Discussion Papers In Economics And Business

Supplementary Commodity Taxes, Labor Tax, and the Tax-mix

Yukihiro Nishimura

Discussion Paper 25-08-Rev.

March 2026

Graduate School of Economics
The University of Osaka, Toyonaka, Osaka 560-0043, JAPAN

Supplementary Commodity Taxes, Labor Tax, and the Tax-mix

Yukihiro Nishimura[+*]

March 2026

Abstract

This paper extends the Atkinson–Stiglitz Theorem by relaxing the assumption of weak separability of preferences. We show that, to supplement income taxation, the optimal commodity tax burden relative to disposable income is regressive. This regressivity reinforces the classic U-shaped pattern of optimal labor income taxation. Moreover, when goods are complementary to labor, e.g., child care, the overall marginal labor tax burden increases once commodity taxes are introduced; when they are substitutes, it decreases. Under the Rawlsian objective, these variations in marginal tax rates have stronger effects at lower income levels.

Keywords: Commodity tax, Income tax, Marginal income tax rates

JEL Classification: H21, H24, D63

1. Introduction

[+] This paper was presented at SOPE October 2025 (Kwansei Gakuin University) and Tokyo University. The author acknowledges Robin Boadway, Pierre Pestieau, Antoine Ferey, Laurence Jacquet, Andreas Haufler, Shuichi Tsugawa, Shigeki Kunieda, Motohiro Sato, and Yasuhiro Sato for comments and suggestions. The author is also grateful for financial support from MEXT/JSPS KAKENHI JP23K22126.
The authors have no financial or proprietary interests in any material discussed in this article. Also, we do not have material from other sources that require permission to reproduce.

[*] Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka 560-0043, Japan. E-mail: ynishimu@econ.osaka-u.ac.jp

This paper extends the Atkinson-Stiglitz (1976) Theorem and explores practical features of optimal taxation of labor income. We depart from the weak separability of preferences. In such environments, commodity taxes or subsidies become necessary for achieving efficiency. Our theoretical results provide clear implications for designing a practical policy for the optimal tax mix.

Atkinson and Stiglitz (1976) demonstrated that differential commodity taxes (and, in dynamic models, capital income taxation) are not needed under a set of assumptions. These include: (i) the presence of an optimal nonlinear income tax or the possibility of nonlinear income-tax reform, (ii) utility functions exhibiting weak separability between labor and goods, (iii) individual abilities (skills) being the only unobservable characteristics, and (iv) the absence of externalities or merit good considerations (Boadway and Pestieau 2003; Boadway 2012, Section 3.1.2; Tuomala 2016, Chapter 12). This paper departs from (ii), as weak separability is too restrictive to hold in practice. Consumption patterns often exhibit complementarity or substitutability with labor decisions. For instance, Crawford et al. (2011), in the Mirrlees Review, found evidence of complementarity between child care and labor supply in assessing household demand structures.[1] Also, as discussed in the recent study by Ferey et al. (2024),

[1] We adopt the classical mechanism-design approach to the optimal tax mix, where Hicksian terms and income effects capture key features of optimal policy. Our methodology is closely aligned with the recent study of Boadway and Smart (2025), which has an applied study for child care subsidies using Canadian data. Kaplow (2010) pointed out that optimal commodity tax analysis in our tax-mix context made confusing claims, with some authors offering seemingly opposing conclusions regarding the sign of optimal taxes. To be correct, if goods are complementary to leisure, the commodities are taxed.

items such as savings can vary across income classes, and their substitutability or complementarity with labor supply contributes to such heterogeneity.

This departure from the conditions under which uniform commodity taxes are optimal raises several policy-relevant questions. First, can we determine whether optimal commodity taxes are progressive or not? Under differential commodity taxation, the average tax burden, defined as the tax payment relative to disposable income, is no longer uniform. Second, does Diamond's (1998) U-shaped pattern of marginal income taxes remain robust? Third, which income brackets are most affected by commodity taxes? Recent work on optimal commodity taxation (e.g., Allcott et al. (2019), Ferey et al. (2024), Spiritus (2025)) has not fully addressed these issues.[2]

Mirrlees (1976, Section 5), building on his earlier conclusion that labor income taxes are not "effective weapons" for redistribution (Mirrlees, 1971, p. 206), examined how optimal tax structure adjusts when commodity demands differ across individuals with varying abilities, maintaining weak separability. Mirrlees (1976, Eq. (113)) demonstrated that commodity tax payments per unit of post-tax income (the average tax burden) rise with post-tax income. Moreover, under the same assumption, he showed that the overall tax system should impose lower marginal burdens on labor once commodity taxes are introduced (Mirrlees, 1976, pp. 350–351).

---

[2] Saez (2002) and Piketty and Saez (2013) focus primarily on how taste heterogeneity interacts with labor supply, and in Doligalski et al.'s (2025) extension, commodities are subsidized when they serve as signals of low skills. In contrast, in the present paper, the optimal commodity tax is no longer uniform by departing from weak separability. Moreover, we identify which income brackets are most affected by commodity taxes.

We interpret these results as indicating a potential role of commodity taxes to supplement the progressivity of taxation. As discussed above, our analysis focuses on the joint structure of commodity and income taxation under the simplest setting that incorporates only skill heterogeneity[3] but that relaxes weak separability. Regarding the marginal labor income tax, subsequent studies following Mirrlees (1976), including Edwards et al. (1994), Nava et al. (1996), Jacobs and Boadway (2014), and Ferey et al. (2024), do not provide a conclusive answer as to whether optimal nonlinear income tax rates are higher or lower when optimal linear commodity taxes are introduced.

Regarding the progressivity of commodity taxation, we first derive an implication of the Corlett–Hague (1953) and Christiansen (1984) rule: when preferences are not weakly separable, goods that are complements (substitutes) for leisure should be taxed (subsidized). Because such a commodity tax system imposes a lower tax burden on individuals who supply more labor, the optimal commodity tax structure is regressive when measured by tax payments relative to disposable income. This regressivity arises because commodity taxes serve efficiency rather than redistributive objectives (Lemma 2). Since Corlett-Hague-Christiansen result itself is silent on taxation on goods' income elasticity of leisure, we derive Lemma 2 under the assumption that the commodity demand exhibits homogeneity of degree 1. A recent result by Nishimura (2026) shows an illustrative case in which different income effects

[3] In contrast to Mirrlees' (1976) implicit motivation to identify environments conducive to progressive taxation, it is worth noting that subsequent research has found optimal income taxes to be substantially more progressive than those implied by Mirrlees (1971).

neutralize substitution effects and labor-incentive effects, leading to the tax rates depending only on complementarity with labor.

Surprisingly, this property reinforces Diamond's (1998) finding of a U-shaped pattern in optimal labor income taxation: the total labor tax wedge (the sum of the effective marginal income tax and the additional commodity tax burden from a unit labor supply) increases in income at the upper end of the skill distribution when Pareto weights are constant, and decreases in income in regions where the skill density increases, such as around the median and mode of the income distribution.

Our main result concerning whether marginal labor tax wedges fall or not is presented as follows. If goods are complementary to labor—for example, due to the importance of child care—then the overall marginal labor tax burden becomes higher than in the absence of commodity taxes or subsidies. The reason is as follows. The Corlett–Hague rule and the resulting regressivity of the commodity tax system increase the social marginal value of taxing income, thereby allowing for higher optimal marginal income taxes. Conversely, when goods are overall substitutes for leisure, optimal commodity taxation reduces the marginal labor wedge, stimulating labor supply. We further derive results concerning the conditions under which marginal income tax rates increase or decrease.

Incentive-based taxation that penalizes individuals with low labor supply needs to be accompanied by offsetting targeted transfers to the poorer households. Regarding this, we show that, under the Rawlsian social welfare objective, the absolute magnitude of these increases or decreases in marginal income tax rates is greater for individuals with lower incomes.

The next section sets up the model and derives key properties on the labor incentive effects and income effects of optimal commodity taxation. The following section demonstrates that the classic results of positive marginal tax wedges on labor and the U-shaped pattern are robust under the optimal commodity taxes without assuming weak separability. Then, we show that, when goods are complements of labor, the tax system as a whole should bear higher taxes on labor at the margin after commodity taxes are introduced. The last section concludes.

2. Revisiting and Revising Classic Results

There is a continuum of individuals with mass one distributed by their ability (skill level) n. Let $F(n)$ be the distribution function, with $f(n)$ being its density function. There are I+1 goods, with good 0 serving as the untaxed numeraire, denoted by $c = x^0$. The utility of each individual is represented by a common utility function $u(x, l)$ for a I+1-dimentional vector $x = (x^0, x^1, \dots, x^I)$ and labor l. An individual of skill level $n \in [\underline{n}, \bar{n}]$ supplies labor l_n and earns labor income $z_n = nl_n$, and out of the after-tax income $y_n \equiv z_n - T(z_n)$, the individual chooses the commodity consumption x_n^i $(i = 0 \dots I)$. We set $0 \leq \underline{n} < \bar{n} = \infty$ so the skill distribution is unbounded.

Commodities are produced linearly from labor with a unit marginal cost normalized to one. The government levies a commodity tax rate t_i (a subsidy when it is negative) per unit of good i. The vector of after-tax commodity prices, $1 + t_j = q_j$ for good j, is written as q. Given the after-tax price vector q and (l_n, y_n), individual consumption is conditional in the sense that labor supply (gross income) is being held constant. The conditional demand function is denoted by $x_n(q, l_n, y_n)$, and $x_n^i(q, l_n, y_n)$ for each taxed item i. The Hicksian demand is written as $x_n^{c*,i}(q, l_n, v_n)$ for the targeted level of utility v_n.

For the indirect utility $v_m = u(x_m(q, l_m, y_m), l_m)$, the government, by chooses q and (l_m, y_m) for all individuals, maximizes a Bergson-Samuelson Social Welfare Function (SWF), $\int_{m=\underline{n}}^{\bar{n}} \Psi(v_m)\, dF(m)$, $\Psi' > 0$ and $\Psi'' \leq 0$, subject to the resource constraint $\int_{m=\underline{n}}^{\bar{n}} \{ml_m - \sum_{i=0\ldots I} x_m^i(q, y_m, l_m) - R\}\, dF(m) \geq 0$ for an exogenous revenue requirement $R > 0$, and the conventional incentive constraint $u\big(x_m^i(q, y_m, l_m), l_m\big) \geq u\big(x_n^i(q, y_m, ml_m/n), ml_m/n\big)$ for all m and all n. Let η be the Lagrange multiplier of the resource constraint. The first-order approach for the self-selection problem is $\frac{du(x_m^i(q, z_m - T(z_m), z_m/m), z_m/m)}{dm} = -\frac{\partial u(x_m^i(q, y_m, z_m/m), z_m/m)}{\partial l} \frac{l}{m}$. Let θ_m be the co-state variable associated with this constraint for type *m*, and η be the Lagrange multiplier on the resource constraint. The Lagrangian for the optimal control problem is written as (Jacobs and Boadway (2014), Eq. (20));

$$L \equiv \int_{m=\underline{n}}^{\bar{n}} \{\Psi(v_m) + \eta(ml_m - \sum_{i=0\ldots I} x_m^{c*,i}(q, l_m, v_m) - R)\}\, dF(m) + \int_{m=\underline{n}}^{\bar{n}} \theta_m \frac{\partial u/\partial l}{m} l_m - \frac{d\theta_m}{dm} v_m\, dm$$
$$+\theta_{\bar{n}} u_{\bar{n}} - \theta_{\underline{n}} u_{\underline{n}},$$

where the integration by parts is applied to the incentive constraints.

The first-order conditions (FOC) for l_n and t_i are: for the marginal income tax rate $T'(z_n)$ that type *n* faces at the optimum as $\frac{(-\partial u/\partial l)/n}{\partial v_n/\partial y_n} = 1 - T'(z_n)$, and for $s_n^{ij} = \frac{\partial x_n^{c*,i}}{\partial q_j}$ and $v_y^n = \frac{\partial u(x_n^i(q, y_n, z_n/n), z_n/n)}{\partial y_n}$, in the case where a fraction of people $F(n_0)$ do not work,

(1) $\frac{\partial L}{\partial l_n} = 0 \;\leftrightarrow\; \frac{T'(z_n)}{1-T'(z_n)} + \sum_{i=1\ldots I} t_i \frac{1}{(1-T'(z_n))n} \frac{\partial x_n^{c*,i}}{\partial l_n} = \frac{(v_y^n \theta_n)/\eta}{nf(n)} \frac{(-l_n \partial u/\partial l)/\partial l_n}{-\partial u/\partial l}$

(2) $\frac{\partial L}{\partial q^i} = 0 \;\leftrightarrow\; \int_{m=\underline{n}}^{\bar{n}} \sum_{j=1\ldots I} t_j s_n^{ij}\, dF(m) = \int_{m=n_0}^{\bar{n}} \frac{(v_y^m \theta_m)/\eta}{mf(m)} \frac{\partial x_m^i}{\partial l_m} l_m dF(m)$

where (2) holds for $i = 1, \ldots, I$.

Besides the income tax rate, labor wedges in this setting include an additional component arising from commodity taxation. The left-hand side (LHS) of (1) is called the marginal income wedge on labor income which type n faces, by Jacobs and Boadway (2014). We assume that $\theta_n \geq 0$ for all n, with strict inequality except at $\underline{n}$ and $\bar{n}$. Extending Seade (1982), this property holds under Atkinson-Stiglitz (1976) weak separability when consumption goods are normal goods (an assumption we put below). Mirrlees (1976) showed that $\theta_n \geq 0$ is the case in his specific case of $u = U(a(x_n, n) + b(nl_n, n))$, whereas, in the absence of the restrictions on the demand, Jacobs and Boadway (2014) implicitly assumed $\theta_n \geq 0$ in interpreting the optimal commodity taxes. We show below that $\theta_n \geq 0$ serves as a reinforcing assumption that remains valid at the optimum.

A caveat is that, unlike the total tax wedge on labor in (1), the right-hand side (RHS) (2) involves uncompensated derivative of the conditional demands.[4]

2.1 Labor Incentive Effects

Throughout this paper, in order to get clear-cut results, we assume that the sign of derivatives, such as that of $\sum_{i=1 \ldots I} t_i \, \partial x_n^i / \partial l_n$ (good's overall complementarity with labor supply) is invariant across ability types n. We first show the following:

[4] In the discrete model in Edwards et al. (1994) and Nava et al. (1996), the RHS of (2) represents the difference in commodity demands between type n and type n's mimicker with less labor effort. The uncompensated derivative captures this demand difference arising from the mimicker's reduced labor effort.

Lemma 1: *At the optimal commodity tax, regardless of the sign of* t_i*, the uncompensated labor derivative of commodity demand,* $\sum_{i=1...I} t_i \, \partial x_n^i / \partial l_n$ *is negative for all n.*

Proof: Multiplying each equation of (2) by t_i and summing up, we have:

$$\int_{m=\underline{n}}^{\bar{n}} \sum_{i=1...I} t_i \sum_{j=1...I} s_n^{ij} t_j \, dF(m) = \int_{m=n_0}^{\bar{n}} \frac{(v_y^m \theta_m)/\eta}{mf(m)} \sum_{i=1...I} t_i \frac{\partial x_m^i}{\partial l_m} l_m dF(m)$$

The LHS of the above formula is negative due to the negative definiteness of the Slutsky matrix. Assuming that the sign of $\partial x_n^i / \partial l_n$ is the same for all *n* for each commodity, we have the desired result.

QED

This is an alternative expression of the classic Corlett-Hague (1953) type rule, in that taxes on goods that are complementary with leisure and subsidies on those that are relatively more substitutable for leisure are optimal.

2.2 Income Effects

Regarding the income effects of commodity taxes or subsidies, we establish the following result:

Lemma 2: *Suppose that commodity demands are homogeneous of degree one in post-tax income y. Then, the commodity tax payments per unit of post-tax income,* $\frac{\sum_{i=1...I} t_i x_n^i}{y_n}$*, are decreasing in post-tax incomes (and thus the commodity taxes are regressive) if and only if labor supply increases with skill* *levels.*

Proof: The uncompensated demand function is denoted by $x_n^i = x_n^i(q, y_n, l_n)$. To show how $\sum_{i=1\dots I} t_i \partial x_n^i / \partial y_n$ varies with *n*:

$$(3)\quad \frac{\partial}{\partial n}\left(\sum_{i=1\dots I} t_i \frac{\partial x_n^i}{\partial y_n}\right) = \frac{\partial}{\partial n}\left(\frac{\sum_{i=1\dots I} t_i\, x_n^i}{y_n}\right) = \sum_{i=1\dots I} t_i \frac{\partial x_n^i}{\partial l_n}\frac{dl_n}{dn}\frac{1}{y_n} + \sum_{i=1\dots I} t_i \left(\frac{\partial x_n^i}{\partial y_n} - \frac{x_n^i}{y_n}\right)\frac{dy_n/dn}{y_n}$$

$$= \sum_{i=1\dots I} t_i \frac{\partial x_n^i}{\partial l_n}\frac{dl_n}{dn}\frac{1}{y_n}$$

We have shown that $\sum_{i=1\dots I} t_i \frac{\partial x_n^i}{\partial l_n} < 0$. So the sign of $\frac{\partial}{\partial n}\left(\sum_{i=1\dots I} t_i \frac{\partial x_n^i}{\partial y_n}\right)$ is opposite to that of $\frac{dl_n}{dn}$. Notice that y_n is increasing in *n* when l_n is. Therefore, the desired conclusion holds. *QED*

The homogeneity assumption in Lemma 2 is in Mirrlees (1976).[5] Mirrlees (1976, Section 5), on the top of weak separability of the utility with respect to labor supply, assumed that demands are homogeneous of degree one in y_n. In the specification of $u = U(a(x_n, n) + b(nl_n, n))$, he shows that commodity tax payments per unit of post-tax income, $\frac{\sum_{i=1\dots I} t_i x_n^i}{y_n}$, are increasing in y_n (eq. (113)). However, Crawford et al. (2011) showed that commodity demands are not separable from labor supply, with child care as a complement to labor supply. Here, we go back to the original uni-dimentional case *without* weak separability, a core case of the Corlett-Hague-Christiansen.

[5] Notice that this assumption is different from Deaton (1979) who maintains weak separability in preferences. We show here that homogeneity is compatible with non-separability between consumption and labor. Suppose, for example, that the consumption of each good relates to the leisure time $\bar{l} - l_n$, and we have $u = \sum_i \alpha_i (\bar{l} - l_n)^{\beta_i} (x_m^i)^{\gamma} / \gamma - l_n^{\gamma^l} / \gamma^l$. Some goods with $\beta_i < 0$ are complementary to labor, like child care in Crawford et al. (2011).

Commodity taxes are employed for efficiency reasons, as redistribution is achieved through nonlinear income taxation. At the first glance, labor supply looks like a signal of different skills, but at the same time, the goods which are complementary to labor are subsidized. As a result, when preferences are not weakly separable, the commodity tax system imposes a lower tax burden per unit of post-tax income on individuals who supply more labor. Assuming away the possibility that individuals reduce labor supply as wages rise, the commodity tax system is unambiguously regressive when tax burdens are measured relative to disposable income. The lemma does not state if higher income people pay higher commodity taxes, i.e., the sign of $\sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial y_n}$ is not certain.

Assumption A3 was adopted here since the optimally differentiated commodity taxes in the unidimensional case by Christiansen (1984), Jacobs and Boadway (2014) and others do not have definite results with respect to different income elasticities of the commodities. As one recent example, Nishimura (2026) considered the case of $u = u\left(h^a({x_n}^a), {x_n}^b, l_n\right)$, with a set of *partially separable* commodities in group *a* exhibiting different income elasticities. Consumers choose the expenditure y_n^a on group *a*,[6] where $x_n^i(q, l_n, y_n) = \tilde{x}_n^i(q^a, y_n^a(q, y_n, l_n))$. Nishimura (2026) shows that, if $\frac{\partial \tilde{x}_n^i}{\partial y_n^a}$ is identical across $n \in [\underline{n}, \bar{n}]$ where income elasticities can differ,[7] then it is optimal *not* to differentiate commodity taxes according to the differing income effects, and the tax rates depend on $\frac{\partial y_n^a}{\partial l_n}$, the change of total

[6] It is convenient to divide the optimization of labor supply and consumption into three stages. In the first stage, type n with labor supply l_n earns labor income $z_n = nl_n$ and the after-tax income $y_n = z_n - T(z_n)$. In the consumption stage, (s)he first chooses $y_n^a = \sum_A q_i^a x_n^i$ and $y_n^b \equiv y_n - y_n^a$ for the consumption of x^a and x^b respectively. Then the consumption is decided in respective groups.

[7] Like the Linear Engel curve $\tilde{x}_n^i = \chi^i + \gamma^i(y_n^a - \sum_A q_i^a \chi^i)/q_i^a$ (*A* means group *a*).

propensity to consume goods in group a with respect to labor supply. Here, $\frac{\partial x_n^i}{\partial y_n} = \frac{\partial \tilde{x}_n^i}{\partial y_n^a}\frac{\partial y_n^a}{\partial y_n}$, and this is a suggestive case since the LHS of (2) for the separated goods is proportional to $\frac{\partial \tilde{x}_n^i(q^a, y_n^a)}{\partial y_n^a}$ as well as $\frac{\partial x_n^i}{\partial l_n} = \frac{\partial \tilde{x}_n^i}{\partial y_n^a}\frac{\partial y_n^a}{\partial l_n}$ in the RHS. In words, those with higher income effects have both higher substitution effects and higher labor-incentive effects, which make the impacts to (2) neutral. As a result, $\frac{\partial y_n^a}{\partial l_n}$, group a's complementarity or substitutability with labor, decides the tax rates regardless of the income elasticities of each good.

3. The marginal burden on labor income of the low-income and the highest income

The optimal total tax wedge on labor income in (1) is rewritten as a Diamond (1998)-Saez (2001) type of the ABC formula.

Let ρ be non-labor income, $\epsilon_l^I = (1 - T'(z_n))\frac{\partial z_n}{\partial \rho} < 0$[8] be the income effect of z_n, $\epsilon_{zT'}^* = \frac{\partial ln z_n^{C*}}{\partial ln(1-T'(z_n))} > 0$ be the compensated tax elasticity of z_n, and $A(n) \equiv 1 + \frac{1+\epsilon_l^I}{\epsilon_{zT'}^*} > 0$ (Seade (1982)).

Let $C(n)$ be the inverse of the ability Pareto weight and $C(n) = \frac{1-F(n)}{nf(n)}$. For $z_n = nl_n$, the optimal allocation generates the income distribution $H(z_n) = F(n)$ in the area where z_n is increasing in n, assuming the Spence-Mirrlees condition $z_{n_1} \geq z_{n_2}$ for all n_1 and $n_1 > n_2$, and $C_z(z_n) =$

[8] $y_n = z_n - T(z_n) + \rho$. The derivative is evaluated at $\rho = 0$.

$\frac{1-H(z_n)}{z_n dH(z_n)/dz_n}$,[9] $A(n)C(n) = \frac{1}{\epsilon^*_{zT'}} C_z(z_n)$.

We have (Jacobs and Boadway (2014, Proposition 2), Mirrlees (1976, eq. (96))):

$$\mathcal{W}_n \equiv \frac{T'(z_n)}{1-T'(z_n)} + \sum_{i=1...I} t_i \left(\frac{1}{(1-T'(z_n))n} \frac{\partial x_n^i}{\partial l_n} + \frac{\partial x_n^i}{\partial y_n} \right) = A(n)B(n)C(n) \quad (4)$$

For the term $B(n)$, taxing 1 Euro from type m reduces social marginal welfare by $g_m = \Psi'(v_m)v_y^m/\eta$, normalized by the Lagrange multiplier η. The reduction of y_m also reduces tax revenue by $\tau_x^m = \sum_{i=1...I} t_i \frac{\partial x_m^i}{\partial y_m}$, and income effects of the labor supply accumulates in an incentive-compatible manner by $D_{nm} = exp\left[\int_n^m \frac{-\epsilon_l^I A(m')}{m'} dm'\right]$ (tax-perturbation approach of Saez (2001)). Then

$$B(n) \equiv \int_{m=n}^{\bar{n}} (1 - g_m - \tau_x^m) \frac{D_{nm} dF(m)}{1-F(n)} \quad (5)$$

is the average social marginal value of taxing 1 Euro from type m, taking into account its effect on the revenue through the income effects of labor supply and commodity demands.

In (1) and (4), the term related to the commodity tax burden is decomposed by the Slutsky Equation $\frac{1}{(1-T'(z_n))n} \frac{\partial x_n^{C*,i}}{\partial l_n} = \frac{1}{(1-T'(z_n))n} \frac{\partial x_n^i}{\partial l_n} + \frac{\partial x_n^i}{\partial y_n}$ into total effects and income effects. Alternatively, it can be expressed as $\frac{dx_n^i}{dz_n} = \frac{\partial x_n^i}{\partial z_n} + \frac{\partial x_n^i}{\partial y_n}\frac{\partial y_n}{\partial z_n} = \frac{\partial x_n^i}{n\,\partial l_n} + \frac{\partial x_n^i}{\partial y_n}\left(1 - T'(z_n)\right)$ as a total effect of efficient-unit labor supply on commodity demands.

Lemma 2 shows that several fundamental properties of optimal income taxation remain valid under non-separable preferences and in the presence of optimal commodity taxes and subsidies. We assume that there is no range in which higher-income individuals supply less labor, so that the income effect

[9] Income effect of total consumption is written as $\frac{\partial y_n}{\partial \rho} = \left(1 - T'(z_n)\right)\frac{\partial z_n}{\partial \rho} + 1 = \epsilon_l^I + 1$. The Agent Monotonicity condition by Seade (1982) is implied under normality of total consumption $\frac{\partial y_n}{\partial \rho} \geq 0$.

of commodity taxes τ_x^m is decreasing in *n*:[10]

We first note that if commodity demands are homogeneous of degree one in y_n, then the optimal marginal labor tax wedge is positive for all income levels. The proof of this property is standard.[11] The limiting tax rates at $z_n = 0$ and $z_{\bar{n}}$ are derived later. As noted above, the assumption $\theta_n > 0$ for all n serves as a reinforcing condition that holds at the optimum.

Corollary 1: *Suppose that commodity demands are homogeneous of degree one in* y_n*, and* $\frac{dl_n}{dn} > 0$. *Assume also that the elasticities of labor supply (both the substitution and income effects) are constant. Then, total tax wedges on labor income are increasing when the skill distribution follows a Pareto form and when the income elasticity of labor is sufficiently small in its absolute value. Moreover, above the critical skill level* n_C *at which* $g_{n_C} + \tau_x^{n_C} = 1$*, total labor tax wedges on labor income are decreasing at the skill levels where* $nf(n)$ *is rising.*

Proof: We only need to prove that the first statement is true. Using an alternative definition of *B(n)*$= \int_{m=n}^{\bar{n}} \left(1 - g_m - \tau_x^m - \mathcal{W}_m \epsilon_l^I\right) \frac{dF(m)}{1-F(n)}$ by Jacobs and Boadway (2014), it is increasing in *n* if:

$$1 - g_n - \tau_x^n < \int_{m=n}^{\bar{n}} \left(1 - g_m - \tau_x^m - \mathcal{W}_m \epsilon_l^I\right) \frac{dF(m)}{1 - F(n)} \left(1 + A(n)C(n)\epsilon_l^I\right)$$

[10] An alternative case when the result of Corollary 1 holds is when $\sum_{i=1\ldots I} t_i \frac{\partial x_m^i}{\partial y_m} < 0$ for all *m* (goods are overall subsidized) and $\mathcal{W}_{\bar{n}}|_{optimum} - \frac{T'_{\bar{n}}}{1-T'_{\bar{n}}}|_{optimum, t_i=0, i=1\ldots I} > 0$ along with A1-A4 that constitutes Proposition 2 below.

[11] The transversality condition implies that $B(\underline{n}) = \int_{m=\underline{n}}^{\bar{n}} \left(1 - \Psi'(v_m) \frac{\partial v_m}{\partial y_n} / \eta - \sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial y_n}\right) D_{\underline{n}m} dF(m) = 0$, and our Lemma 2 shows that $1 - g_m - \tau_x^m$ is non-decreasing in *n*. Therefore, as in Mirrlees (1971) and Seade (1982), $B(n)$ is increasing in *n* for low *n*, and decreasing for high *n* going towards 0 at $\bar{n}$.

We have $1 - g_n - \tau_x^n < \int_{m=n}^{\bar{n}} (1 - g_m - \tau_x^m) \frac{dF(m)}{1-F(n)}$ and the term in the last bracket is assumed to be positive (see footnote 13). *QED*

As in Diamond (1998), the result hinges on conditions on skill distribution, elasticities and social objective. Those assumptions that warrant Lemma 2 and additional assumptions A4-A6 result in Diamond's (1998) conclusion. More specifically, when the skill density $f(n)$ rises above n_C until the mode, the decreasing marginal tax rates apply from n_C above the modal skill level. At least in this range, if $n\big(1 - T'(z_n)\big)$ increases, $\frac{dl_n}{dn} > 0$ holds as reinforcing assumption.[12] The Pareto distribution is often applied to the highest income, where Atkinson (1995) shows, in the absence of commodity taxes, *constant* marginal tax rates when $g_m = 0$ for high incomes, including the Rawlsian social welfare function. In our case, the income effects τ_x^m implies that the marginal income tax rates rise when the income elasticity of labor is sufficiently small in its absolute value.

For the highest income level $z_{\bar{n}}$, assuming that, as in Saez (2001), both the marginal income tax and the demand for goods asymptotically converge at the highest income level (therefore, we do not impose a no-distortion property as Jacobs and Boadway (2014) did), we employ an alternative definition of

[12] Using a usual decomposition, $\sum_{i=0\ldots I} q_i x_n^i \leq n(1 - T'(z_n))l_n - (T(z_n) - z_n T'(z_n))$ is the budget constraint. If $n(1 - T'(z_n))$ increases in *n*, assuming that $T''(z_n) < 0$ (expected in the range where $\mathcal{W}_n$ is decreasing), then the substitution effect suggests that labor supply creases in *n*. The "virtual" non-labor income $-T(z_n) + z_n T'(z_n)$ varies across individuals by $\frac{\partial(-T(z_n) + z_n T'(z_n))}{\partial n} = T''(z_n)\frac{dz_n}{dn}$. $\frac{dz_n}{dn} \geq 0$ due to the Spence-Mirrlees condition, then income effect (the non-labor income decreasing in *n* with $-\epsilon_l^I > 0$) reinforces the substitution effect to have $\frac{dl_n}{dn} > 0$.

the progressivity index $B(n)$ by Jacobs and Boadway (2014), $\mathcal{W}_{\bar{n}}A(\bar{n})^{-1}C(\bar{n})^{-1} = B(\bar{n}) = lim_{n\to\bar{n}} \int_{m=n}^{\bar{n}}\left(1-\Psi'(v_m)v_y^m/\eta - \mathcal{W}_m\epsilon_l^I - \tau_x^m\right)\frac{dF(m)}{1-F(n)} = 1-g_{\bar{n}} - \mathcal{W}_{\bar{n}}\epsilon_l^I - \tau_x^{\bar{n}}$, so we have:[13]

$$\mathcal{W}_{\bar{n}} = \frac{1-g_{\bar{n}}-\tau_x^{\bar{n}}}{A(\bar{n})^{-1}C(\bar{n})^{-1}+\epsilon_l^I}$$

In the case where a fraction of people $F(n_0)$ do not work, the optimal marginal income wedge for type n_0 is expressed by replacing $A(n_0)$ with 1 (or applying Piketty and Saez (2013, Appendix A.2));

$$\mathcal{W}_{n_0} = \frac{1}{C(n_0)}\int_{m=n_0}^{\bar{n}}(1-g_m-\tau_x^m)\frac{D_{n_0 m}dF(n_0)}{1-F(n_0)}$$

We compare (4) for each type *n*, in two cases with and without commodity taxation. In doing so, similar to Mirrlees (1976, pp. 351-352), we assume that the terms involving elasticities, namely $A(n)$and D_{nm}, do not change at n after the introduction of commodity taxes.[14]

4 The Effects of Commodity Taxes on Labor Tax Wedges

We now derive whether the introduction of commodity taxes/subsidies increases the commodity-tax inclusive labor wedge $\mathcal{W}_n$. We show the following: (i) under the Rawlsian SWF in which the changes in g_n $(n > \underline{n})$ are negligible, the introduction of commodity taxes (subsidies) decreases (increases) marginal income wedges. (ii) Also under the Rawlsian objective, the absolute impact is larger for lower-income individuals. (iii) We then demonstrate that result (i) extends to a general social welfare

[13] For $(C_z(z_{\bar{n}}))^{-1} > 1$, $\epsilon_{zT'}^* + \epsilon_l^I > 0$ is sufficient for $A(\bar{n})^{-1}C(\bar{n})^{-1} + \epsilon_l^I = \epsilon_{zT}^*\,[(C_z(z_{\bar{n}}))^{-1}-1] + (\epsilon_{zT}^* + \epsilon_l^I)$ to be positive (e.g., Saez (2001, Eq. (9)).

[14] Given that there is no natural clue on the change of the elasticity parameters of the same person in different allocations, our exercise is similar to Mirrlees (1976) who examined how the term with the distributional effect (our *B(n)*) changes, before and after the commodity taxation.

function. We discuss the changes of both $\mathcal{W}_n$ and $T'_n = T'(z_n)$ [15] with the introduction of the commodity taxes.

4.1 Lower Marginal Income Wedge under the Rawlsian Objective

Define:

$$(6) \qquad \mathcal{W}_n|_{optimum} - \frac{T'_n}{1-T'_n}|_{optimum,t_i=0,i=1...I} \equiv \Delta \mathcal{W}_n$$

$$= A(n)C(n)\int_{m=n}^{\bar{n}}\left(-\tau_x^m - \Delta g_m\right)\frac{D_{nm}dF(m)}{1-F(n)}$$

$$= A(n)C(n)\int_{m=\underline{n}}^{n} v_y^n/v_y^m(\tau_x^m + \Delta g_m)\frac{S_{mn}dF(m)}{1-F(n)}$$

as the increase/decrease in the total labor wedges that type n faces after the introduction of the commodity taxes. The last line is from an alternative representation of the optimal tax formula (see, for example, Tuomala (1990, Chapter 6)), where $S_{mn} \equiv exp\left[\int_m^n \frac{l_{m'}v_{yl}^{m'}}{m'v_y^{m'}}dm'\right] = v_y^m/v_y^n(D_{mn})^{-1}$ (Saez (2001, Eq. (26)) is an expression of the term D_{mn} in the traditional mechanism-design approach.

The term $\frac{T'_n}{1-T'_n}|_{optimum,t_i=0,i=1...I}$ represents the optimal marginal tax on labor incomes when commodities are not taxed or subsidized at all. The differences are represented by (i) the commodity tax/subsidy and (ii) the change in marginal utilities of total consumption. As we see, the second term reflects the changes in labor supply and consumption associated with the tax changes. We first take a look at (i):

Proposition 1: *Consider the first term $\tau_x^m = \sum_{i=1...I} t_i \frac{\partial x_m^i}{\partial y_m}$ in (6). The introduction of commodity taxes*

[15] Here, we compare the tax wedge of each type n. Regarding income levels, we can reasonably say that z_n's will increase at least on average after the introduction of optimal commodity taxes, by the encouragement of work efforts in the commodity tax system.

decreases <u>*the optimal total labor wedges*</u> $\mathcal{W}_n$ *for all positive income levels if and only if the commodities are overall taxed, in the sense that a unit increase in income raises total tax receipts.*

Furthermore, the absolute effect is larger for lower-income individuals in the range where the inverse of the income Pareto weight, $C_z(z_n)$, *is decreasing in* n.

Proof: Given our assumption that the sign of the income effects is the same for all individuals, $\tau_x^m > 0$ will decrease the total tax wedge of all working individuals. To show that the term $-A(n)C(n)\int_{m=n}^{\bar{n}} \tau_x^m \frac{D_{nm}dF(m)}{1-F(n)} = -\frac{C_z(z_n)}{\epsilon_{zT'}^*}\int_{m=n}^{\bar{n}} \tau_x^m \frac{D_{nm}dF(m)}{1-F(n)}$ is greater for lower incomes, we note: in addition to $C_z(z_n)$, the weight $D_{nm} = exp\left[\int_n^m \frac{-\epsilon_l^I A(m')}{m'} dm'\right]$ is decreasing in *n*. In addition, as long as the labor supply increases in *n*, from Lemma 2, the conditional average of the income effects above *n* in (6) is decreasing in incomes, too. *QED*

Incentive-based taxation that penalizes individuals with low labor supply would need to be accompanied by offsetting targeted transfers or appropriate adjustments to the income tax schedule. Proposition 1 guides such tax reform: low-income individuals face a lower (higher) marginal income tax burden when commodity taxation (subsidization) is optimal. In the case of taxing commodities, the tax mix alleviates the labor tax burdens. In the case of subsidization, the stimulation of labor results in, surprisingly, a higher tax on labor at the margin. As such, Proposition 1 shows how the income effects of the supplementary taxes are allocated across income classes. This conclusion can be applied to the regressive sin taxes (Allcott et al., 2019) where consumption of the taxed good is concentrated among poorer households. Our answer to this equity-efficiency trade-off is to reduce income tax burdens progressively.

4.2 Lower Marginal Income Wedge under the General Bergson-Samuelson SWF

Now we turn to the welfare component including $g_m = \Psi'(v_m)v_y^m/\eta$. Using Proposition 1 as a benchmark, our focus here is to analyze how incorporating distributional concerns affects the total tax wedge. Specifically, we examine whether the introduction of commodity taxes (resp. subsidies) increases (resp. decreases) the optimal wedge once welfare effects are taken into account.

We focus on the case in which $\tau_x^m > 0$, since the argument is symmetric when commodities are subsidized ($\tau_x^m < 0$). Furthermore, we examine $\Delta \mathcal{W}_n$ for individuals with $1 - g_n > 0$. This condition indicates that the social marginal utility (in money-equivalent terms) of transferring one unit of income to type n is strictly less than one. In other words, the planner assigns a lower welfare weight to these individuals than the marginal cost of public funds. As in Diamond (1998), an interesting case arises when $1 - g_n > 0$ holds for a large fraction of the working population: the potential welfare recipients lie far below the average of the income distribution.[16]

Plugging $-\Delta g_m - \tau_x^m$ ($m \geq n$ or $m \leq n$) derived in Appendix B into (6), $\Delta \mathcal{W}_n$ is determined by taking into account the cumulative effects of taxes imposed on higher- or lower-income classes. Under

---

[16] From the transversality condition, when $\tau_x^m > 0$, we have $\int_{m=\underline{n}}^{\bar{n}}(1 - g_m)D_{\underline{n}m}dF(n) > 0$ and $D_{\underline{n}m} > 1$ is increasing in m. Therefore, $g_m < 1$ for the majority of the population. The condition of $g_m < 1$ is more stringent in the case of a subsidy in which $\int_{m=\underline{n}}^{\bar{n}}(1 - g_m)dF(m) < 0$.

a Rawlsian social welfare function (where $g_m = 0$ for all $m > \underline{n}$), for example,

$$\Delta \mathcal{W}_{n_0} = \frac{1}{n_0 f(n_0)} \int_{m=n_0}^{\bar{n}} (-\tau_x^m)\, D_{n_0 m} dF(n_0)$$

$$= \frac{1}{n_0 f(n_0)} \Big(\Delta\, g_{\underline{n}} + \tau_x^{n_0} F(n_0) \Big).$$

The first line shows $\mathcal{W}_{n_0} < 0$ when $\tau_x^m > 0$, and $\mathcal{W}_{n_0} > 0$ when $\tau_x^m < 0$, as in Proposition 1.

For further illustration with general Social Welfare Functions, below we show that $\Delta\,\mathcal{W}_n < 0$ is likely for $n < \bar{n}$ when $\tau_x^n > 0$. We assume:

- A1: $\Psi'(v_n) S_{nm}$ and $-\frac{v_{yy}^n}{v_y^n}$ are decreasing in *n,* and $\frac{-\Psi''(v_n) v_n}{\Psi'(v_n)} \frac{v_y^n}{v_n}$ is non-increasing in *n.*
- A2: Commodity demands are normal goods with homogeneous of degree 1 in *y.*
- A3: $\frac{v_{yl}^n}{v_l^n} \leq \frac{-\Psi''(v_n) v_n}{\Psi'(v_n)}$
- A4: $\Delta\eta >\ 0$ when $\tau_x^m > 0$.

The last part of A1 includes Atkinson's CRRA-type SWFs $\Psi'(v_n) = (v_n)^{-d}$ $(d < 1)$ and the Utilitarianism $(\Psi''(v_n) = 0)$, and the decreasing absolute risk aversion includes the changes of labor supply across types. We continue to assume A2. A4 is a natural condition from $\Delta B(\underline{n}) = A(n)C(n) \int_{m=\underline{n}}^{\bar{n}} \Big(-\tau_x^m - \Delta\,(\Psi'(v_m) v_y^m / \eta)\Big) \frac{D_{nm} dF(m)}{1-F(n)} = 0$. The version of the next proposition without A4 is available upon request to the author.

We show the following in Appendix C:

Proposition 2 *Suppose that commodities are overall taxed so that* $\tau_x^m > 0$. *Assume that* $\Delta\mathcal{W}_{\bar{n}} < 0$ *and* *let type* n_1 *denote the highest skill level for which* $\Delta\mathcal{W}_n \geq 0$*. Assume further that* $1-g_{n_1}>0$*. Then, with respect to* $s\epsilon_{gy}^{m} = \partial \ln g_m/\partial y_m$ *and* $s\epsilon_l^I \equiv \partial z_n/\partial\rho/z_n$*, either of the following (7.1) or (7.2) or (7.3) must hold:*

$$(7.1) \quad \frac{\Delta v_{n_1}}{v_y^{n_1}} > \int_{m=\underline{n}}^{n_1} \Psi'(v_m) s\epsilon_{gy}^{m} S_{mn_1}\left[\frac{\Delta v_m}{v_y^m}\right]\frac{dF(m)}{1-F(n_1)}\left(\Psi'(v_{n_1}) s\epsilon_{gy}^{n_1}\right)^{-1}$$

$$(7.2) \quad \left(1-T'_{n_1}\right)\Delta z_{n_1} > \int_{m=\underline{n}}^{n_1} \frac{-s\epsilon_l^I}{\epsilon_{zT'}^*}\Psi'(v_m) S_{mn_1}(1-T'_m)\Delta z_m \frac{dF(m)}{1-F(n_1)}\left(\frac{-s\epsilon_l^I}{\epsilon_{zT'}^*}\Psi'(v_{n_1})\right)^{-1}$$

$$(7.3) \quad \int_{m=n_2}^{n_1}\int_{n=m}^{n_1}\frac{\partial}{\partial n}\left(\frac{\sum_i t_i \partial x_n^{c*,i}}{\partial v_n}(1-g_n)S_{nn_1}\right)dn\frac{dF(m)}{1-F(n_1)} < 0$$

for the compensated demand x_n^{c*} *, where* n_2 *is the skill level that satisfies* $g_{n_2}=1$ *after the introduction of commodity taxes.*

If we cannot find such n_1 *that satisfies either of them, so* $\Delta\mathcal{W}_n < 0$ *for* *<u>all</u>* $n<\bar{n}$ *with* $1-g_n>0$*.*

We discuss below that it is not likely to be able to find such $n_1<\bar{n}$ that satisfy (7.1)-(7.3) as long as $1-g_{n_1}>0$. So the consequence of Proposition 1, $\Delta\mathcal{W}_n<0$, holds.

(7.1) says how $\Delta v_m/v_y^m$ is distributed. First, following Kreiner and Verdelin (2012), after the income tax schedule is adjusted in response to the introduction of commodity taxes, let $\Delta\bar{T}_m$ denote the mechanical change in income tax payments, that is, the increase (or decrease, if negative) in taxes if labor income z_m were held constant. The change in the income tax burden by type m, ΔT_m, is given by:

$$\Delta T_m \approx \Delta\bar{T}_m + T'(z_m)\cdot\Delta z_m$$

We show in Appendix B the following:

(8) $\Delta v_m = v_y^m(-\Delta \bar{T}_m - \sum_i t_i x_m^i)$

The decrease in income tax burden $(-\Delta \bar{T}_m)$ in excess of type *m*'s commodity tax payment $(\sum_i t_i\, x_m^i)$ represents a (possible) dividend arising from a revenue-neutral tax changes. (7.1) says that type n_1 receives $-\Delta \bar{T}_{n_1}$ above $\sum_i t_i\, x_{n_1}^i$ that exceed a weighted average[17] of $\Delta v_m / v_y^m$ for *all* lower incomes (the weights given by the elasticity of social marginal utility). There is no justification for favoring a particular individual over all lower incomes this manner.

(7.2) says that type *n_1* increases the retained income $(1 - T'(z_m))z_m$(the net-of-tax rates are those in post-commodity tax regime) greater than the weighted average of all lower incomes.

As for (7.3), the term $\frac{\sum_i t_i \partial x_n^{c*,i}}{\partial v_n}(1 - g_n)$ represents the marginal increase of the commodity-tax revenue, multiplied by the gross social value of taxing one additional Euro $1 - g_m = 1 - \Psi'(v_m) v_y^m / \eta$. (7.3) states that this value is decreasing in n; that is, the tax is regressive in that sense.

For $\epsilon_{il} \equiv \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$, $\epsilon_{zn} \equiv \frac{\partial ln\, z_n}{\partial ln\, n} = 1 + \epsilon_{ln}$, $\alpha^n \equiv 1 - \frac{v_{yl}^n v_y^n / v_l^n}{v_{yy}^n} \propto -\epsilon_l^I > 0$, $\xi_n \equiv \frac{\Psi''(v_n) v_y^n}{\Psi'(v_n)} / \frac{v_{yy}^n}{v_y^n} \geq 0$, and $\sigma_n \equiv \frac{1 - T'(z_n)}{1 - T(z_n)/z_n}$, we have (Appendix C):

$$(7.3)' \quad \frac{\partial}{\partial n}\left(\left[\frac{\sum_i t_i \partial\, x_n^{c*,i}}{\partial v_n}(1 - g_n)\right] S_{nn_1}\right)$$

$$= \frac{(1 - g_n)\epsilon_{ln}}{n v_y^n y_n} \sum_i t_i x_n^i \left(\epsilon_{il} + \frac{\epsilon_{zn}}{\epsilon_{ln}} \sigma_n \frac{-y_n v_{yy}^n / v_y^n}{1 - g_n}\left[\alpha^n + (1 - \alpha^n + \xi_n)\frac{g_n}{\epsilon_{zn}}\right]\right) S_{nn_1}$$

If (7.3)' is negative, then for taxed goods, conditional labor elasticity of demand $\epsilon_{il} = \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$ must be sufficiently large relative to $\frac{\epsilon_{zn}}{\epsilon_{ln}}$, multiplied by the coefficient of residual income progression. However,

[17] In fact, $\int_{m=\underline{n}}^{n_1} \Psi'(v_m) s \epsilon_{gy}^m S_{mn_1} \frac{dF(m)}{1 - F(n_1)} \left(\Psi'(v_{n_1}) s \epsilon_{gy}^{n_1}\right)^{-1} > 1$ is multiplied to the right hand side, so that the gap is even larger.

for this condition to be satisfied, ϵ_{il} got to be unrealistically high; with a reasonable income tax burden and even a slightly high tax elasticity of compensated labor-supply elasticity, for example, if $\epsilon^*_{zT'}$=0.55, $\frac{-y_n v^n_{yy}}{v^n_y} = 1.2$, $\epsilon^I_l = -0.1$, $T'(z_n)$=0.3, and $T(z_n)/z_n = 0.2$, $g_n = 0.6$,[18] $\xi_n = 0.25$ (α^n becomes 0.17), ϵ_{il} has to be lower than -5.23. The lower substitution and income effects imply even larger (in negative) threshold values. If this condition is not satisfied, subsidized goods must have unrealistically high ϵ_{il} to make up with.

Proposition 2 for the case of a subsidy, $\tau^n_x < 0$, is stated as follows. If $\Delta \mathcal{W}_{\bar{n}} > 0$, the necessary conditions for type n_1's $\Delta \mathcal{W}_{n_1} \leq 0$ are as follows: (C.3) of Appendix C shows that, either the value $-\frac{\Delta v_{n_1}}{v^{n_1}_y} = \Delta \bar{T}_{n_1} + \sum_i t_i x^i_{n_1}$ (the total tax increase or the welfare loss) or $-\left(1 - T^{'}_{n_1}\right) \Delta z_{n_1}$ (the reduction in gross income) must exceed the weighted average of these terms for all lower-skilled types, Such an outcome is generally implausible and violates incentive compatibility. In essence, it is neither desirable nor feasible for a single type to bear a disproportionate tax burden relative to all lower-income individuals. By contrast, the condition corresponding to (7.3) necessarily holds as long as $\frac{\partial\, l_n}{\partial\, n}$ >0. Therefore, when $\tau^n_x < 0$, we conclude that $\Delta \mathcal{W}_n > 0$ for all $n < \bar{n}$ with $1 - g_n > 0$. When $\tau^n_x < 0$, we have $\sum_{i=1 \ldots I} t_i \frac{1}{(1-T'(z_n))n} \frac{\partial x^i_n}{\partial l_n} + \tau^n_x < 0$ in (4). Thus, it is surprising that the marginal labor tax rate rises even though the labor wedge includes negative terms from the commodity subsidy. Consequently, our

---

[18] Appendix C shows that the term $-\Delta g_{n_1} - \tau^{n_1}_x$ in (7.2)' has to be positive for the premise $\Delta \mathcal{W}_{n_1} > 0$, so the value of $g_{n_1} (< g_n)$ has to be sufficiently high. For g_{n_1}=0.6, $-\Delta \bar{T}_{n_1} / \sum_i t_i x^i_{n_1} > 1.53$. If the required value of $-\Delta \bar{T}_{n_1} / \sum_i t_i x^i_{n_1}$ is lower, then we can have a higher g_{n_1} which demands smaller threshold value of ϵ_{il}.

conclusion that the total tax wedge $\mathcal{W}_n$ increases is stronger than stating that the marginal income tax T'_n may rise.

4.3 Marginal Income Wedges at the Highest Skill

For the highest skilled $\bar{n}$, if the social weight $g_{\bar{n}}$ is sufficiently small so that, in (6), its possible variation after the introduction of commodity taxes, $\Delta g_{\bar{n}}$, is smaller in the absolute value than the commodity tax's income effect $\tau_x^{\bar{n}}$. Then we obtain $\Delta \mathcal{W}_{\bar{n}} < 0$ when goods are taxed.

A possible exception to the above scenario arises when the choice of the numeraire affects the sign of τ_x^m. Specifically, consider a regime a in which $\tau_x^{m,a} > 0$ and another regime b in which $\tau_x^{m,b} < 0$ within the same economy. This implies that a good that is taxed in regime a (denote it by i^*, with $t_{i^*}^a > 0$) becomes untaxed in regime b. Such a situation occurs in the simplest case where $I + 1 = 2$ and $t_1^a > 0$. Then the relative prices are made consistent across regimes: $\frac{T'^a + t_{i^*}^a}{1 + t_{i^*}^a} = T'^b$, $t_i^b = \frac{t_i^a - t_{i^*}^a}{1 + t_{i^*}^a}$.

Appendix D shows that:

$$\mathcal{W}_m^a = \mathcal{W}_m^b \frac{1}{1 + t_{i^*}^a}$$

so $0 > \mathcal{W}_m^a - \frac{T_n'}{1 - T_n'}|_{optimum, t_i = 0, i = 1 \ldots I}$ (from the tax regime) and $\mathcal{W}_m^b - \frac{T_n'}{1 - T_n'}|_{optimum, t_i = 0, i = 1 \ldots I} > 0$ (from the equivalent subsidy regime) are compatible. At the same time, it is also possible that $\mathcal{W}_m^b$ has the same sign as $\mathcal{W}_m^a$. This latter scenario is consistent with Proposition 2. For example, for the saving taxation in the two-period model, with $I + 1 = 2$ goods representing consumption in each period. Suppose that taxing retirement consumption as the leisure complement in regime a—that is, taxing capital income—reduces the effective tax burden on labor income. Then, the equivalent subsidy in regime b, which subsidizes young-period consumption, the labor wedge may either: (i) increase relative to the case without commodity taxation, or (ii) decrease, but by a smaller amount than $\Delta\mathcal{W}_m^a$.

4.4 Marginal Income Tax

In cases where $\Delta\mathcal{W}_n$ is negative (positive) under a commodity tax (subsidy), we refer back to (1) to examine whether the addition of the commodity tax (the second term) increases or decreases the marginal income tax. Namely,

$$\frac{T_n'}{1-T_n'}\Big|_{optimal\ t_i} - \frac{T_n'}{1-T_n'}\Big|_{optimum\ with\ t_i=0} \equiv \Delta\,\frac{T_n'}{1-T_n'}$$

$$= \Delta\mathcal{W}_n - \sum_{i=1\ldots I} t_i\left(\frac{1}{(1-T'(z_n))n}\frac{\partial x_n^i}{\partial l_n} + \frac{\partial x_n^i}{\partial y_n}\right)$$

Recall $\sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial l_n} < 0$ by Lemma 1. Then the case of the subsidy is unambiguous: $\Delta\mathcal{W}_n > 0$ concludes that $\Delta\,\frac{T_n'}{1-T_n'}>0$. Our next proposition deals with the case of taxes.

Proposition 3: *Suppose that the commodities are subsidized and* $\Delta\mathcal{W}_n > 0$*. Then the marginal income tax rates increase for all income levels.*

When the commodities are taxed and $\Delta\mathcal{W}_n < 0$*, the marginal income tax rates decrease if:* (i) *the absolute values of uncompensated elasticities of commodity demands with respect to labor are sufficiently low, and* (ii) *the income elasticities of demands are high, and* (iii) *the coefficient of residual income progression* $\sigma_n = \frac{1-T'(z_n)}{1-T(z_n)/z_n}$ *is high.*

These comparisons are performed using weighted sums, where the weights are given by the tax rates.

Proof: We only need to prove for the case of the taxes. Given $\Delta \mathcal{W}_n < 0$, from (1),

$$\Delta \frac{T_n'}{1-T_n'} \leq -\sum_{i=1\ldots I} t_i \frac{x_n^i}{y_n}\left(\frac{\epsilon_{il}}{\sigma_n} + \frac{\partial x_n^i}{\partial y_n}\frac{y_n}{x_n^i}\right)$$

The sign of the above formula depends on: (i) $-\epsilon_{il}$ (absolute values of elasticities of commodity demands with respect to labor), (ii) $\frac{y_n^i}{x_n^i}\frac{\partial x_n^i}{\partial y_n}$ (the income elasticities of demands). (iii) the coefficient of residual income progression. *QED*

We only used $-\sum_{i=1\ldots I} t_i \frac{x_n^i}{y_n}\left(\frac{\epsilon_{il}}{\sigma_n} + \frac{\partial x_n^i}{\partial y_n}\frac{y_n}{x_n^i}\right)$ in the proof of Proposition 3, but the term $-A(n)C(n)\int_{m=n}^{\bar{n}} \tau_x^m \frac{D_{nm}dF(m)}{1-F(n)} = -\frac{C_z(z_n)}{\epsilon_{zT'}^*}\int_{m=n}^{\bar{n}} \tau_x^m \frac{D_{nm}dF(m)}{1-F(n)}$ used in Proposition 1 may also be useful, especially when the redistributive objective isis close to the Rawlsian.

5. Conclusion

This paper derived the features of the optimal income taxation in a practical scenario in which Atkinson and Stiglitz' (1976) weak separability does not hold. The departure from the optimality of uniform commodity taxes raises highly policy-relevant questions, given that many countries adopt non-uniform consumption taxes (such as Value Added Tax).

Optimal commodity taxes are driven primarily by efficiency considerations, aiming to reduce tax burdens on individuals who supply more labor. Consequently, the commodity tax burden per unit of post-tax income decreases with post-tax income. This regressivity strengthens an increasing average social marginal value of taxing income at higher income levels, supporting higher marginal labor income taxes for high-income individuals. It also leads to lower marginal labor tax rates for middle and low-income levels. Our result and intuition carry to the case of regressive sin taxes (Allcott et al., 2019) where consumption of the taxed good is concentrated among poorer households. Such clear-cut analytical results are obtained due to the uni-dimensionality. This pattern will remain with taste heterogeneity.

If goods are complementary to labor, the commodities are subsidized, which raises marginal labor tax wedges, as the subsidies increase the social marginal value of taxing income. Conversely, when goods are overall substitutes for labor, lower income taxes stimulate labor supply, thereby increasing commodity tax revenue.

Our methodological framework can also be applied to marginal tax reform, complementing the recent analysis of Boadway and Smart (2025), who similarly employ the Slutsky decomposition to study policy reforms involving differential subsidies for child care. The tax-perturbation approaches of Saez (2001) and Jacquet et al. (2013) remain applicable in our setting, ensuring that our results extend to tax reform analysis.[19]

Appendix A: Behavioral Elasticities

Term	Definition	alternative expression

[19] The tax reform analysis is more complicated than in Laroque (2005) and Kaplow (2006) where commodity tax reform and compensated change of post-tax income y_n, since labor supplies l_m (including those with $m \neq n$) have to be adjusted for incentive-compatibility.

Conditional labor elasticity of commodity demand (compensated)	$\frac{\partial ln\, x_n^{i,c*}}{\partial ln\, l_n}$	$= \frac{z_n}{x_n^i}\left(\frac{\partial x_n^i}{n \partial l_n} + \left(1 - T'(z_n)\right)\frac{\partial x_n^i}{\partial y_n}\right) = \frac{z_n}{x_n^i}\left(\frac{\partial x_n^i}{\partial z_n} + \frac{\partial x_n^i}{\partial y_n}\frac{\partial y_n}{\partial z_n}\right) \gtreqless 0$
Conditional labor elasticity of commodity demand (uncompensated)	$\epsilon_{il} \equiv \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$	$\gtreqless 0$
Compensated tax elasticity of labor supply	$\epsilon_{zT'}^* \equiv \frac{\partial ln z_n^{c*}}{\partial ln\left(1 - T'(z_n)\right)}$	$= \frac{-v_l^n / l_n}{-v_{ll}^n + 2v_{yl}^n v_l^n / v_y^n - \left(v_l^n / v_y^n\right)^2 v_{yy}^n} > 0$
Income elasticity of labor supply (ρ is non-labor income: see footnote 8 for ρ.)	$\epsilon_l^I \equiv (1 - T'(z_n))\frac{\partial\, z_n}{\partial\, \rho}$	$= \frac{(v_{yy}^n - v_{yl}^n v_y^n / v_l^n)\left(v_l^n / v_y^n\right)^2}{-v_{ll}^n + 2v_{yl}^n v_l^n / v_y^n - \left(v_l^n / v_y^n\right)^2 v_{yy}^n} < 0$
Uncompensated tax elasticity of labor supply, $1 + \frac{\partial ln\, z_n}{\partial ln\, (1 - T'(z_n))} > 0$ by Seade (1982)	$\frac{\partial ln\, z_n}{\partial ln\left(1 - T'(z_n)\right)}$	$= \frac{(1-T')}{z_n}\left(\frac{\partial z_n^{c*}}{\partial(1-T')} + z_n \frac{\partial\, z_n}{\partial\, \rho}\right) = \epsilon_{zT'}^* + \epsilon_l^I$
Uncompensated wage elasticity of labor supply	$\epsilon_{ln} \equiv \frac{\partial ln\, l_n}{\partial ln\, n}$	$= \frac{n(1 - T')}{l_n}\frac{\partial l_n}{\partial\left(n(1 - T')\right)} = \frac{\partial ln\, z_n}{\partial ln(1 - T')}$
Uncompensated elasticity of earnings supply	$\epsilon_{zn} \equiv \frac{\partial ln\, z_n}{\partial ln\, n}$	$= \frac{n}{n l_n}\left(l_n + n\frac{\partial l_n}{\partial n}\right) = 1 + \epsilon_{ln} > 0$

Appendix B

We first show the following:

$$(B.1) \qquad -\, \Delta\, g_m - \tau_x^m = -(1 - g_m)\tau_x^m + g_m\left(\frac{\Delta\, \eta}{\eta} + s\epsilon_{gy}^m \frac{\Delta v_m}{v_y^m} + \frac{-\partial z_m / \partial \rho / z_m}{\epsilon_{zT'}^*}(1 - T_m')\, \Delta\, z_m\right)$$

where $s\epsilon_{gy}^m \equiv -v_{yy}^m / v_y^m - \Psi''(v_m) v_y^m / \Psi'(v_m)$>0 is a semi-elasticity of social marginal utility g_m. The income effect of labor $\epsilon_l^I = (1 - T'(z_m))\frac{\partial\, z_m}{\partial\, \rho}$ and a compensated elasticity $\epsilon_{zT'}^*$ were introduced above, where the former is written in its semi-elasticity form.

Derivation of (B.1): For type *n*'s marginal utility of post-tax income *y*, its change by the introduction of optimal commodity taxes is written as, with respect to the commodity taxes t_i $(i = 1 \ldots I)$, the post-tax income y_n and gross income z_n,

$$\Delta v_y^n \approx \sum_{i=1\ldots I} v_{q_i y}^n t_i + v_{yy}^n \Delta y_n + v_{yl}^n / n \Delta z_n$$

(notice that the status-quo is no commodity taxes, so there is no Δ sign for t_i's). From Roy's identity, $\sum_{i=1\ldots I} \partial v_{q_i}^n / \partial y_n \, t_i = -\sum_{i=1\ldots I} \partial\left(v_y^n \, x_n^i\right) / \partial y_n t_i = -\sum_{i=1\ldots I}\left(v_{yy}^n x_n^i + v_y^n x_{n,y}^i\right) t_i = -v_{yy}^n \sum_{i=1\ldots I} t_i x_n^i - v_y^n \tau_x^n$. For $g_n \equiv \Psi'(v_n) v_y^n / \eta$,

$$\Delta\left(\Psi'(v_n) v_y^n / \eta\right) \approx \Psi'(v_n) \Delta v_y^n / \eta + \Psi''(v_n) v_y^n / \eta \Delta v_n - \Psi'(v_n) v_y^n / \eta^2 \Delta \eta$$

$$= g_n(-\tau_x^n + v_{yy}^n / v_y^n (\Delta y_n - \textstyle\sum_{i=1\ldots I} t_i x_n^i) + v_{yl}^n / (n v_y^n) \Delta z_n + \Psi''(v_n) v_y^n / \Psi'(v_n)\left(-\Delta \bar{T}_n - \sum_i t_i x_n^i\right)) - g_n \Delta \eta / \eta.$$

So we have:

$$(B.2) \qquad -\Delta g_m - \tau_x^m$$

$$= -(1 - g_m)\tau_x^m + g_m \Delta \eta / \eta - g_m[v_{yy}^m / v_y^m (\Delta y_m - \sum_i t_i x_m^i) + \frac{v_{yl}^m}{m v_y^m} \Delta z_m$$

$$+ \Psi''(v_n) v_y^m / \Psi'(v_n)(-\Delta \bar{T}_m - \sum_i t_i x_m^i))]$$

The part of $-v_{yy}^m / v_y^m\left(\Delta y_m - \sum_i t_i x_m^i\right) - v_{yl}^m / \left(m v_y^m\right) \Delta z_m$ in the above formula is rearranged to:

$$(B.3) - v_{yy}^m / v_y^m \left(\Delta y_m - \sum_i t_i x_m^i\right) - \frac{v_{yl}^m}{m v_y^m} \Delta z_m = \left(-v_{yy}^m + v_{yl}^m v_y^m / v_l^m\right)\left(1 - T'(z_m)\right) \Delta z_m / v_y^m$$

$$+ v_{yy}^m / v_y^m \left(\Delta z_m - \Delta y_m - T'(z_m) \Delta z_m + \sum_i t_i x_m^i\right)$$

$$= -v_{yy}^m / v_y^m \left(-\Delta \bar{T}_m - \sum_i t_i x_m^i\right) + \left(-v_{yy}^m + v_{yl}^m v_y^m / v_l^m\right) / v_y^m \left(1 - T'(z_m)\right) \Delta z_m$$

where we used $m\left(1 - T'(z_m)\right) = -v_l^m / v_y^m$, $\Delta z_m - \Delta y_m = \Delta T_m \approx \Delta \bar{T}_m + T'(z_m) \cdot \Delta z_m$ in the text.

For non-labor income ρ, let $\epsilon_l^I \equiv \left(1 - T'(z_n)\right)\frac{\partial z_n}{\partial \rho} = (v_{yy}^n - v_{yl}^n v_y^n / v_l^n)\left(v_l^n / v_y^n\right)^2 / (-v_{ll}^n + 2v_{yl}^n v_l^n / v_y^n - \left(v_l^n / v_y^n\right)^2 v_{yy}^n) < 0$ be the income effect of (retained) earnings. $y_n = z_n - T(z_n) + \rho$ so $\frac{\partial y_n}{\partial \rho} = \epsilon_l^I + 1$, so $0 > \epsilon_l^I > -1$ for normality of consumption and leisure. For $\epsilon_{zT'}^* = (-v_l^m / l_m) / (-v_{ll}^m + 2v_{yl}^m v_l^m / v_y^m - \left(v_l^m / v_y^m\right)^2 v_{yy}^m)$,

$$\text{(B.4)} \quad -(v_{yy}^m - v_{yl}^m v_y^m / v_l^m) / v_y^m = -\epsilon_l^I \left(v_y^m / -v_l^m\right)(-v_{ll}^m + 2v_{yl}^m v_l^m / v_y^m - \left(v_l^m / v_y^m\right)^2 v_{yy}^m) / (-v_l^m)$$

$$= \frac{-\epsilon_l^I}{\epsilon_{zT'}^*} \frac{1}{(1 - T_m') z_m} = \frac{-\partial z_m / \partial \rho / z_m}{\epsilon_{zT'}^*}$$

Plugging (B.3) into (B.2) derives (B.1).

QED

Derivation of (8): Since the welfare increase is written as, using $-v_l^m / v_y^m = m\left(1 - T'(z_m)\right)$ and $\Delta T_m \approx \Delta \bar{T}_m + T'(z_m) \cdot \Delta z_m$ and $\Delta y_m = (1 - T'(z_m)) \Delta z_m + (-\Delta \bar{T}_m)$:

$$\Delta v_m = \sum_i v_{q_i}^m t_i + v_y^m \Delta y_m + v_l^m / m \Delta z_m$$

$$= -v_y^m \sum_i t_i x_m^i + v_y^m \left(\left(1 - T'(z_m)\right) \Delta z_m - \Delta \bar{T}_m\right) + v_l^m / m \Delta z_m$$

$$= v_y^m \left(-\Delta \bar{T}_m - \sum_i t_i x_m^i\right),$$

QED

Appendix C

Derivation of (7.1) and (7.2) for $\Delta \eta > 0$:

Substituting (B.1) into (6), we can write the distributional effect as, for $\psi_m = g_m / v_y^m$:

$$\Delta B(n)=\int_{m=\underline{n}}^{n} v_y^n\left((1-g_m)\frac{1}{v_y^m}\tau_x^m-\psi_m\frac{\Delta\eta}{\eta}-\psi_m s\epsilon_{gy}^m(-\Delta\bar{T}_m-\sum_i t_i x_m^i)\right.$$

$$\left.-\psi_m\frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*}(1-T_m')\Delta z_m\right)\frac{S_{mn}\,dF(m)}{1-F(n)}$$

Suppose that $n_1<\bar{n}$. $\Delta\mathcal{W}_n<0$ before n_1 and $\Delta\mathcal{W}_{n_1}\geq 0$, so $\Delta g_{n_1}+\tau_x^{n_1}\geq\Delta B(n_1)$ must be the case:

$$(C.1)\qquad (1-g_{n_1})\frac{1}{v_y^{n_1}}\tau_x^{n_1}-\psi_{n_1}\left[\frac{\Delta\eta}{\eta}+s\epsilon_{gy}^{n_1}\left(-\Delta\bar{T}_{n_1}-\sum_i t_i x_{n_1}^i\right)+\frac{-\partial z_{n_1}/\partial\rho/z_{n_1}}{\epsilon_{zT'}^*}(1-T_{n_1}')\Delta z_{n_1}\right]$$

$$\geq\int_{m=\underline{n}}^{n_1}\left((1-g_m)\frac{1}{v_y^m}\tau_x^m\right.$$

$$\left.-\psi_m\left[\frac{\Delta\eta}{\eta}+s\epsilon_{gy}^m(-\Delta\bar{T}_m-\sum_i t_i x_m^i)+\frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*}(1-T_m')\Delta z_m\right]\right)\frac{S_{n_1 m}dF(m)}{1-F(n_1)}$$

As the first possibility for (C.1) to hold, $\frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*}=-(v_{yy}^m-v_{yl}^m v_y^m/v_l^m)/v_y^m\equiv\beta^m s\epsilon_{gy}^m$ (see (B.4)) derives, with $\Delta v_m/v_y^m=-\Delta\bar{T}_m-\sum_i t_i x_m^i$, $\mu_m=\Delta v_m/v_y^m+\left(1-T'(z_m)\right)\Delta z_m$,

$$(C.2)\qquad s\epsilon_{gy}^{n_1}\left(-\Delta\bar{T}_{n_1}-\sum_i t_i x_{n_1}^i\right)+\frac{-\partial z_{n_1}/\partial\rho/z_{n_1}}{\epsilon_{zT'}^*}(1-T_{n_1}')\Delta z_{n_1}$$

$$=s\epsilon_{gy}^{n_1}\left(\beta^{n_1}\mu_{n_1}+(1-\beta^{n_1})\frac{\Delta v_{n_1}}{v_y^{n_1}}\right)$$

$$>\int_{m=\underline{n}}^{n_1}s\epsilon_{gy}^m\left(\beta^m\mu_m+(1-\beta^m)\frac{\Delta v_m}{v_y^m}\right)\frac{\psi_m}{\psi_{n_1}}\frac{S_{mn_1}dF(m)}{1-F(n_1)}$$

$\beta^{n_1}\leq 1$ under A3, $\frac{v_{yl}^{n_1}}{v_l^{n_1}}\leq-\Psi''(v_{n_1})v_y^{n_1}/\Psi'(v_{n_1})$).

The Case of Subsidy $\tau_x^n<0$*:* Suppose that $\Delta\eta<0$, complementary to A4. With the same procedure, if n_1 is the highest skill level that has $\Delta\mathcal{W}_{n_1}\leq 0$,with $-\Delta v_m/v_y^m=\Delta\bar{T}_m+\sum_i t_i x_m^i$, $-\mu_m=-\frac{\Delta v_m}{v_y^m}-\left(1-T'(z_m)\right)\Delta z_m$, the following must hold: For $k_{n_1}=\int_{m=\underline{n}}^{n_1}\Psi'(v_m)s\epsilon_{gy}^m S_{mn_1}\frac{dF(m)}{1-F(n_1)}\left(\Psi'(v_{n_1})s\epsilon_{gy}^{n_1}\right)^{-1}$,

$$(C.3)\qquad -\beta^{n_1}\mu_{n_1} - (1-\beta^{n_1})\frac{\Delta v_{n_1}}{v_y^{n_1}}$$

$$> k_{n_1}\int_{m=\underline{n}}^{n_1} \psi_m s\epsilon_{gy}^m S_{mn_1}\Bigg[-\beta^m \mu_m$$

$$- (1-\beta^m)\frac{\Delta v_m}{v_y^m}\Bigg]\frac{dF(m)}{1-F(n_1)}\left(\int_{m=\underline{n}}^{n_1} \psi_m s\epsilon_{gy}^m S_{mn_1}\frac{dF(m)}{1-F(n_1)}\right)^{-1}$$

Either $-\frac{\Delta v_{n_1}}{v_y^{n_1}}$ or $-\mu_{n_1}$ must exceed the weighted average of the corresponding value multiplied by $k_{n_1} > 1$. This condition can be decomposed to (7.1) and (7.2).

Derivation of (7.3) and (7.3)': The second possibility for (C.1) to hold is the following.

$$\left(1-g_{n_1}\right)\frac{1}{v_y^{n_1}}\tau_x^{n_1} - \int_{m=\underline{n}}^{n_1}\left((1-g_m)\frac{\tau_x^m}{v_y^m}\right)\frac{S_{mn_1}dF(m)}{1-F(n_1)} \leq \ \psi_{n_1}\frac{\Delta\eta}{\eta} - \int_{m=\underline{n}}^{n_1}\psi_m\frac{\Delta\eta}{\eta}\frac{S_{mn_1}dF(m)}{1-F(n_1)}$$

The RHS of the above formula is not positive. For *m* such that $g_m > 1$, $\left(1-g_{n_1}\right)\frac{1}{v_y^{n_1}}\tau_x^{n_1} > (1-g_m)\frac{1}{v_y^m}\tau_x^m$, $\left(1-g_{n_1}\right)\frac{1}{v_y^{n_1}}\tau_x^{n_1} - (1-g_m)\frac{\tau_x^m}{v_y^m}S_{mn_1} > 0$. For the compensated demand x_n^{c*}, $\frac{\partial x_n^i/\partial y_n}{v_y^n} = \frac{\partial x_n^{c*,i}}{\partial v_n}$ from the Slutsky Equation, so we have (7.2) as a necessary condition:

$$\int_{m=n_2}^{n_1}\int_{n=m}^{n_1}\frac{\partial}{\partial n}\left([(1-g_n)\frac{\tau_x^n}{v_y^n}]S_{nn_1}\right)\frac{dF(m)}{1-F(n_1)} < 0$$

Under A2:

$$\frac{\partial}{\partial n}\left(\left[(1-g_n)\frac{\tau_x^n}{v_y^n}\right]S_{nn_1}\right)\left(S_{nn_1}\right)^{-1}$$

$$= \frac{\partial}{\partial n}\left(\left(\frac{1}{v_y^n} - \psi_n\right)\sum_i t_i\frac{x_n^i}{y_n}\right) + \left[(1-g_n)\frac{\tau_x^n}{v_y^n}\right]\frac{\partial S_{nn_1}}{\partial n}\left(S_{nn_1}\right)^{-1}$$

$$= \sum_i t_i\frac{\partial x_n^i}{\partial l_n}\frac{\partial l_n}{\partial n}\frac{1-g_n}{y_n v_y^n} + \sum_i t_i\left(\frac{\partial x_n^i}{\partial y_n} - \frac{x_n^i}{y_n}\right)\frac{\partial y_n}{\partial n}\frac{1-g_n}{y_n v_y^n}$$

$$+ \frac{\tau_x^n}{v_y^n}\left(\frac{-v_{yy}^n}{v_y^n}(1-T'(z_n))\frac{\partial z_n}{\partial n} - \Psi''(v_n)/\eta v_y^n\frac{-v_l^n l_n}{n} - \frac{v_{yl}^m}{v_y^n}\frac{\partial l_n}{\partial n} - \frac{v_{yl}^n}{v_y^n n}l_n(1-g_n)\right)$$

For $\alpha^m = 1 - \frac{v_{yl}^m v_y^m / v_l^m}{v_{yy}^m}$, $\frac{n}{l_n}\frac{\partial l_n}{\partial n} = \epsilon_{ln}$, $\epsilon_{zn} \equiv \frac{n}{z_n}\frac{\partial z_n}{\partial n} = 1 + \epsilon_{ln}$, since $(1-\alpha^n)\left(1 - T'^{(z_n)}\right)n = -\frac{v_{yl}^n}{v_{yy}^n}$, so $-\frac{\tau_x^n}{v_y^n}\frac{v_{yl}^n}{v_y^n n} l_n(\epsilon_{ln} + 1 - g_n) = \frac{\tau_x^n}{v_y^n}\frac{v_{yy}^n}{v_y^n n}\left(1 - T'(z_n)\right)z_n(1-\alpha^n)(\epsilon_{zn} - g_n)$. Therefore, for $\epsilon_{il} \equiv \frac{l_n}{x_n^i}\frac{\partial x_n^i}{\partial l_n}$ and $\frac{\partial y_n}{\partial n} = \left(1 - T'(z_n)\right)\frac{\partial z_n}{\partial n}$ we have:

$$\frac{\partial}{\partial n}\left(\left[(1-g_n)\frac{\tau_x^n}{v_y^n}\right]S_{nn_1}\right)\left(S_{nn_1}\right)^{-1}$$
$$= \sum_i t_i \frac{x_n^i}{y_n}\frac{1}{v_y^n n}\left((1-g_n)\epsilon_{il}\epsilon_{ln} + \frac{-v_{yy}^n}{v_y^n}\left(1 - T'(z_n)\right)z_n\alpha^n\epsilon_{zn}\right.$$
$$\left. + g_n\left(\frac{-v_{yy}^n}{v_y^n}(1-\alpha^n) + \frac{-\Psi''(v_n)v_y^n}{\Psi'(v_n)}\right)\left(1 - T'(z_n)\right)z_n\right)$$

For $(1 - T'(z_n))z_n = y_n\frac{1-T'(z_n)}{1-T(z_n)/z_n}$, the above formula is (7.2)'. From (B.4), $\alpha^n = \frac{-\epsilon_l^I}{\epsilon_{zT'}^*}\left(\frac{1-T'(z_n)}{1-T(z_n)/z_n}y_n\left(-v_{yy}^n\right)/v_y^n\right)^{-1}$.

QED

Appendix D:

Consider regime *a*, and transform it to regime *b* with an equivalent allocation in which one of the taxed goods (denote it by i^* with $t_{i^*}^a$>0) becomes untaxed regime *b*. Then the relative prices are made consistent across regimes: $\frac{T'^a + t_{i^*}^a}{1+t_{i^*}^a} = T'^b$, $t_i^b = \frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}$.

Since $y_m^b = y_m^a/(1+t_{i^*}^a)$, $\frac{\partial x_m^{i,b}}{\partial y_m^b}\left(\left(\frac{1+t_i^a}{1+t_{i^*}^a}\right)_{i=1,\dots,I}, l_m, \frac{y_m^a}{1+t_{i^*}}\right)\frac{1}{1+t_{i^*}} = \frac{\partial x_m^{i,a}((1+t_i^a)_{i=1,\dots,I}, l_m, y_m^a)}{\partial y_m^a}$, so $\sum_i \frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,b}}{\partial y_m^b} = \sum_i t_i^a\frac{\partial x_m^{i,a}}{\partial y_m^a} + \sum_{i=1\dots I} t_{i^*}^a t_i^a\frac{\partial x_m^{i,a}}{\partial y_m^a} - t_{i^*}^a$, meaning that

$$\left(1 - T'^b\right)\sum_i \frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,b}}{\partial y_m^b} = (1 - T'^a)\sum_i t_i^a\frac{\partial x_m^{i,a}}{\partial y_m^a} - \left(1 - T'^b\right)t_{i^*}^a$$

$\frac{\partial x_m^{i,b}}{\partial l_m} = \frac{\partial x_m^{i,a}}{\partial l_m}$ implies that $\sum_{i=1\dots I}\frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,b}}{\partial l_m} = \sum_{i=1\dots I}\frac{t_i^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,a}}{\partial l_m} + \sum_{i=1\dots I} t_i^a\frac{t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,a}}{\partial l_m} = \sum_{i=1\dots I} t_i^a\frac{\partial x_m^{i,a}}{\partial l_m}$.

So $1 - T'^b - \sum_{i=1\dots I}\frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,b}}{n\partial l_m} - \left(1 - T'^b\right)\sum_i \frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}\frac{\partial x_m^{i,b}}{\partial y_m^b}$

$= 1 - T'^b - \sum_{i=1...I} t_i^a \frac{\partial x_m^{i,a}}{n\partial l_m} - (1 - T'^a) \sum_i t_i^a \frac{\partial x_m^{i,a}}{\partial y_m^a} + \left(1 - T'^b\right) t_{i^*}^a$

$$= 1 - T'^a - \sum_{i=1...I} t_i^a \frac{\partial x_m^{i,a}}{n\partial l_m} - (1 - T'^a) \sum_i t_i^a \frac{\partial x_m^{i,a}}{\partial y_m^a}$$

Therefore, $\mathcal{W}_m^b(1 - T'^b) = \mathcal{W}_m^a(1 - T'^a) = \mathcal{W}_m^a(1 - T'^b)(1 + t_{i^*}^a)$. *QED*

References

Allcott, H., Lockwood, B., Taubinsky, D., (2019), Regressive Sin Taxes, with an Application to the Optimal Soda Tax, Quarterly Journal of Economics, 134, 1557–1626.

Atkinson, A., Stiglitz, J. (1976), The Design of Tax Structure: Direct versus Indirect Taxation. Journal of Public Economics, 6, 55–75.

Atkinson, A., (1995), Public Economics in Action: The Basic Income/Flat Tax Proposal, The Lindahl Lectures, Oxford University Press.

Boadway, R., (2012), From Optimal Tax Theory to Tax Policy: Retrospective and Prospective Views (Munich Lectures in Economics), Cambridge, MIT Press.

Boadway, R., Smart, M., (2025), Distributionally Sensitive Cost-Benefit Analysis, CESifo Working Paper No. 11692.

Boadway, R., Pestieau, P., (2003), Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem, in Arnott, Greenwald, B., Kanbur, R., and Nalebuff, B., (eds.), Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz (Cambridge, MA: MIT Press), 387-403.

Christiansen, V., (1984), Which Commodity Taxes Should Supplement the Income Tax? Journal of

Public Economics, 24, 195-220.

Corlett, W., Hague, D., (1953), Complementarity and The Excess Burden of Taxation. Review of Economic Studies, 21, 21–30.

Crawford, I., Keen, M., Smith, S., (2011), Value Added Tax and Excises, Chapter 4 in Dimensions of Tax Design: The Mirrlees Review, Edited by Institute for Fiscal Studies (IFS) and James Mirrlees, Oxford University Press, Oxford.

Deaton, A. 1979. Optimally Uniform Commodity Taxes, Economics Letters, 2, 357-361.

Diamond, P.A., (1998), Optimal Income Taxation: an Example with a U-shaped Pattern of Optimal Marginal Tax Rates, American Economic Review 88, 83–95.

Doligalski, P., Dworczak, P. Akbarpour, M., Kominers, S., (2025), Optimal Redistribution via Income Taxation and Market Design, Discussion Paper 25/787, University of Bristol School of Economics.

Edwards, J., Keen, M., Tuomala, M., 1994. Income tax, Commodity Taxes and Public Good Provision: a Brief Guide. FinanzArchiv, 51, 472–487.

Ferey, A., Lockwood, B., Taubinsky, D., (2024), Sufficient Statistics for Nonlinear Tax Systems with General Across - Income Heterogeneity, American Economic Review, 114, 3206–49.

Jacobs, B., Boadway, R., (2014), Optimal Linear Commodity Taxation under Optimal Nonlinear Income Taxation, Journal of Public Economics, 117, 201–210.

Jacquet, L., Linden, V., Lehmann, E., (2013), The Optimal Marginal Tax Rates with both Extensive and Intensive Responses, Journal of Economic Theory, 148, 1770-1805.

Kaplow, L., (2006), On the Undesirability of Commodity Taxation Even When Income Taxation Is Not optimal, Journal of Public Economics, 90, 1235–1250.

Kaplow, L., (2010), Taxing Leisure Complements, Economic Inquiry, 48, 1065-1071.

Kreiner, C., Verdelin, N., (2012), Optimal Provision of Public Goods: A Synthesis, The Scandinavian Journal of Economics, 114, 384-408.

Laroque, G. (2005), Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof, Economics Letters, 87, 141–144.

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

Mirrlees, J., (1976), Optimal Tax Theory: a Synthesis. Journal of Public Economics, 6, 327–358.

Nava, M., Schroyen, F., Marchand, M., (1996), Optimal Fiscal Policy and Expenditure in a Two Class Economy, Journal of Public Economics, 61, 119–137.

Nishimura, Y., (2026), Linear Commodity Tax Reform and Optimal Commodity Taxes under Partial Separability, Discussion Papers In Economics And Business, University of Osaka, Discussion Paper 25-09-Rev.2 www2.econ.osaka-u.ac.jp/econ_society/dp/2509R2.pdf

Piketty, T., Saez, E., (2013), Optimal Labor Income Taxation, Handbook of Public Economics Volume 5, Chapter 7, 391-474. Elsevier.

Saez, E., (2001), Using Elasticities to Derive Optimal Income Tax Rates. Review of Economic Studies, 68, 205–229.

Saez, E., (2002), The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes, Journal of Public Economics, 83, 217–230.

Seade, J., (1982), On the Sign of the Optimum Marginal Income Tax. Review of Economic Studies, 49, 637-643.

Spiritus, K., (2025), Optimal Commodity Taxation when Households Earn Multiple Incomes, International Tax and Public Finance, 32 98–119.

Tuomala, M., (1990), Optimal Taxation and Redistribution, Oxford, Clarendon Press.

Tuomala, M., (2016), Optimal Redistributive Taxation, Oxford, Oxford University Press.