (or endowment distribution), these restrictions have been shown to lead to rather simple and intuitive taxation rules (such as the positive relationship between tax rates and income elasticities). Second, we have presented some illustrative examples for which the solution (including the pattern of binding incentive constraints) has been determined numerically.

To achieve a more precise characterization of the optimal commodity tax structure and, most importantly, to make effective policy recommendations, both approaches would have to be pursued further. In particular, the analytical part could be refined by using empirical evidence on the joint distribution of unobservable characteristics

(and or preference structure). More ambitiously, it might be insightful to base the numerical calculations on a more “realistic” specification, calibrated on observed pattern of demand and wealth/productivity distribution.

References

- Arrondel, L., A. Masson and P. Pestieau (1997) "Bequests and inheritance: empirical issues and French-US comparison," in *Is Inheritance Justified?*, ed. Erreygers, G. and T. Vandevelde. Berlin: Springer Verlag, 233–249.
- Atkinson, A.B. (1977) "Optimal taxation and the direct versus indirect tax controversy," *Canadian Journal of Economics*, **10**, 590–606.
- Atkinson, A.B. and J.E. Stiglitz (1976) "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics*, **6**, 55–75.
- Atkinson, A.B. and J.E. Stiglitz (1980) *Lectures on Public Economics*. New York: McGraw-Hill.
- Boadway, R., M. Marchand and P. Pestieau (1994) "Towards a theory of the direct-indirect tax mix," *Journal of Public Economics*, **55**, 71–88.
- Browning, M. and C. Meghir (1991) "The effects of male and female labor supply of commodity demands," *Econometrica*, **59**, 925–951.
- Cremer, H. and F. Gahvari (1995b) "Uncertainty, optimal taxation and the direct versus indirect tax controversy," *Economic Journal*, **95**, 1165–79.
- Due, J. and A.F. Friedlaender (1973) *Government Finance*. Toronto: Irwin.
- Mirrlees, J.A. (1971) "An elaboration in the theory of optimal income taxation," *Review of Economic Studies*, **28**, 175–208.
- Mirrlees, J.A. (1976) "Optimal tax theory: a synthesis," *Journal of Public Economics*, **6**, 327–358.
- Stiglitz, J.E. (1987) "Pareto efficient and optimal taxation and the new new welfare economics," in *Handbook of Public Economics*, ed. Auerbach, A. and M. Feldstein. Amsterdam: North-Holland, 991–1042.
- Stiglitz, J.E. (1988) *Economics of the Public Sector*, 2nd edition. New York: W.W. Norton & Company.