WORKING PAPERS

N° 1636

April 2025 – revised November 2025

# “Tax avoidance and commodity tax differentiation”

Georges Casamatta and Helmuth Cremer

Toulouse School of Economics

# Tax avoidance and commodity tax differentiation*

Georges Casamatta† and Helmuth Cremer‡

April 2025, revised November 2025

**Abstract**

We analyze the optimal combination of direct and indirect taxes in the presence of tax avoidance. Proportional commodity taxes remain part of the optimal tax structure even when avoidance is possible and the Atkinson–Stiglitz conditions hold. Taxing consumption, despite avoidance, enhances the screening of unobserved productivity relative to income taxation alone. Under weak separability and homothetic subutility, optimal commodity taxes are positive and uniform. With non-homothetic preferences, uniformity may not hold, and the optimal differentiation between luxuries and necessities depends on the distribution of productivity.

Keywords: Optimal taxation, Tax avoidance, indirect taxation, Atkinson–Stiglitz theorem, Commodity taxation.

JEL classification: H21, H26

*We thank the reviewer and the Editor, Joseph Harrington, for their constructive comments and suggestions. We also thank François Maniquet and the participants of the 21st LAGV conference (Marseille) for helpful comments. Helmuth Cremer gratefully acknowledges the funding received by TSE from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

†UMR CNRS 6240 LISA, University of Corsica Pasquale Paoli, Corte, France, and Toulouse School of Economics; email: casamatta_g@univ-corse.fr

‡Toulouse School of Economics, University of Toulouse Capitole.

# 1 Introduction

The literature on optimal taxation, pioneered by Mirrlees (1971), has traditionally focused on how taxes influence labor supply—what Slemrod (1995) refers to as the *real* response to taxation. However, individuals can also adjust their reported income without altering their labor supply, either legally (tax avoidance) or illegally (tax evasion).

This paper focuses on tax avoidance—legal strategies used to minimize tax liabilities—an issue of growing importance in advanced economies. A 2011 UK Treasury report (*Tackling Tax Avoidance*) estimates that tax avoidance accounts for approximately 17.5% of the total tax gap (£42 billion), the difference between owed and paid taxes.

The distinction between avoidance and evasion is largely a matter of interpretation. Following Usher (1986), the cost of tax evasion can be viewed as the expense required to eliminate detection risks when misreporting income.

Recent academic work has increasingly examined tax avoidance, with one strand focusing on the elasticity of taxable income (see Saez et al. (2012)). This measure captures both real and avoidance responses and is central to assessing tax distortions. Other studies analyze the tax avoidance margin. For instance, Slemrod (2001) models income taxation with both labor supply and avoidance responses but does not address optimal taxation. Slemrod and Kopczuk (2002) determines the optimal level of tax avoidance, highlighting that unlike labor supply responses, avoidance can be mitigated by government policy. If avoidance responses are significant, optimal policy may favor broadening the tax base over reducing tax rates, contrary to the standard Mirrleesian approach. Similarly, Roine (2006) examines how avoidance-driven tax distortions influence voting behavior on tax policy.

Despite these contributions, few studies adopt a fully-fledged optimal taxation approach. Grochulski (2007) and Casamatta (2021) derive optimal income tax schedules when avoidance is costly. Grochulski (2007) finds that under a sub-additive cost function, individuals should fully report their income at the optimum. Casamatta (2021), assuming a convex cost function, argues that middle-class individuals should be allowed to avoid taxes. However, his results rely on a setting where labor income is the sole tax instrument.[1]

In this paper, we introduce indirect taxation, allowing both labor income and consumption goods to be taxed. We incorporate both labor supply and tax avoidance responses. Our work builds on Boadway et al. (1994), who show that income tax avoidance justifies a role for commodity taxes, even when they would otherwise be redundant. However, their analysis assumes only income taxes can be avoided, making the case for commodity taxation unsurprising. We extend their framework to allow for commodity tax avoidance and generalize their results from a two-type model to a continuous-type setting.

Our analysis contributes to the long-standing debate on direct versus indirect taxation (see Atkinson (1977)). A key result in this debate is the Atkinson-Stiglitz theorem (Atkinson and Stiglitz, 1976), which states that with a nonlinear income tax, taxing consumption is unnecessary if utility is weakly separable between consumption and leisure. This condition is weaker than the requirements in Ramsey taxation, which also demands homothetic subutility for goods. The relevance of the Atkinson-Stiglitz theorem has been questioned in various contexts, including general equilibrium effects (Naito, 1999), multidimensional heterogeneity (Cremer et al., 2001), and income tax evasion (Boadway et al., 1994). We maintain weak separability and abstract from general equilibrium effects and multidimensional heterogeneity, instead focusing on tax avoidance.

[1] These findings have been extended to endogenous labor supply settings by Casamatta (2023). Related contributions include Selin and Simula (2020) and Doligalski and Rojas (2023).

A key insight of our study is that tax avoidance disrupts the traditional Atkinson-Stiglitz result. Since declared income differs from true income, income taxes alone may be insufficient, creating a role for commodity taxation—even when the assumptions of the Atkinson-Stiglitz theorem hold. Consumption taxes provide better screening of true income than a system relying solely on labor income taxation. Remarkably, this remains true even when commodity taxes themselves can be avoided, providing a novel justification for a mix of direct and indirect taxes.

This reasoning applies to a single consumption good, where in the absence of tax avoidance, the consumption tax can be normalized to zero. We concentrate on the more complex case of multiple consumption goods and the implications for tax differentiation. When Engel curves are affine (straight lines that do not necessarily go through the origin), we recover the result from Boadway et al. (1994) that tax rates should be uniform across goods if firms do not avoid taxes. However, when firms engage in tax avoidance and have access to the same avoidance technology, uniform taxation requires homothetic preferences over goods and thus linear Engel curves (as defined in Mathematics). In other words, when both firms and individuals avoid taxes, the conditions for uniform taxation become more stringent, aligning with the Ramsey framework.

With nonlinear Engel curves, optimal commodity taxation is non-uniform but difficult to characterize precisely. In a setting where only individuals avoid taxes, we derive conditions for tax differentiation. Our findings suggest that whether luxuries should be taxed more heavily than necessities depends on the shape of the skill distribution. We illustrate this result using numerical simulations based on the Almost Ideal Demand System (Deaton and Muellbauer, 1980), estimated on food consumption data.

In a related paper Doligalski et al. (2025) show that when a subset of goods is observable and weakly separable, no distortion should be introduced on these goods—even in the presence of tax avoidance—as long as the planner can use nonlinear commodity instruments (or equivalently, type-dependent expenditure subsidies). We concentrate on linear commodity taxes, because only anonymous transactions are observable.[2] In this case, the Atkinson–Stiglitz/Doligalski equivalence no longer holds: commodity taxation is no longer redundant, and, unless utilities are homothetic, optimal tax rates are positive and may differ across goods.

The remainder of this paper is structured as follows. Section 2 presents the model. Section 3 solves the social planner's problem. Section 4 examines the case for uniform taxation, while Section 5 addresses tax differentiation.

## 2 Model

We consider an economy with $1+n$ goods: labor supply and $n$ consumption goods. All individuals share the same utility function $u(\mathbf{c}, l)$, where $\mathbf{c} \equiv (c_1, \dots, c_n)$ is the vector of goods consumed and $l$ denotes labor supply. The utility function is standard: it is increasing in consumption goods, decreasing and strictly concave in labor supply, strictly quasi-concave, and twice continuously differentiable. Individuals differ in their productivity $w$, which is distributed according to the distribution function $F(\cdot)$ and the density $f(\cdot)$ on the support $[w_-, w_+]$. An individual with productivity $w$ who supplies $l$ units of labor generates income $y = wl$. Individuals can (legally) reduce their taxable income, denoted $\hat{y}$. The cost of hiding $\Delta$ euros is $\phi(\Delta)$, where $\phi(0) = 0$, $\phi'(0) = 0$, and $\phi''(\cdot) > 0$.

Firms also have the ability to conceal sales from the authorities. Each unit of output concealed by

[2] Throughout the paper we use the term linear in a strict mathematical sense, while economists otherwise often refer to affine functions as linear. In particular a linear tax is proportional.

firms in industry $i$ entails a resource cost of

$$G_i(1-\alpha),$$

which is an increasing and convex function of the proportion $(1-\alpha)$ of unreported sales.[3]

Policy instruments consist of a (nonlinear) tax $\hat{T}(\hat{y})$ on reported income as well as linear tax rates $t_i$ on each commodity $i = 1, \ldots, n$.[4] Let $\hat{R} \equiv y - \hat{T}(\hat{y}) - \phi(y - \hat{y})$ denote the net labor income of an individual with true income $y$ and reported income $\hat{y}$.

## 2.1 Consumption and labor supply decisions

An individual with productivity $w$ and exogenous income $I$ solves the following program:

$$\max_{\mathbf{c},y,\hat{y}} u(\mathbf{c}, y/w) \quad \text{st} \quad \sum_k p_k c_k = x, \tag{1}$$

where $p_k$ is the consumer price of good $k$ and $x \equiv \hat{R} + I$ is disposable income.

Let $U(w)$ denote the level of utility achieved by an individual with productivity $w$. Applying the envelope theorem, we have

$$\frac{dU}{dw} = -\frac{y}{w^2} u_l, \tag{2}$$

where $u_l \equiv \partial u / \partial l$ is the marginal disutility of labor. This equation represents the local incentive constraint in the government's problem.

## 2.2 Firms' behavior

The marginal cost is constant and normalized to 1. Firms producing good $i$ maximize

$$\begin{aligned} \pi_i &= (p_i - 1 - \alpha_i t_i - (1-\alpha_i)\, G_i\,(1-\alpha_i))\, c_i, \\ &= (p_i - 1 - \alpha_i t_i - g_i\,(1-\alpha_i))\, c_i, \end{aligned}$$

where $\alpha_i$ is the reported share of sales and $g_i(1-\alpha_i) = (1-\alpha_i)G_i(1-\alpha_i)$. The first- and second-order conditions with respect to $\alpha_i$ require that

$$\begin{aligned} g_i'(1-\alpha_i) &= t_i \\ -g_i''(1-\alpha_i) &< 0 \end{aligned} \tag{3}$$

where (3) is satisfied because $G_i$ is an increasing and convex function. We also assume that $g'(0) = 0$ and $g'(1) = \infty$ to ensure an interior solution.

Defining

$$t_i^e \equiv \alpha_i t_i, \tag{4}$$

the market is in equilibrium when

$$p_i = 1 + g_i + t_i^e. \tag{5}$$

[3] On the firm's side, our specification is inspired by Cremer and Gahvari (1993).

[4] This reflects the by now traditional information structure in mixed taxation models. Reported income is observable at an individual level, while for goods only anonymous transactions are observed.

### 2.3 The social planner's program

According to the revelation principle, individuals directly and truthfully report their type $w$. The government maximizes social welfare, defined as the sum of a concave transformation $\Gamma(\cdot)$ of individual utilities, subject to resource and incentive constraints. Policy instruments are the vector of commodity tax rates $\mathbf{t} = (t_1, \ldots, t_n)$, the labor income tax schedule $T(w)$, and reported income $\hat{y}(w)$.[5] The utility of an individual with productivity $w$ is given by:

$$U(w) = \max_{\mathbf{c},y} u(\mathbf{c}, y/w) \quad \text{st} \quad \sum_k p_k c_k = y - T(w) - \phi(y - \hat{y}(w)). \tag{6}$$

Writing the government budget constraint (GBC) as:

$$\int \left( \sum_k t_k^e c_k + T(w) \right) f(w) dw \geq 0, \tag{7}$$

the planner's program can then be stated as follows:

$$\max_{\mathbf{t}, \hat{y}(.), T(.)} \int \Gamma(U(w)) dF(w) \quad \text{st} \quad (2), (4), (5), (6), (7).$$

## 3 Optimal commodity taxes

The Lagrangian associated with the social planner's program is:

$$\mathcal{L} = \int \left[ \Gamma(U(w)) + \mu \left( \sum_k t_k^e c_k + T(w) \right) \right] f(w) dw + \int \lambda(w) \left( \frac{dU}{dw} + \frac{y}{w^2} u_l \right) dw,$$

where $\mu$ and $\lambda(w)$ are the multipliers associated with GBC constraint and the incentive constraints, respectively. Integrating by parts:

$$\int \lambda(w) \frac{dU}{dw} dw = -\int \lambda'(w) U(w) dw + \lambda(w_+) U(w_+) - \lambda(w_-) U(w_-),$$

the Lagrangian can be rewritten as:

$$\begin{aligned} \mathcal{L} = & \int \left[ \Gamma(U(w)) + \mu \left( \sum_k t_k^e c_k + T(w) \right) \right] f(w) dw \\ & + \int \left( \lambda(w) \frac{y}{w^2} u_l - \lambda'(w) U(w) \right) dw + \lambda(w_+) U(w_+) - \lambda(w_-) U(w_-). \end{aligned}$$

[5] As usual in the optimal tax literature we use a mechanism design approach for the determining the equilibrium tax and thus the allocation. The marginal tax rate can then be determined from the individuals FOCs which must be satisfied to implement the solution. We neglect this step because we focus on commodity taxes.

The first-order conditions with respect to $t_i$ and $T$ are given by:

$$\frac{\partial \mathcal{L}}{\partial t_i} = \int \left[ \Gamma'(U)\frac{dU}{dp_i}\frac{dp_i}{dt_i} + \mu \left( \frac{dt_i^e}{dt_i}c_i + \sum_k t_k^e \frac{dc_k}{dp_i}\frac{dp_i}{dt_i} \right) \right] f(w)dw + \int \left[ \lambda \frac{d}{dp_i}\left(\frac{y}{w^2}u_l\right)\frac{dp_i}{dt_i} - \lambda' \frac{dU}{dp_i}\frac{dp_i}{dt_i} \right] dw = 0,\ i = 1, \ldots, n \tag{8}$$

$$\frac{\partial \mathcal{L}}{\partial T} = \left[ \Gamma'(U)\frac{dU}{dT} + \mu \left( 1 + \sum_k t_k^e \frac{dc_k}{dT} \right) \right] f(w) + \lambda \frac{d}{dT}\left(\frac{y}{w^2}u_l\right) - \lambda' \frac{dU}{dT} = 0. \tag{9}$$

Dividing (8) by $dp_i/dt_i$ yields:

$$\int \left[ \Gamma'(U)\frac{dU}{dp_i} + \mu \left( A_i c_i + \sum_k t_k^e \frac{dc_k}{dp_i} \right) f(w) + \lambda \frac{d}{dp_i}\left(\frac{y}{w^2}u_l\right) - \lambda' \frac{dU}{dp_i} \right] dw = 0, \tag{10}$$

where

$$A_i \equiv \frac{dt_i^e/dt_i}{dp_i/dt_i}.$$

We now multiply (9) by $c_i$ :

$$c_i \left[ \Gamma'(U)\frac{dU}{dT} + \mu \left( 1 + \sum_k t_k^e \frac{dc_k}{dT} \right) \right] f(w) + \lambda c_i \frac{d}{dT}\left(\frac{y}{w^2}u_l\right) - \lambda' c_i \frac{dU}{dT} = 0,$$

take the integral of this condition and subtract it from (10). Using Roy's identity, $dU/dp_i = c_i(dU/dT)$, we obtain:

$$\int \left[ \mu \sum_k t_k^e \left( \frac{dc_k}{dp_i} - c_i \frac{dc_k}{dT} \right) f + \lambda \left( \frac{d}{dp_i}\left(\frac{y}{w^2}u_l\right) - c_i \frac{d}{dT}\left(\frac{y}{w^2}u_l\right) \right) + \mu c_i (A_i - 1) f \right] dw = 0.$$

Using the Slutsky equation for consumption goods:

$$\frac{dc_k}{dp_i} = S_{ki} + c_i \frac{dc_k}{dT},$$

this results in:

$$\int \left[ \mu \left( \sum_k t_k^e S_{ki} \right) f + \lambda \left( \frac{d}{dp_i}\left(\frac{y}{w^2}u_l\right) - c_i \frac{d}{dT}\left(\frac{y}{w^2}u_l\right) \right) + \mu c_i (A_i - 1) f \right] dw = 0. \tag{11}$$

Note that:

$$\frac{d}{d\kappa}\left(\frac{y}{w^2}u_l\right) = \frac{dy}{d\kappa}\left( \frac{1}{w^2}u_l + \frac{y}{w^3}u_{ll} \right) + \frac{y}{w^2}\frac{d^2u}{d\kappa dl},\ \kappa = p_i, T.$$

Using the Slutsky equation for labor supply:

$$\frac{dy}{dp_i} = S_{yi} + c_i \frac{dy}{dT},$$

condition (11) becomes:

$$\int \left[ \mu \left( \sum_k t_k^e S_{ki} \right) f + \lambda \left( S_{yi} \left( \frac{1}{w^2}u_l + \frac{y}{w^3}u_{ll} \right) + \frac{y}{w^2} \left( \frac{d^2u}{dp_i dl} - c_i \frac{d^2u}{dT dl} \right) \right) + \mu c_i (A_i - 1) f \right] dw = 0.$$

We differentiate Roy's identity:

$$\frac{d^2u}{dp_i dl} = \frac{du}{dT}\frac{dc_i}{dl} + c_i \frac{d^2u}{dTdl},$$

so that the optimality condition is finally:

$$\int \left[ \mu \left( \sum_k t_k^e S_{ki} \right) f + \lambda \left( \frac{y}{w^2} \frac{du}{dT} \frac{dc_i}{dl} + S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) \right) + \mu c_i \left( A_i - 1 \right) f \right] dw = 0. \tag{12}$$

The effect of increasing the tax rate on good $i$ can be decomposed into three parts. First, it generates substitution distortions across goods $\left( \sum_k t_k^e S_{ki} \right)$. Second, it affects incentive constraints through two channels. The first reflects complementarities between consumption and leisure, $\left( (y/w^2)(du/dT)(dc_i/dl) \right)$: goods more complementary to leisure should be taxed more (see Christiansen (1984); Jacobs and Boadway (2014)), although this channel vanishes under weak separability (see (14)). The second is specific to tax avoidance: a higher $t_i$ induces a compensated change in labor income, $S_{yi}$. When $S_{yi} < 0$ (resp. $> 0$), labor income decreases (resp. increases), which—via (2)—relaxes (resp. tightens) incentive constraints. Put differently, a compensated reduction (resp. increase) in labor income by the mimicked makes mimicking less (resp. more) attractive. Finally, a higher $t_i$ also induces firms to conceal a larger share of their tax base, adding the term $c_i(A_i - 1)$.

With a utility function weakly separable between goods and labor:

$$u(\mathbf{c}, l) = \Phi(f(\mathbf{c}), l), \tag{13}$$

the term $dc_i/dl$ is equal to 0 and the optimality condition reduces to:

$$\int \left[ \mu \left( \sum_k t_k^e S_{ki} \right) f + \lambda S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) + \mu c_i \left( A_i - 1 \right) f \right] dw = 0. \tag{14}$$

For the remainder of the paper we consider weakly separable preferences so that according to the Atkinson-Stiglitz theorem, absent of tax avoidance, commodity taxes should be uniform. We indeed show in Appendix B that the terms within the integral vanish when the individuals and firms do not practice tax avoidance and the commodity taxes are uniform. This demonstration builds upon the following Lemma, which proof is in Appendix A:

**Lemma 1.** *Let $S_{ki}$ and $S_{yi}$ denote the compensated effects of a marginal increase in the price of good $i$ on the consumption of good $k$ and on labor income, and let $R$ and $I$ denote net labor income and exogenous income. We then have:*

$$\sum_k p_k S_{ki} = \frac{dR}{dp_i} + c_i \frac{dR}{dI}. \tag{15}$$

*Under tax avoidance by individuals, this becomes:*

$$\sum_k p_k S_{ki} = \left(1 - \phi'\right) S_{yi}.$$

## 4 Uniform tax structure

We now go back to the avoidance setting and ask whether the optimal tax structure is uniform. To answer this question, we assume that all industries offer the same tax avoidance opportunities: $G_i = G$ for all $i$, and we consider a uniform tax structure: $t_1 = \ldots = t_n \equiv t$. Noticing that in this framework $\alpha_i = \alpha$ and $A_i = A$ for all $i$, the optimality condition (14) becomes:

$$\int \left[ \mu\alpha \left( \sum_k tS_{ki} \right) f + \lambda S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) + \mu c_i (A-1) f \right] dw = 0,$$

which implies:

$$t = \frac{\int \lambda \left( S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) - \mu c_i (A-1) f \right) dw}{-\mu\alpha \int \left( \sum_k S_{ki} \right) f dw}.$$

From Lemma 1, $\sum_k p_k S_{ki} = (1-\phi') S_{yi}$. This condition becomes $\sum_k S_{ki} = (1-\phi') S_{yi}/(1+g+t)$ in the case of a uniform tax structure. Substituting into the previous expression, we obtain:

$$\frac{t}{1+g+t} = \frac{\int \lambda \left( S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) - \mu c_i (A-1) f \right) dw}{-\mu\alpha \int (1-\phi') S_{yi} f dw}.$$

We finally observe that, with a weakly separable utility function (Goldman and Uzawa, 1964; Sandmo, 1974):

$$S_{yi} = \varphi(w) \frac{dc_i}{dx}. \tag{16}$$

Considering affine Engel curves:[6]

$$c_i = \gamma_i(\mathbf{p}) + \zeta_i(\mathbf{p}) x,$$

equation (16) becomes

$$S_{yi} = \varphi(w) \zeta_i(\mathbf{p}).$$

Therefore a uniform tax structure is optimal with affine Engel curves if:

$$\frac{t}{1+g+t} = \frac{\int \lambda \varphi(w) \left( \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) - \frac{\mu c_i (A-1) f}{\varphi(w) \zeta_i(\mathbf{p})} \right) dw}{-\mu\alpha \int (1-\phi') \varphi(w) f dw}.$$

Two observations are in order. First, when firms do not avoid taxes ($A = 1$), a uniform tax structure is optimal, thus restoring the result of Boadway et al. (1994). However, this result no longer holds when firms engage in tax avoidance, since the numerator now depends on $i$. In such a case, the optimal tax structure is uniform only if consumption is proportional to disposable income, meaning that preferences over produced goods are (strictly) homothetic. Interestingly, the possibility of commodity tax avoidance in addition to income tax avoidance thus strengthens the conditions for obtaining a uniform tax. More precisely, we return to the condition which yields uniform commodity taxes in the Ramsey model, that is absent of income taxation, see Atkinson and Stiglitz (1972).

This leads to the following proposition.

**Proposition 1.** *With affine Engel curves and firms practicing avoidance, all having access to the same avoidance technology, a uniform commodity tax structure is optimal if and only if preferences over goods are homothetic.*

[6] As pointed out by Boadway et al. (1994) this is the case when the subutility for produced goods is quasi-homothetic, that is homothetic to some arbitrary point which does not need to be the origin.

When only households engage in avoidance, a uniform commodity tax can be optimal under affine Engel curves, as in Boadway et al. (1994). However, once firms are also able to conceal part of their sales, a uniform statutory rate no longer implies a uniform effective tax rate: each sector remits only $\alpha_i t$, where $\alpha_i$ depends on its avoidance technology. This breaks the equivalence across goods and creates heterogeneous distortions in the tax base.

With homothetic preference, on the other hand, all goods are consumed in constant proportions at every income level. Consequently, commodity demands do not convey differential information about true income because no good is more useful than another to screen high- and low-productivity individuals. In other words, differentiating tax rates yields no informational benefit, while still generating efficiency costs.

## 5 Tax differentiation with general Engel curves

We now focus on the case where only individuals engage in tax avoidance, so that $t_i^e = t_i$ and $A_i = 1$ for all $i = 1, \ldots, n$. The next proposition (proof in Appendix C) states (sufficient) conditions under which, with nonlinear Engel curves, optimal commodity taxes are differentiated.

**Proposition 2.** *Consider two goods $j$ and $k$, and assume that: (i) the utility function is weakly separable between goods and labor; (ii) both goods are (Hicksian) substitutes for leisure: $S_{yj} < 0$ and $S_{yk} < 0$; (iii) the ratio $S_{yk}/S_{yj}$ increases with productivity. Then, starting from a uniform tax structure, it is optimal to increase (resp. decrease) the tax rate on good $k$ if the density function is sufficiently skewed to the right (resp. left).*

This proposition shows that, with nonlinear Engel curves, a sufficiently skewed distribution of productivities leads to differentiated optimal taxes. We conjecture that the result extends to more general distributions and that uniform taxation arises only in degenerate cases. A formal proof, however, would require technical developments beyond the scope of this paper.

The intuition is that good $k$ should be taxed more heavily than good $j$ whenever the corresponding redistributive gain exceeds the additional efficiency loss from taxing $k$.[7] As condition (23) in Appendix C makes clear, the relative efficiency cost (left-hand side) depends on the productivity distribution $f(\cdot)$, since distortions aggregate across types and are therefore weighted by the density. By contrast, the redistributive effect (right-hand side) reflects how a change in the allocation of one type affects the incentive constraints of all higher types, and is thus independent of $f$. It follows that the skewness of the productivity distribution is the key determinant of whether the optimal tax rate on good $k$ should be raised or lowered relative to that on good $j$.

In the next Lemma, we establish sufficient conditions for assumption (*iii*) to be true.

**Lemma 2.** *Consider two goods $j$ and $k$, and assume that: (i) the utility function is weakly separable between goods and labor; (ii) both goods are normal; (iii) disposable income is strictly increasing in productivity; (iv) the Engel curve for good $j$ (resp. $k$) is strictly concave (resp. convex). Then the ratio $S_{yk}/S_{yj}$ increases with productivity.*

[7]The term $S_{yi}$ can be both negative or positive. To reduce the number of cases to be considered, Proposition 2 assumes that $S_{yi} < 0$. This is not a general property, even with weakly separable preferences, but it holds for instance for all goods with a CES utility and for the AIDS preferences with the parameters considered in the illustration. When this assumption holds the goods are all Hicksian substitutes to leisure: an increase in $t_i$ creates a downward substitution effect on labor supply. Note that this is the worst-case scenario for the desirability of taxing that good because the traditional Corlett-Hagues rule would plead for a subsidy. This is because a tax on that good exacerbates labor supply distortions. In this case a tax is only desirable if it provides redistributive benefits.

*Proof.* With a weakly separable utility function (assumption *(i)*), (16) implies:

$$\frac{S_{yk}}{S_{yj}} = \frac{\frac{dc_j}{dx}}{\frac{dc_i}{dx}}.$$

Assumptions *(ii)*, *(iii)*, and *(iv)* then imply that $S_{yk}/S_{yj}$ increases with productivity. □

This lemma establishes that, when both goods are normal and disposable income rises with productivity, the ratio $S_{yk}/S_{yj}$ increases with productivity if good $j$ is a necessity (concave Engel curve) and good $k$ a luxury (convex Engel curve). Together with Proposition 2, this implies that, starting from uniform taxation, the luxury should be taxed more heavily if the productivity distribution is sufficiently right-skewed. Conversely, and perhaps more surprisingly, the tax on the luxury should be reduced when the distribution is sufficiently left-skewed.

The intuition is that taxing a luxury good distorts labor supply disproportionately at the top of the skill distribution. When the distribution is heavily left-skewed, with most individuals highly skilled, aggregate distortions can outweigh the redistributive benefit, reversing the ranking of optimal rates. Although such extreme left-skewness is empirically implausible—and the case for taxing luxuries less than necessities is of limited practical relevance—the polar results in Proposition 2 do not rule out similar configurations under more realistic distributions. We illustrate this with a numerical example based on a lognormal distribution, where the optimal tax structure calls for taxing some luxury goods less heavily than necessities.

# 6 Application to AIDS preferences

In this section, we continue to assume that only individuals avoid taxes.

## 6.1 The Almost Ideal Demand System

A convenient representation of preferences for introducing nonlinear Engel curves is the Almost Ideal Demand System (AIDS) developed by Deaton and Muellbauer (1980). In this model, the demand function for good $i$ is given by:

$$c_i = \frac{x}{p_i}(\alpha_i + \sum_j \gamma_{ij} \ln p_j + \beta_i(\ln x - \ln P)), \tag{17}$$

where $P$ is a price index defined as:

$$\ln P = \alpha_0 + \sum_k \alpha_k \ln p_k + \frac{1}{2}\sum_k \sum_j \gamma_{kj} \ln p_k \ln p_j,$$

and AIDS parameters meet the following restrictions:

$$\sum_i \alpha_i = 1,\ \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = 0,\ \sum_i \beta_i = 0,\ \gamma_{ij} = \gamma_{ji}. \tag{18}$$

The indirect utility function $v$ is:

$$v(\mathbf{p}, x) = \frac{\ln x - \ln P}{b}, \tag{19}$$

where

$$b = \beta_0 \prod_i p_i^{\beta_i}.$$

### 6.2 Analytical derivations

Differentiating (17), we have:

$$\frac{dc_i}{dx} = \frac{c_i}{x} + \frac{\beta_i}{p_i}$$
$$\frac{d^2 c_i}{dx^2} = \frac{1}{x}\frac{\beta_i}{p_i}.$$

Goods with $\beta_i > 0$ (resp. $\beta_i < 0$) have strictly convex (resp. concave) Engel curves and are thus classified as luxuries (resp. necessities), so condition *(iv)* of Lemma 2 holds. With AIDS preferences, however, some goods may be inferior (violating condition *(ii)*), and it is generally difficult to verify analytically whether assumption *(iii)* is satisfied. We therefore rely on numerical simulations in the next section.

### 6.3 Numerical simulations

We consider an additively separable utility function:

$$u(\mathbf{c}, l) = f(\mathbf{c}) + \varphi(l)$$

where the preferences for goods satisfy AIDS and $\varphi(l) = -l^2/2$. The social welfare transformation is $\Gamma(U) = U$. We assume that the cost of avoidance takes the form: $\phi(\Delta) = 0.2\left(\Delta^2/2\right)$.

The Almost Ideal Demand System was estimated by Anderson and Blundell (1983) using data on consumers' expenditures in Canada for the period 1947-1979. Demands for five categories of non-durable goods are considered: food, clothing, energy, transport, and recreation. The estimates of the AIDS parameters are displayed in Table 1. According to these estimates, two goods have negative values of $\beta$ and are therefore necessities (food and clothing), while the other three goods (energy, transport, and recreation) are luxuries.

| Commodity $i$ | $\alpha_i$ | $\beta_i$ | $\gamma_{i1}$ | $\gamma_{i2}$ | $\gamma_{i3}$ | $\gamma_{i4}$ | $\gamma_{i5}$ |
|---|---|---|---|---|---|---|---|
| 1: Food | 0.342 | −0.096 | 0.095 | −0.011 | 0.023 | 0.007 | −0.114 |
| 2: Clothing | 0.129 | −0.102 | −0.011 | −0.008 | 0.017 | −0.029 | 0.031 |
| 3: Energy | 0.055 | 0.014 | 0.023 | 0.017 | 0.006 | −0.030 | −0.016 |
| 4: Transport | 0.147 | 0.036 | 0.007 | −0.029 | −0.030 | 0.032 | 0.020 |
| 5: Recreation | 0.327 | 0.148 | −0.114 | 0.031 | −0.016 | 0.020 | 0.079 |

Table 1: AIDS parameters

Finally, productivites are distributed according to a (truncated) lognormal distribution on the support $[1, 8]$.

The optimal tax rates are shown in Table 2. In the first row of this table, we set $\beta_i = 0$ for all five goods. So none of them is a necessity or a luxury. We check that the optimal tax rate is the same for all goods. The second row shows the optimal tax rates when $\beta_i$ are set to the values estimated by Anderson and Blundell (1983). We obtain that one of the luxury goods (energy) should be taxed less

| $\beta_i$ | $t_1^*$ | $t_2^*$ | $t_3^*$ | $t_4^*$ | $t_5^*$ |
|---|---|---|---|---|---|
| 0 | 0.203 | 0.203 | 0.203 | 0.203 | 0.203 |
| estimates | 0.184 | 0.194 | 0.154 | 0.222 | 0.226 |

Table 2: Optimal tax rates

than the two necessity goods (food and clothing), while a larger tax rates should be applied to the other two luxury goods (transport and recreation).

Finally, we compute the welfare gain from implementing the optimal commodity taxes. We measure this as the change in social welfare when moving from zero to optimal tax rates, normalized by the government budget constraint multiplier to obtain the monetary equivalent. Expressed relative to mean labor income in the no-tax benchmark, the aggregate welfare gain is modest, at only $+0.6\%$ (both when the $\beta_i$'s are set to zero or to their estimated values). In contrast, the individual equivalent variation relative to labor income shows much larger redistributive effects, ranging from about $+12\%$ for the lowest-productivity types to -9% for the highest types. Note that the results suggest that the welfare gains are coming predominantly from setting consumption taxes at the positive (and uniform) level, while benefits from tax differentiation appear to be negligible. Recall that with tax avoidance a uniform commodity tax is not a redundant instrument.

# 7 Conclusion

We analyze the optimal combination of income and commodity taxes when both can be avoided at some cost. Apart from tax avoidance, our framework satisfies the conditions of the Atkinson-Stiglitz theorem, implying that in the absence of avoidance, commodity taxes would be unnecessary. However, we show that when avoidance is possible both tax instruments become essential.

Commodity taxes should generally be non-uniform unless the subutility function for goods is homothetic, ensuring linear Engel curves that pass through the origin. Consequently, the presence of both income and commodity tax avoidance strengthens the conditions for uniform taxation. Recall indeed that affine Engel curves suffice when only income tax can be avoided, as shown by Boadway et al. (1994).

When Engel curves are nonlinear, commodity taxes are typically non-uniform, though their optimal structure is complex. Interestingly, it is not always optimal to tax luxuries more than necessities—this depends on the shape of the productivity distribution.

# References

Anderson, G. and R. Blundell (1983, July). Testing Restrictions in a Flexible Dynamic Demand System: An Application to Consumers' Expenditure in Canada. *The Review of Economic Studies 50*(3), 397.

Atkinson, A. and J. Stiglitz (1972, April). The structure of indirect taxation and economic efficiency. *Journal of Public Economics 1*(1), 97–119.

Atkinson, A. and J. Stiglitz (1976, July). The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics 6*(1-2), 55–75.

Atkinson, A. B. (1977). Optimal Taxation and the Direct versus Indirect Tax Controversy. *The Canadian Journal of Economics / Revue canadienne d'Economique 10*(4), 590–606. Publisher: [Wiley, Canadian Economics Association].

Boadway, R., M. Marchand, and P. Pestieau (1994, September). Towards a theory of the direct-indirect tax mix. *Journal of Public Economics 55*(1), 71–88.

Casamatta, G. (2021). Optimal income taxation with tax avoidance. *Journal of Public Economic Theory 23*(3), 534–550. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jpet.12495.

Casamatta, G. (2023, August). Optimal income taxation with tax avoidance and endogenous labour supply. *Canadian Journal of Economics/Revue canadienne d'économique 56*(3), 913–939.

Christiansen, V. (1984). Which commodity taxes should supplement the income tax? *Journal of Public Economics 24*, 195–220.

Cremer, H. and F. Gahvari (1993, February). Tax evasion and optimal commodity taxation. *Journal of Public Economics 50*(2), 261–275.

Cremer, H., P. Pestieau, and J.-C. Rochet (2001). Direct versus Indirect Taxation: The Design of the Tax Structure Revisited. *International Economic Review 42*(3), 781–799.

Deaton, A. and J. Muellbauer (1980). An Almost Ideal Demand System. *The American Economic Review 70*(3), 15.

Doligalski, P., P. Dworczak, J. Krysta, and F. Tokarski (2025, August). Incentive Separability. *Journal of Political Economy Microeconomics 3*(3), 539–567.

Doligalski, P. and L. E. Rojas (2023). Optimal redistribution with a shadow economy. *Theoretical Economics 18*(2), 749–791.

Goldman, S. M. and H. Uzawa (1964). A Note on Separability in Demand Analysis. *Econometrica 32*(3), 387–398. Publisher: [Wiley, Econometric Society].

Grochulski, B. (2007). Optimal Nonlinear Income Taxation with Costly Tax Avoidance. *Economic Quarterly 93*(1), 77–109.

Jacobs, B. and R. Boadway (2014, September). Optimal linear commodity taxation under optimal non-linear income taxation. *Journal of Public Economics 117*, 201–210.

Mirrlees, J. A. (1971, April). An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies 38*(2), 175–208.

Naito, H. (1999, February). Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics 71*(2), 165–188.

Roine, J. (2006). The Political Economics of Not Paying Taxes. *Public Choice 126*(1/2), 107–134.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50.

Sandmo, A. (1974). A Note on the Structure of Optimal Taxation. *The American Economic Review 64*(4), 701–706. Publisher: American Economic Association.

Selin, H. and L. Simula (2020, February). Income shifting as income creation? *Journal of Public Economics 182*, 104081.

Slemrod, J. (1995). Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986. *The American Economic Review 85*(2), 175–180.

Slemrod, J. (2001, March). A General Model of the Behavioral Response to Taxation. *International Tax and Public Finance 8*(2), 119–128.

Slemrod, J. and W. Kopczuk (2002, April). The optimal elasticity of taxable income. *Journal of Public Economics 84*(1), 91–112.

Usher, D. (1986, October). TAX EVASION AND THE MARGINAL COST OF PUBLIC FUNDS. *Economic Inquiry 24*(4), 563–586.

# Appendix

## A Proof of Lemma 1

We differentiate the individual budget constraint (1) with respect to $p_i$:

$$c_i + \sum_k p_k \frac{dc_k}{dp_i} = \frac{dR}{dp_i}. \tag{20}$$

We then differentiate the individual budget constraint with respect to $I$:

$$\sum_k p_k \frac{dc_k}{dI} = \frac{dR}{dI} + 1.$$

Multiplying this condition by $c_i$ gives:

$$c_i \sum_k p_k \frac{dc_k}{dI} = c_i \frac{dR}{dI} + c_i. \tag{21}$$

Adding up (20) and (21) and using the Slutsky equation, we get:

$$\sum_k p_k S_{ki} = \frac{dR}{dp_i} + c_i \frac{dR}{dI}.$$

Under tax avoidance by individuals, we have $R(w) = y - T(w) - \phi(y - \hat{y}(w))$. Therefore $dR/dp_i = (dR/dy)(dy/dp_i) = (1 - \phi')(dy/dp_i)$ and $dR/dI = (dR/dy)(dy/dI) = (1 - \phi')(dy/dI)$. So, again using Slutsky:

$$\sum_k p_k S_{ki} = (1 - \phi') S_{yi}.$$

## B Optimal tax schedule with no avoidance

We show here how to recover the Atkinson-Stiglitz theorem in our framework. In this purpose, we assume that tax avoidance is not possible, either for the individuals or for the firms. This implies that labor income $y$ is observable and is therefore a choice variable of the social planner. In this setting, condition (11) becomes:

$$\int \left[ \mu \left( \sum_k t_k S_{ki} \right) f + \lambda \frac{y}{w^2} \left( \frac{d^2 u}{dp_i dl} - c_i \frac{d^2 u}{dT dl} \right) \right] dw = 0,$$

which in turn leads (12) to become:

$$\int \left[ \mu \left( \sum_k t_k S_{ki} \right) f + \lambda \frac{y}{w^2} \frac{du}{dT} \frac{dc_i}{dl} \right] dw = 0.$$

We see that the term related to the relaxation of the incentive constraints,

$$S_{yi} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right)$$

has disappeared. With a weakly separable utility function, the second term within brackets also disappears, so that it remains to be shown that the first term equals 0 with a uniform tax structure. To do so, we note that the net labor income of an individual with productivity $w$ is $R(w) = y(w) - T(w)$. Lemma 1 then implies $\sum_k (1+t_k) S_{ki} = 0$. With a uniform tax structure, this becomes $\sum_k S_{ki} = 0$, so that the optimality condition of the social planner is satisfied.

# C Proof of Proposition 2

We start from a uniform tax structure with common rate $t$ and assume that the FOC (14) is satisfied for good $j$. Under uniform taxation and (15), this condition can be written as

$$\int \left[ \mu \frac{t}{1+t} (1-\phi') S_{yj} f \ + \ \lambda S_{yj} \Big( \frac{u_l}{w^2} + \frac{y\, u_{ll}}{w^3} \Big) \right] dw = 0. \tag{22}$$

Since, by assumption, $S_{yj} < 0$, it is optimal to set a tax rate above $t$ on good $k$ if and only if

$$\begin{aligned} &\int \left[ \mu \frac{t}{1+t} (1-\phi') S_{yk} f \ + \ \lambda S_{yk} \Big( \frac{u_l}{w^2} + \frac{y\, u_{ll}}{w^3} \Big) \right] dw \ > \ 0 \\ \Leftrightarrow \quad & \frac{\int (1-\phi') S_{yk} f \, dw}{\int (1-\phi') S_{yj} f \, dw} \ < \ \frac{\int \lambda S_{yk} \big( \frac{u_l}{w^2} + \frac{y\, u_{ll}}{w^3} \big) dw}{\int \lambda S_{yj} \big( \frac{u_l}{w^2} + \frac{y\, u_{ll}}{w^3} \big) dw}, \end{aligned} \tag{23}$$

where the second line uses (22) to substitute for $\mu$.

Under assumptions *(ii)* and *(iii)*, $\forall w > w_-$:

$$\begin{aligned} & \frac{-S_{yk}(w)}{-S_{yj}(w)} > \frac{-S_{yk}(w_-)}{-S_{yj}(w_-)} \\ \Rightarrow & - S_{yk}(w) > -S_{yj}(w) \frac{S_{yk}(w_-)}{S_{yj}(w_-)} \\ \Rightarrow & \int_{w_-}^{w_+} \lambda (-S_{yk}) \left( -\frac{1}{w^2} u_l - \frac{y}{w^3} u_{ll} \right) dw > \frac{S_{yk}(w_-)}{S_{yj}(w_-)} \int_{w_-}^{w_+} \lambda (-S_{yj}) \left( -\frac{1}{w^2} u_l - \frac{y}{w^3} u_{ll} \right) dw \\ \Rightarrow & \frac{\int_{w_-}^{w_+} \lambda S_{yk} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) dw}{\int_{w_-}^{w_+} \lambda S_{yj} \left( \frac{1}{w^2} u_l + \frac{y}{w^3} u_{ll} \right) dw} > \frac{S_{yk}(w_-)}{S_{yj}(w_-)}, \end{aligned}$$

where we have used the fact that $\lambda > 0, \forall w \in (w_-, w_+)$, and $-(1/w^2)u_l - (y/w^3)u_{ll} > 0$.

When $f(w) \to 0 \ \forall w \neq w_-$ (extremely right skewed distribution), the left side of (23) tends to $S_{yk}(w_-)/S_{yj}(w_-)$. The above inequality then implies that (23) is satisfied and therefore that the tax rate on good $k$ should be increased. A symmetric argument can be used to show that the tax rate should be decreased if the density function is sufficiently skewed to the left.