Discussion Papers In Economics And Business

Supplementary Commodity Taxes, Labor Tax on the Middle Class, and the Tax-mix

Yukihiro Nishimura

Discussion Paper 25-08

August 2025

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

Supplementary Commodity Taxes, Labor Tax on the Middle Class, and the Tax-mix

Yukihiro Nishimura[+*]

August 2025

Abstract

Given that the incentive consideration reduces the scope for redistribution, Mirrlees (1976, Optimal tax theory: a synthesis. Journal of Public Economics 6, 327–358) emphasized the redistributive effects of commodity taxes (which include capital income tax), which reduces the effective tax wedge on labor income. We revert to the unidimensional case to show that the optimal labor wedge can become higher after the introduction of the optimal commodity taxes/subsidies when labor complements are subsidized. This is partly because supplementary commodity taxes are not increasing in ability as Mirrlees (1976) thought. Among the classic results, decreasing marginal taxes on the middle class, including the mode of the income distribution, remains valid with commodity taxes and without separability in the utility function.

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

JEL Classification: H21, H24, D63

[+] The author is grateful for financial support from MEXT/JSPS KAKENHI JP23K22126.
The authors have no financial or proprietary interests in any material discussed in this article.
[*] Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka 560-0043, Japan. E-mail: ynishimu@econ.osaka-u.ac.jp

1. Introduction

The tax mix of labor income and consumption is a controversial issue. The Atkinson-Stiglitz (1976) separability is merely a benchmark, and heterogeneity across income classes exist in, for example, saving (treating current and future consumption as two different commodities). The optimal mix of nonlinear income taxation and linear commodity taxation is a complex issue which we address in this paper. Mirrlees (1976) discussed in which direction the optimal income tax rates go if commodity taxes are introduced. He concludes, under special assumptions of the utility function, that the effective marginal tax on labor (by the total of marginal income tax and commodity tax increases) should be lower with the introduction of the optimal commodity tax (p. 350). His reason is that supplementary commodity tax tends to increase (and in the case of the negative tax, subsidy receipts tend to decrease) tax burdens with ability, which helps to reduce the total tax wedge on labor income. Subsequently, Jacobs and Boadway (2014) elaborates on the optimal tax formulas in Diamond (1998)-Saez (2001) type of expressions. However, as in the discrete versions in Edwards et al. (1994) and Nava et al. (1996), Jacobs and Boadway have not concluded whether optimal non-linear income taxes are higher or lower under optimal linear commodity taxes.

In this paper, we fully characterize whether optimal non-linear income taxes are higher or lower under optimal linear commodity taxes. Mirrlees (1976) made particular assumptions on preferences for commodity consumption under which commodity taxes are "effective weapons" (Mirrlees (1971), p. 206) to supplement taxing the rich. Subsequent works found that the optimal income taxes are much more progressive than Mirrlees (1971), and it is worthwhile to reexamine how the optimal labor income tax reacts when leisure complements are taxed and leisure substitutes are subsidized, without assuming an ability-preference relationship as in Mirrlees (1976).

We first show that the burden of optimal commodity taxes/subsidies per disposable income is decreasing in ability, contrary to Mirrlees' (1976) taste heterogeneity model. Among the classic results, decreasing marginal taxes on the middle class including the mode of the income distribution (Diamond (1998, Proposition 3)) remains valid with commodity taxes and without separability in the utility function. We further show that the optimal labor wedge becomes higher after the introduction of the optimal commodity taxes when labor complements are subsidized. The opposite scenario of decreasing optimal income taxes occurs when leisure complements are taxed.

2. Revisiting and Revising Classic Results

Individuals of skill level $n \in [\underline{n}, \bar{n}]$ whose labor supply is l_n earn the labor income $z_n = nl_n$, and out of the after-tax income $y_n \equiv z_n - T(z_n)$, they choose the commodity consumption x_i^n ($i = 0 \dots I$, I+1 is the number of goods with the numeraire good $c = x_0$ to be untaxed).

The government maximizes a Bergson-Samuelson Social Welfare Function (SWF), $\int_{m=\underline{n}}^{\bar{n}} \Psi(v_m)\, dF(m)$ subject to the resource constraint $\int_{m=\underline{n}}^{\bar{n}} \{ml_m - \sum_{i=0 \dots I} x_m^i(q, y_m, l_m) - R\}\, dF(m)$ for an exogenous revenue requirement $R > 0$, and the conventional incentive constraint. Let η be the Lagrange multiplier of the resource constraint.

The marginal income wedge type n faces under the labor income tax $T(z_n)$ and linear commodity taxes t_i $(i = 1 \dots I)$ is as follows (Jacobs and Boadway (2014, eq. (25)), Mirrlees (1976, eq. (96))):

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

The issue discussed in Mirrlees (1976) is whether the introduction of commodity taxes/subsidies increases the commodity-tax inclusive labor wedge $\mathcal{W}_n$ or not. In this paper, we discuss the changes of both $\mathcal{W}_n$ and $T'_n = T'(z_n)$[1] with the introduction of the commodity taxes.

In (1), the first term is the effective marginal income tax. The second term associates the commodity tax burden with the additional unit of labor income. The total of the terms in parentheses is the compensated demand elasticity, decomposed by the Slutsky Equation. Alternatively, it can be expressed as $\frac{dx_n^i}{dz_n} = \frac{\partial x_n^i}{\partial z_n} + \frac{\partial x_n^i}{\partial y_n}\frac{\partial y_n}{\partial z_n} = \frac{\partial x_n^i}{n\,\partial l_n} + \frac{\partial x_n^i}{\partial y_n}\left(1 - T'(z_n)\right)$ (and dividing by $1 - T'(z_n)$ to align with $\frac{T'(z_n)}{1-T'(z_n)}$), the total effect of labor supply on the goods' demands. The sign of $\partial x_n^i / \partial l_n$ represents the good's complementary with work. Assuming that its sign would be the same for all n, we first show the following:

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

Proof: With respect to the Hicksian derivative s_n^{ij} of good j's demand with respect to good i's price,

[1] 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.

the optimal commodity taxes satisfy[2] (Jacobs and Boadway (2014, eq. (31)))

$$(2) \qquad \int_{m=\underline{n}}^{\bar{n}} \sum_{j=1\ldots I} t_j s_n^{ij} \, dF(m) = \int_{m=n_0}^{\bar{n}} \frac{\mathcal{W}_m}{A_m} \frac{\partial x_m^i}{\partial l_m} l_m dF(m)$$

for $i = 1, \ldots, I$, where A(m) > 0 is defined below in (4). Multiplying each equation by t_j and summing up, we have:

$$\int_{m=\underline{n}}^{\bar{n}} \sum_{i=1\ldots I} t_i \sum_{j=1\ldots I} s_n^{ij} t_j \, dF(m) = \int_{m=n_0}^{\bar{n}} \frac{\mathcal{W}_m}{A_m} \sum_{i=1\ldots I} t_i \frac{\partial x_m^i}{\partial l_m} l_m dF(m)$$

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

The implication of Lemma 1 is that, given the desire for income redistribution, if the government aims for the same rate of the total tax wedge on labor income (by the total of income and commodity taxes), the introduction of commodity taxes/subsidies *increases* the marginal income tax rates T_n'.

On the income effects of the commodity taxes (or subsidies), we show the following:

Lemma 2: *Suppose that commodity demands are homogeneous of degree 1 in the after-tax income*

---

[2] A small caveat is that, unlike the total tax wedge on labor in (1), the RHS of (2) is an uncompensated derivative of the conditional demands. In the counterpart of 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 represents the demand difference through the reduced labor efforts of the mimicker.

y. Then commodity tax burdens per post-tax income are <u>*lower*</u> *(and with the case of negative tax, the subsidy receipts are* <u>*higher*</u>*) in ability if and only if the labor supply increases with ability.*

Proof: Under income and commodity taxes, the uncompensated demand function is denoted by $x_n^i = x_n^i(q, y_n, l_n)$. To show how $\sum_{i=1...I} t_i \partial x_n^i / \partial y_n$ varies with n:

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

$$= \sum_{i=1...I} t_i \frac{\partial x_n^i}{\partial l_n}\frac{\partial l_n}{\partial n}\frac{1}{y_n}$$

We have shown that $\sum_{i=1...I} t_i \frac{\partial x_n^i}{\partial l_n} < 0$. So the sign of $\frac{\partial}{\partial n}\left(\sum_{i=1...I} t_i \frac{\partial x_n^i}{\partial y_n}\right)$ is opposite to that of $\frac{\partial l_n}{\partial n}$.

QED

The homogeneity assumption (for illustrative purposes) in Lemma 2 is the same as Mirrlees (1976). Mirrlees (1976) showed his progressivity result under strong assumptions that commodity demands differ by tastes across abilities (conditioned by incomes) and that they are independent of labor supplies. His results are overturned to conclude Lemma 2 when the commodity demands are dependent on labor supplies, a core reason for supplementary commodity taxes.[3] The commodity tax

---

[3] It may be informative to write $\frac{\partial x_n^i}{\partial l_n}\frac{\partial l_n}{\partial n} = \frac{\partial x_n^i}{\partial l_n}\left(-\frac{l_n}{n} + \frac{\partial z_n/\partial n}{n}\right)$. The first part, $\frac{\partial x_n^i}{\partial l_n}\left(-\frac{l_n}{n}\right)$, is the differentiation of $x_n^i(q, y_n, z_n/n)$ with respect to n given z_n, which Mirrlees (1976) focuses on. However, in the second term, the fact that z_n being increasing in *n*, together with Lemma 1, overturns its sign.

burden becomes progressive only in the area where labor income does not increase as much as ability, where the tax motive on higher ability coincides with the tax on leisure-complement.

3. Displaying the ABC Formula

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

Let $F(n)$ be the distribution function of ability, with the density function *f(n).* Individual optimization generates the income distribution $H(z_n(n)) = F(n)$, assuming the Spence-Mirrlees condition $z_n(n_1) \geq z_n(n_2)$ for all n_1 and $n_2 < n_1$.

For $C(n)$ to be the inverse of the ability Pareto-weight, and $C_z(z_n) = \frac{1-H(z_n)}{z_n dH(z_n)/dz_n}$, $\epsilon_l^I = (1 - T'(z_n))\frac{\partial z_n}{\partial \rho}$ to be the income effect of z_n with respect to the non-labor income ρ, $\epsilon_{zT'}^{*} = \frac{\partial ln z_n^{C*}}{\partial ln\,(1-T'(z_n))}$ to be the compensated tax elasticity of z_n, we have (Jacobs and Boadway (2014, Proposition 2)):

(4) $$\mathcal{W}_n = A(n)B(n)C(n)$$

$$A(n) \equiv 1 + \frac{1+\epsilon_l^I}{\epsilon_{zT'}^{*}} > 0 \text{ (Seade (1982))},\ C(n) \equiv \frac{1-F(n)}{nf(n)},\ A(n)C(n) = \frac{1}{\epsilon_{zT'}^{*}} C_z(z_n)$$

As for *B(n)*:

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

$g_m = \Psi'(v_m) v_y^m / \eta$ = the social marginal utility of additional y_m to type *m,* normalized by the Lagrange multiplier.

$\tau_x^n = \sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial y_n}$, $D_{nm} = exp\left[\int_n^m \frac{-\epsilon_l^I A(m')}{m'} dm'\right]$,

B(n) is the average social marginal value of taxing 1 Euro above type *n*, taking into account its effect on the revenue through the income effects of labor supply and commodity demands.

Lemma 2 shows that Diamond's (1998) Proposition 3 remains valid under non-separable preferences and the presence of the optimal commodity taxes and subsidies:

Corollary 1: *Suppose that commodity demands are homogeneous of degree 1 in* y_n*, and* $\frac{\partial l_n}{\partial n} > 0$*. Above the critical skill level* n_2 *at which* $g_{n_2} + \tau_x^{n_2} = 1$*, total tax wedges on labor income are decreasing where the elasticities of labor supply (both the substitution and income effects) are constants and nf(n) rising with skill.*

For $\bar{n}$, assuming that both the marginal income tax and the demand for goods asymptotically converge at the highest income level, using an alternative definition of *B(n)* by Jacobs and Boadway (2014), $\mathcal{W}_{\bar{n}} A(\bar{n})^{-1} C(\bar{n})^{-1} = B(\bar{n}) = lim_{n\to\bar{n}} \int_{m=n}^{\bar{n}} (1 - \Psi'(v_m) v_y^m/\eta - \mathcal{W}_m \epsilon_l^I - \tau_x^m) \frac{dF(m)}{1-F(n)} = 1 - g_{\bar{n}} - \mathcal{W}_{\bar{n}} \epsilon_l^I - \tau_x^{\bar{n}}$, so we have:[4]

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

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));

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

[4] 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)).

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.

Term	definition	alternative expression
Conditional labor elasticity of commodity demand (compensated)	$\frac{\partial ln\, x_n^{i,c*}}{\partial ln\, l_n}$	$= \frac{z_n}{x_n^i}\left(\frac{\partial x_n^i}{n\partial l_n} + \left(1 - T'(z_n)\right)\frac{\partial x_n^i}{\partial y_n}\right) = \frac{z_n}{x_n^i}\left(\frac{\partial x_n^i}{\partial z_n} + \frac{\partial x_n^i}{\partial y_n}\frac{\partial y_n}{\partial z_n}\right)$
Conditional labor elasticity of commodity demand (uncompensated)	$\epsilon_{il} \equiv \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$	
Compensated tax elasticity of labor supply	$\epsilon_{zT'}^* \equiv \frac{\partial ln z_n^{c*}}{\partial ln\left(1 - T'(z_n)\right)}$	$= \frac{-v_l^n/l_n}{-v_{ll}^n + 2v_{yl}^n v_l^n/v_y^n - \left(v_l^n/v_y^n\right)^2 v_{yy}^n}$
Income elasticity of labor supply (ρ is non-labor income)	$\epsilon_l^I \equiv (1 - T'(z_n))\frac{\partial z_n}{\partial \rho}$	$= \frac{(v_{yy}^n - v_{yl}^n v_y^n/v_l^n)\left(v_l^n/v_y^n\right)^2}{-v_{ll}^n + 2v_{yl}^n v_l^n/v_y^n - \left(v_l^n/v_y^n\right)^2 v_{yy}^n}$
Income effect of total consumption	$\frac{\partial y_n}{\partial \rho}$	$= \left(1 - T'(z_n)\right)\frac{\partial z_n}{\partial \rho} + 1 = \epsilon_l^I + 1$
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	$\epsilon_{ln} \equiv \frac{\partial ln\, l_n}{\partial ln\, n}$	$= \frac{n(1 - T')}{l_n}\frac{\partial l_n}{\partial(n(1 - T'))} = \frac{\partial ln\, z_n}{\partial ln\,(1 - T')}$
Uncompensated elasticity of earnings supply	$\epsilon_{zn} \equiv \frac{\partial ln\, z_n}{\partial ln\, n}$	$= \frac{n}{nl_n}(l_n + n\frac{\partial l_n}{\partial n}) = 1 + \epsilon_{ln}$

Table 1: Behavioral Elasticities

3.1 Differential of the Marginal Income Wedges

Define:

(7) $$\mathcal{W}_n|_{optimum} - \frac{T_n'}{1-T_n'}|_{t_i=0,i=1\dots I} \equiv \Delta\mathcal{W}_n$$

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

as the increase/decrease in the total labor wedges that type n faces after the introduction of the commodity taxes. The term $\frac{T_n'}{1-T_n'}|_{t_i=0,i=1\dots I}$ 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 payment/subsidy receipts and (ii) the change in marginal utilities of 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: *Focus on the first term* $\tau_x^m = \sum_{i=1\dots I} t_i \frac{\partial x_m^i}{\partial y_m}$ *in (7). The introduction of commodity taxes decreases the optimal total labor wedges* $\mathcal{W}_n$ *of all positive income levels if and only if the commodities are overall taxed (in the sense that a unit income increase yields an increase in tax receipts). Its impact (in absolute terms) is greater for lower incomes.*

Proof: Given our assumption that the sign of the income effects is the same for all individuals, $\tau_x^m > 0$ will decrease (increase) the total tax wedge of all working individuals. To show that the term of $-A(n)C(n)\int_{m=n}^{\bar{n}}\tau_x^m\frac{D_{nm}dF(m)}{1-F(n)} = -\frac{C_z(z_n)}{\epsilon_{zT'}^*}\int_{m=n}^{\bar{n}}\tau_x^m\frac{D_{nm}dF(m)}{1-F(n)}$ is greater for lower incomes, we note: the inverse of the income Pareto weight $C_z(z_n)$ and the weight $D_{nm} = exp\left[\int_n^m \frac{-\epsilon_i^I A(m')}{m'}dm'\right]$ are higher with the lower 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 (7) goes higher with lower incomes, too. *QED*

The same result is obtained with n_0 in (6), but only if n_0 is invariant. n_0 may or may not increase by the introduction of commodity taxes.

3.2 Lower Marginal Income Wedge under the Rawlsian Objective

Now we discuss the part including $g_m = \Psi'(v_m)v_y^m/\eta$. Setting Proposition 1 as a benchmark, our focus is on how the introduction of the welfare effects will ever increase (resp. decrease) the total tax wedge after the introduction of commodity taxes (resp. subsidies). Before we proceed, we introduce one notation, following Kreiner and Verdelin (2012). After the income tax system gets adjusted by the introduction of commodity taxes, let $\Delta\bar{T}_m$ be the increase (decrease if it is negative) in the income tax payment *if* z_m is not changed. The change of the income tax payment by type *m*, ΔT_m, is approximated by the mechanical change of the system ($\bar{T}_m$) and the individual's behavioral change:

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

We discuss the case of $\tau_x^m > 0$, since the argument is completely symmetric in the case of the subsidy. Further, we examine $\Delta\mathcal{W}_n$ for people with $1 - g_n > 0$. It means that a (money-equivalent) social marginal utility of transferring a unit income to type *n* is less than 1. As in Diamond (1998), an interesting case is that this value is positive for a large fraction of the labor force (a potential welfare recipient is a way below the population average).[5] We show in Appendix A that:

[5] 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 $\int_{m=\underline{n}}^{\bar{n}}(1-g_m)dF(m) < 0$.

(8) $- \Delta g_m - \tau_x^m$

$$= -(1-g_m)\tau_x^m + g_m\left(\frac{\Delta\eta}{\eta} + s\epsilon_{gy}^m\left(-\Delta\bar{T}_m - \sum t_i x_m^i\right) + \frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*}(1-T_m')\,\Delta z_m\right)$$

(7)' $\quad \Delta\mathcal{W}_n = A(n)C(n)\int_{m=\underline{n}}^{n} v_y^n/v_y^m\left(\tau_x^m + \Delta g_m\right)\frac{S_{mn}dF(m)}{1-F(n)}$

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. (7)' is an alternative way to write the optimal tax formula (e.g., 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, Equation (26)) is an expression of the term D_{mn} in the traditional mechanism-design approach.

The term $\tau_x^m(1-g_m) = \sum_{i=1...I} t_i\frac{\partial x_m^i}{\partial y_m}(1-g_m)$ is the net contribution of type *n* to the fiscal budget through the commodity-tax payments.[6] As we mentioned above, In comparing the marginal income taxes, the terms involving elasticities, namely, *A(n)* and $v_y^n/v_y^m S_{mn}$, do not change at n after the introduction of commodity taxes, which is similar to Mirrlees (1976, pp. 351-352).[7]

---

[6] $\sum_{i=1...I} t_i\frac{\partial x_m^i}{\partial y_m}$ is a contribution to the tax revenue from an additional unit of income. In turn, this tax revenue, net of the social marginal utility g_m, will contribute to the fiscal budget.

[7] 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.

In the third term of (8), the mechanical reduction of the income tax ($-\Delta \bar{T}_m$) in excess of type *m*'s commodity tax payment ($\sum_i t_i x_m^i$) is a (possible) dividend from a revenue-neutral tax changes. Since the welfare increase is written as, using $-v_l^m/v_y^m = m(1-T'(z_m))$:

$$\text{(9)} \qquad \Delta v_m = \sum_i v_{q_i}^m t_i + v_y^m \Delta y_m + v_l^m/m \Delta z_m$$
$$= -v_y^m \sum_i t_i x_m^i + v_y^m\big((1-T'(z_m))\Delta z_m - \Delta \bar{T}_m\big) + v_l^m/m \Delta z_m$$
$$= v_y^m \left(-\Delta \bar{T}_m - \sum_i t_i x_m^i\right),$$

The objective of tax reforms is to yield as high values of these with as little increase in z_m's as possible.

increase in the income tax payment *if* z_m is not changed after the introduction of optimal commodity taxes	$\Delta \bar{T}_m$
the increase in income tax that type *m* pays after the introduction of optimal commodity taxes	$\Delta T_m = \Delta \bar{T}_m + T'(z_m) \cdot \Delta z_m$
semi-elasticity of social marginal utility $g_m = \Psi'(v_m)v_y^m/\eta$	$s\epsilon_{gy}^m \equiv -v_{yy}^m/v_y^m - \Psi''(v_m)v_y^n/\Psi'(v_m) = \partial \ln g_m/\partial y_m > 0$
the Lagrange multiplier of the optimal-tax problem	η ($\Delta\eta > 0$ or $\Delta\eta < 0$)
increase in well-being after the introduction of commodity taxes	$\mu_m^T \equiv -\Delta \bar{T}_m - \sum_i t_i x_m^i \propto \Delta v_m$
increase in disposal income after the introduction of commodity taxes	$\mu_m^y \equiv -\Delta y_m - \sum_i t_i x_m^i$

Table 2 indicators of welfare changes

Plugging (8) ($m \geqq n$) into (7) (or $m \leqq n$ into (7)'), $\Delta \mathcal{W}_n$ is determined, taking into account on the cumulative effects of taxes of the upper-/lower-income classes. Also, the labor tax wedge of type *n* increases in the transfers $\left(-\Delta \bar{T}_m\right)$ and gross income $\left(\Delta z_m\right)$ of the higher-income people $m \geqq n$. For the Rawlsian SWF ($g_m = 0$ for all $m > \underline{n}$), for example,

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

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

The first line shows $\mathcal{W}_{n_0} < 0$, and $\mathcal{W}_n < 0$ for all $n > n_0$[8] as in Proposition 1.

3.3 Lower Marginal Income Wedges under General Bergson-Samuelson SWF

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 $-\frac{v_{yy}^n}{v_y^n}$ [9] 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 homogeneous of degree 1 in y.
- A3: $\frac{v_{yl}^{n_1}}{v_l^{n_1}} \leq -\Psi''(v_{n_1}) v_y^{n_1} / \Psi'(v_{n_1})$

A1 includes Atkinson's CRRA-type SWFs $\Psi'(v_n) = (v_n)^{-d}$ $(d < 1)$ and the Utilitarianism ($\Psi''(v_n) = 0$). A2 is adopted in Mirrlees (1976).

We show the following in Appendix B:

Proposition 2 (i) *Suppose that* $\Delta \mathcal{W}_{\bar{n}} < 0$ *and type* n_1 *is the highest skill level who has* $\Delta \mathcal{W}_n \geq 0$, *and* $1 - g_{n_1} > 0$. *Then, with respect to* $\mu_m^T \equiv -\Delta \bar{T}_m - \sum_i t_i x_m^i$ *and* $\mu_m^y = \Delta y_m - \sum_i t_i x_m^i$, *either of the*

---

[8] For $n = \bar{n}$, see below.

[9] See Appendix A.

following (10.1) or (10.2) must hold.

For $k_{n_1}=\int_{m=\underline{n}}^{n_1}\Psi'(v_m)s\epsilon_{gy}^{m}S_{mn_1}\frac{dF(m)}{1-F(n_1)}\left(\Psi'(v_{n_1})s\epsilon_{gy}^{n_1}\right)^{-1}>1$, $\beta^m=\frac{1}{s\epsilon_{gy}^m}\left(\frac{-v_{yy}^m}{v_y^m}+\frac{v_{yl}^m}{v_l^m}\right)\propto-\epsilon_l^I>0$, *some values* $s_{n_1}\geq s_m\geq 0$ *defined in Appendix B such that* (a) *when* $\Delta\eta<0, max\{\Delta y_{n_1},-\Delta\bar{T}_{n_1}\}>\sum_i t_i x_{n_1}^i+s_{n_1}$, (b) *when* $\Delta\eta>0$, $s_{n_1}=0=s_m$*:*

(10.1)
$$\beta^{n_1}\mu_{n_1}^{y}+(1-\beta^{n_1})\mu_{n_1}^{T}-s_{n_1}>k_{n_1}\int_{m=\underline{n}}^{n_1}\Psi'(v_m)s\epsilon_{gy}^{m}S_{mn_1}\left[\beta^m\mu_m^y+(1-\beta^m)\mu_m^T-s_m\right]\frac{dF(m)}{1-F(n_1)}\left(\int_{m=\underline{n}}^{n_1}\Psi'(v_m)s\epsilon_{gy}^{m}S_{mn_1}\frac{dF(m)}{1-F(n_1)}\right)^{-1}$$

That is, type n_1 *receives* $-\Delta\bar{T}_{n_1}$ *(reduction in mechanical tax burden) or* Δy_{n_1} *(increase in disposable income), above the commodity tax payment* $\sum_i t_i x_{n_1}^i$, $k_{n_1}>1$ *times greater than a weighted average of* μ_m^T *or* μ_m^y *of all lower incomes.*

For $\iota(\eta)$ *such that* $\iota(\eta)=0$ *when* $\Delta\eta<0$ *and* $\iota(\eta)=1$ *when* $\Delta\eta>0$*:*

(10.2)
$$\int_{m=n_2}^{n_1}\int_{n=m}^{n_1}\frac{\partial}{\partial n}\left(\frac{\sum_i t_i\partial x_n^{c*,i}}{\partial v_n}(1-g_n)-\iota(\eta)\Psi'(v_n)\frac{\Delta\eta}{\eta^2}\right)dn\frac{S_{nn_1}dF(m)}{1-F(n_1)}<0$$

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. That is, the net contribution through the commodity-tax payments by sacrificing one unit of utility,* $\frac{\sum_i t_i\partial x_n^{c*,i}}{\partial v_n}(1-g_n)$, *is not increasing in* n *(the tax system is not progressive).*

In either case (i) *or case* (ii), *we cannot find such* n_1 *that satisfy them, so* $\Delta W_n<0$ *for* *<u>all</u>* $n<\bar{n}$ *with* $1-g_n>0$.

(10.1) says how $\mu_m^T = -\Delta \bar{T}_m - \sum_i t_i x_m^i$, as an indicator of well-being, is distributed. Notice that $\Delta y_m = (1 - T'(z_m))\Delta z_m + (-\Delta \bar{T}_m)$ so $\mu_m^y = \mu_m^T + (1 - T'(z_m))\Delta z_m$. When $\Delta\eta < 0$, type n_1 increases the disposable income at least $\sum_i t_i x_{n_1}^i + s_{n_1}$ ($s_{n_1} > 0$), with or without his (her) increased effort; namely, $-\Delta \bar{T}_{n_1} > \sum_i t_i x_{n_1}^i + s_{n_1}$ or $\Delta y_{n_1} > \sum_i t_i x_{n_1}^i + s_{n_1}$, but type $m \leqq n_1$ may *or may not* receive more than $\sum_i t_i x_m^i + s_m$ ($0 < s_m < s_{n_1}$). (10.1) says that type n_1 receives $-\Delta \bar{T}_{n_1}$ or earns Δy_{n_1}, above $\sum_i t_i x_{n_1}^i + s_{n_1}$, $k_{n_1} > 1$ times greater than a weighted average of μ_m^T or μ_m^y of *all* lower incomes (the weight by the elasticity of social marginal utility). Given that μ_m^T is directly related to the welfare, there is no reason to favor a particular individual above all lower incomes this way. When type n_1 becomes worse-off after the introduction of commodity taxes, for lower incomes, the RHS of (10.1) must be negative (and it should not be). We conclude that (10.1) is not likely to hold at the social optimum.

The term $\frac{\sum_i t_i \partial x_n^{c*,i}}{\partial v_n}(1 - g_n)$ in (10.2) is the net contribution through the commodity-tax payments by sacrificing one unit of utility (see footnote 6). If this vale is not increasing in n (not progressive),then income tax may increase the (marginal) tax burden on type n_1.

For $\epsilon_{il} \equiv \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$, $\epsilon_{zn} \equiv \frac{\partial ln\, z_n}{\partial ln\, n} = 1 + \epsilon_{ln}$, $\alpha^n \equiv 1 - \frac{v_{yl}^n v_y^n / v_l^n}{v_{yy}^n} > 0$, $\xi_n \equiv \frac{\Psi''(v_n) v_y^n}{\Psi'(v_n)} / \frac{v_{yy}^n}{v_y^n} \geq 0$, and $\sigma_n \equiv \frac{1 - T'(z_n)}{1 - T(z_n)/z_n}$, we have, as a condition related to (10.2) (Appendix B):

$$(10.2)' \qquad \frac{\partial}{\partial n}\left([\frac{\sum_i t_i \partial\, x_n^{c*,i}}{\partial v_n}(1 - g_n) - \iota(\eta)\Psi'(v_n)\frac{\Delta\eta}{\eta^2}] S_{nn_1} \right)$$

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

For (10.2)' to be negative, for taxed goods, we need to have high labor elasticity of demand ϵ_{il}, relative to the ratio of earning elasticity ϵ_{zn} to uncompensated wage elasticity of labor ϵ_{ln}, multiplied by the coefficient of residual income progression. However, for this condition to hold, conditional labor elasticity of uncompensated commodity demand, $\epsilon_{il} = \frac{\partial ln\, x_n^i}{\partial ln\, l_n}$, got to be unrealistically high; for example, $\epsilon_{il} < -5.23$ if $\epsilon^*_{zT'}$=0.55, $\frac{-y_n v_{yy}^n}{v_y^n} = 1.2$, $\epsilon_l^I = -0.1$, $T'(z_n)$=0.3, and $T(z_n)/z_n = 0.2$, $g_n = 0.6$,[10] $\xi_n = 0.25$ (α^n becomes 0.17). The lower substitution and income effects give even larger (in negative) threshold values. If this condition is not satisfied, subsidized goods must have unrealistically high ϵ_{il} to make up.

Proposition 2 for the case of the subsidy τ_x^n<0 is as follows. Necessary conditions for type n_1's $\Delta \mathcal{W}_{n_1} \le 0$ are: (1) (B.3) of Appendix B shows that, either the value $-\mu_{n_1}^T = \Delta \bar{T}_{n_1} + \sum_i t_i x_{n_1}^i$ (the total tax increase, or the indicator of the welfare decrease) or $-\mu_{n_1}^y = -\mu_{n_1}^T - (1 - T'(z_{n_1}))\Delta z_{n_1}$ (the total tax increase net of your effort) is (with multiple k_{n_1} in (10.1)) greater than its weighted average of all the lower skilled. The issues are whether such tax burdens on a particular type are desirable and incentive-compatible. (2) (10.2)' is always negative (the condition corresponding to (10.2) is necessarily violated) as long as $\frac{\partial\, l_n}{\partial\, n}$ >0.

[10] The threshold value $\epsilon_{il} < -5.23$ is the case of $\Delta \eta < 0$. Appendix B shows that the term $-\Delta g_{n_1} - \tau_x^{n_1}$ in (8) 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, $-\Delta \bar{T}_{n_1} / \sum_i t_i x_{n_1}^i > 1.53$. If the required value of $-\Delta \bar{T}_{n_1} / \sum_i t_i x_{n_1}^i$ is lower, then we can have a higher g_{n_1} which demands smaller threshold value of ϵ_{il}. $\Delta \eta > 0$ yields smaller threshold values too.

3.3 Marginal Income Wedges at the Highest Skill

For $\bar{n}$, we simply assume that the social weight $g_{\bar{n}}$ is 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 $\tau_x^{\bar{n}}$. Then we have $\Delta \mathcal{W}_{\bar{n}} < 0$ when goods are taxed.

Alternatively, the case when the choice of the numeraire affects the sign of τ_x^m is intriguing. Compare regime a with $\tau_x^{m,a} > 0$ and regime b with $\tau_x^{m,b} < 0$ in the same economy. It means that one of the taxed goods (call i^* with $t_{i^*}^a$>0) becomes untaxed in the latter, and it happens when I+1=2 and $t_{i^*}^a \neq 0$. Then the relative prices are made consistent across regimes: $\frac{T'^a + t_{i^*}^a}{1+t_{i^*}^a} = T'^b$, $t_i^b = \frac{t_i^a - t_{i^*}^a}{1+t_{i^*}^a}$.

It is straightforward to show that:

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

so $0 > \mathcal{W}_m^a - \frac{T_n'}{1-T_n'}|_{t_i=0, i=1\ldots I}$ (from the tax regime) and $\mathcal{W}_m^b - \frac{T_n'}{1-T_n'}|_{t_i=0, i=1\ldots I} > 0$ (from the equivalent subsidy regime) are compatible. On the other hand, it is possible that $\mathcal{W}_m^b$ have the same sign as $\mathcal{W}_m^a$. This possibility is consistent with Proposition 2. For example, for the saving taxation in the two-period model (I+1=2 for each period's consumption), suppose that taxing the leisure complement (regime *a,* taxing capital income when the retirement period's consumption is the leisure complement) eases the effective burdens on labor income. Then, the equivalent subsidy on the young-period consumption would either: (i) levy a higher tax wedge on labor income than in the case of no consumption tax, or (ii) ease burdens on labor income but less extent than $\Delta \mathcal{W}_m^a$.

In the cases at which $\Delta\mathcal{W}_n$ is negative (positive) under commodity tax (subsidy), we go back to (1) to see the addition of the commodity-tax (the second term) increases or decreases the marginal income taxes. Namely,

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

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

Recall that the uncompensated elasticity has $\sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial l_n} < 0$ by Lemma 1. Then the case of the subsidy $\sum_{i=1\ldots I} t_i \frac{\partial x_n^i}{\partial y_n} < 0$ 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 go lower if (and only if):* (i) *the absolute values of uncompensated elasticities of commodity demands with respect to labor are (sufficiently) low,* (ii) *the income elasticities of demands are high, and* (iii) *the coefficient of residual income progression is high.*

The comparisons are either by each taxed item or their weighted average 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 $\frac{1-T'(z_n)}{1-T(z_n)/z_n}$. *QED*

In the proof, we use the lowest amount for the decrease/increase of the marginal income tax rate, given that we do not know how much one can use out of $-\tau_x^m$ $(m > n)$ in (7) for type *n*'s marginal income tax rate. Proposition 1 said that the decrease in the marginal tax rate (increase in the subsidy case) is more likely with lower incomes.

4. Conclusion

The patterns of the increase/decrease in the marginal income wedge after the introduction of commodity taxes differ significantly from those of Mirrlees (1976). Contrary to Mirrlees (1976), it can increase after the commodity taxes, since (i) when the labor complements (leisure complements) are subsidized (taxed), there arises a scope to increase total the tax wedge on labor income, (ii) supplementary commodity taxes are not increasing in ability as Mirrlees (1976) thought, (iii) in the case of optimal subsidy, the additional tax on labor income is required for revenue neutrality, (iv) the responses of labor supply from reduced marginal taxes, as well as the mechanical changes of the revenue-neutral changes, are interrelated, so a taxpayer's increased labor positively affects the marginal income tax for the lower incomes.

Appendix A:

Derivation of (8): 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 $(i = 1 \dots I)$, the post-tax income y_n and gross income z_n,

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

(notice that the status-quo is no commodity taxes, so there is no Δ sign for t_i's). From Roy's identity, $\sum_{i=1\dots I} \partial v_{q_i}^n / \partial y_n \, t_i = -\sum_{i=1\dots I} \partial (v_y^n \, x_n^i) / \partial y_n \, t_i = -\sum_{i=1\dots I} \left(v_{yy}^n x_n^i + v_y^n x_{n,y}^i \right) t_i = -v_{yy}^n \sum_{i=1\dots I} t_i x_n^i - v_y^n \tau_x^n$. From (9) we have, for $g_n \equiv \Psi'(v_n) v_y^n / \eta$,

$$\Delta (\Psi'(v_n) v_y^n / \eta) \approx \Psi'(v_n) \Delta v_y^n / \eta + \Psi''(v_n) v_y^n / \eta \, \Delta v_n - \Psi'(v_n) v_y^n / \eta^2 \, \Delta \eta$$
$$= g_n (-\tau_x^n + v_{yy}^n / v_y^n (\Delta y_n - \textstyle\sum_{i=1\dots I} t_i x_n^i) + v_{yl}^n / (n v_y^n) \Delta z_n + \Psi''(v_n) v_y^n / \Psi'(v_n) \left(-\Delta \bar{T}_n - \textstyle\sum_i t_i x_n^i \right)) - g_n \Delta \eta / \eta.$$

So we have:

$$(A.1) \quad -\Delta g_m - \tau_x^m$$
$$= -(1 - g_m) \tau_x^m + g_m \Delta \eta / \eta - g_m [v_{yy}^m / v_y^m (\Delta y_m - \sum_i t_i x_m^i) + \frac{v_{yl}^m}{m v_y^m} \Delta z_m$$
$$+ \Psi''(v_n) v_y^m / \Psi'(v_n) (-\Delta \bar{T}_m - \sum_i t_i x_m^i))]$$

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

$$(A.2) \quad -v_{yy}^m / v_y^m (\Delta y_m - \sum_i t_i x_m^i) - \frac{v_{yl}^m}{m v_y^m} \Delta z_m = (-v_{yy}^m + v_{yl}^m v_y^m / v_l^m) \left(1 - T'(z_m)\right) \Delta z_m / v_y^m$$
$$+ v_{yy}^m / v_y^m (\Delta z_m - \Delta y_m - T'(z_m) \Delta z_m + \sum_i t_i x_m^i)$$
$$= -v_{yy}^m / v_y^m (-\Delta \bar{T}_m - \sum_i t_i x_m^i) + (-v_{yy}^m + v_{yl}^m v_y^m / v_l^m) / v_y^m \left(1 - T'(z_m)\right) \Delta z_m$$

where we used $m(1 - T'(z_m)) = -v_l^m / v_y^m$, $\Delta z_m - \Delta y_m = \Delta T_m \approx \Delta \bar{T}_m + T'(z_m) \cdot \Delta z_m$ 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 $\epsilon_{zT'}^* = (-v_l^m / l_m) / (-v_{ll}^m + 2v_{yl}^m v_l^m / v_y^m - \left(v_l^m / v_y^m\right)^2 v_{yy}^m)$,

$$\text{(A.3)}\quad -(v_{yy}^m - v_{yl}^m v_y^m / v_l^m) / v_y^m = -\epsilon_l^I\left(v_y^m / -v_l^m\right)(-v_{ll}^m + 2v_{yl}^m v_l^m / v_y^m - \left(v_l^m / v_y^m\right)^2 v_{yy}^m) / (-v_l^m)$$

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

Plugging (A.2) into (A.1) derives (8).

QED

A Sufficient Condition for A1:

For a sufficient condition for $\partial / \partial m\left(-v_{yy}^m / v_y^m\right) < 0$, we have, for $\alpha^m \equiv 1 + \frac{v_{yl}^m v_y^m / (-v_l^m)}{v_{yy}^m} \propto -\epsilon_l^I \in (0,1]$:

$$\partial / \partial m\left(-v_{yy}^m / v_y^m\right) = \partial\left(-v_{yy}^m / v_y^m\right) / \partial y_m (1 - T')\partial z_m / \partial m - \partial l_m / \partial m / v_y^m \left(\frac{-v_{yy}^m v_{yl}^m}{v_y^m} + v_{yyl}^m\right)$$

$$= -v_l^m / \left(v_y^m\right)(-v_{yyy}^m / v_y^m)\left(l_m / m + \left[\alpha^m + \frac{v_{yl}^m v_y^m / v_l^m}{v_{yy}^m}\right]\partial l_m / \partial m\right) - v_{yyl}^m / v_y^m \partial l_m / \partial m$$

$$+ \left(-v_l^m l_m / \left(v_y^m m\right)\right)\left(v_{yy}^m\right)^2 / \left(v_y^m\right)^2 [1 + \epsilon_{ln} + (\alpha^m - 1)\epsilon_{ln}]$$

$$= \partial\left(-v_{yy}^m / v_y^m\right) / \partial y_m [1 + \alpha^m \epsilon_{ln}]\left(-v_l^m l_m / (v_y^m m)\right) - \partial l_m / \partial m / v_y^m \left(\frac{v_{yl}^m v_{yyy}^m}{-v_{yy}^m} + v_{yyl}^m\right)$$

In addition to the decreasing absolute risk aversion with respect to *y*, a sufficient condition for $\partial / \partial m\left(-v_{yy}^m / v_y^m\right) < 0$ is

$$\max\left\{\frac{v_{yl}^m v_{yyy}^m}{-v_{yy}^m}, \frac{-v_{yl}^m v_{yy}^m}{v_y^m}\right\} \geq -v_{yyl}^m.$$

In either sign of v_{yl}^m (positive or negative), this condition is satisfied when $v_{yyl}^m \geq 0$ or when the absolute value of v_{yyl}^m / v_{yl}^m is smaller than $v_{yyy}^m / (-v_{yy}^m)$ (when $v_{yl}^m > 0$) or $-v_{yy}^m / v_y^m$ (when $v_{yl}^m < 0$).

Appendix B

Proof of Proposition 2:

Derivation of (10.1):

Substituting (8) into (7)', we can write the distributional effect as, for $\psi_m = g_m/v_y^m$:

$$\Delta B(n) = \int_{m=\underline{n}}^{n} v_y^n \left((1-g_m)\frac{1}{v_y^m}\tau_x^m - \psi_m \frac{\Delta\eta}{\eta} - \psi_m s\epsilon_{gy}^m(-\Delta \bar{T}_m - \sum_i t_i x_m^i) - \psi_m \frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*}(1-T'_m)\Delta z_m \right) \frac{S_{mn}\, dF(m)}{1-F(n)}$$

Suppose first 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:

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

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

As the first possibility for (B.1) to hold, we have, by setting $\kappa(\eta) = 1$ when $\Delta\eta < 0$ and $\kappa(\eta) = 0$ when $\Delta\eta > 0$, $\frac{-\partial z_m/\partial\rho/z_m}{\epsilon_{zT'}^*} = -(v_{yy}^m - v_{yl}^m v_y^m/v_l^m)/v_y^m \equiv \beta^m s\epsilon_{gy}^m$ (see (A.3)) derives, with $\mu_m^T = -\Delta\bar{T}_m - \sum_i t_i x_m^i$, $\mu_m^y = \mu_m^T + \left(1 - T'(z_m)\right)\Delta z_m$,

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

$$= s\epsilon_{gy}^{n_1}\left(\beta^{n_1}\mu_{n_1}^y + (1-\beta^{n_1})\mu_{n_1}^T - \kappa(\eta)\frac{-\Delta\eta}{\eta}\frac{1}{s\epsilon_{gy}^{n_1}} \right)$$

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

If $\Delta\eta < 0$, for $s_m = \frac{-\Delta\eta}{\eta}\left(s\epsilon_{gy}^m\right)^{-1}$ that is increasing in m under A1, (B.2) must imply (10.1). We show below that $\max\{\Delta y_{n_1}, -\Delta \bar{T}_{n_1}, (1-T'(z_{n_1}))\Delta z_{n_1}\} - \sum_i t_i x_{n_1}^i > s_{n_1}$ when $\beta^{n_1} \leq 1$ (implied by A3, $\frac{v_{yl}^{n_1}}{v_l^{n_1}} \leq -\Psi''(v_{n_1})v_y^{n_1}/\Psi'(v_{n_1})$).

If $\Delta\eta>0$, (B.2) implies (10.1) ($s_{n_1} = 0 = s_m$). A1 warrants $k_{n_1} > 1$.

Proof of $\max\{\Delta y_{n_1}, -\Delta \bar{T}_{n_1}\} - \sum_i t_i x_{n_1}^i > 0$ *when* $\Delta\eta < 0$ and $\beta^{n_1} \leq 1$*:* We also need $-\Delta g_{n_1} - \tau_x^{n_1} > 0$, so,

$$(1-g_{n_1})\frac{1}{g_{n_1}}\tau_x^{n_1} + \frac{-\Delta\eta}{\eta} - s\epsilon_{gy}^{n_1}\left(\beta^{n_1}\mu_{n_1}^y + (1-\beta^{n_1})\mu_{n_1}^T\right) < 0.$$

If $\Delta\eta < 0$, since $\left(1-g_{n_1}\right)\frac{1}{g_{n_1}}\tau_x^{n_1} > 0$, this inequality implies that $\beta^{n_1}\mu_{n_1}^y + (1-\beta^{n_1})\mu_{n_1}^T > \frac{-\Delta\eta}{\eta}\left(s\epsilon_{gy}^{n_1}\right)^{-1} \equiv s_{n_1} > 0$. Therefore, $\mu_{n_1}^T > s_{n_1}$ or $\mu_{n_1}^y > s_{n_1}$.

The Case of Subsidy $\tau_x^n < 0$*:* With the same procedure, if n_1 is the highest skill level that has $\Delta\mathcal{W}_{n_1} \leq 0$, with $-\mu_m^T = \Delta\bar{T}_m + \sum_i t_i x_m^i$, $-\mu_m^y = -\mu_m^T - \left(1-T'(z_m)\right)\Delta z_m$, $s_m = \iota(\eta)\frac{\Delta\eta}{\eta}\left(s\epsilon_{gy}^m\right)^{-1}$, the following must hold:

$$(B.3) \qquad -\beta^{n_1}\mu_{n_1}^y - (1-\beta^{n_1})\mu_{n_1}^T - s_{n_1} > k_{n_1}\int_{m=\underline{n}}^{n_1}\psi_m s\epsilon_{gy}^m S_{mn_1}\left[-\beta^m\mu_m^y - (1-\beta^m)\mu_m^T - s_m\right]\frac{dF(m)}{1-F(n_1)}\left(\int_{m=\underline{n}}^{n_1}\psi_m s\epsilon_{gy}^m S_{mn_1}\frac{dF(m)}{1-F(n_1)}\right)^{-1}$$

Either $-\mu_{n_1}^T - s_{n_1}$ or $-\mu_{n_1}^y - s_{n_1}$ must exceed the weighted average of the corresponding value multiplied by k_{n_1}. If $\Delta\eta > 0$ ($s_{n_1} > 0$) and $\beta^{n_1} \leq 1$, then $\max\{\Delta\bar{T}_{n_1} - \left(1-T'\left(z_{n_1}\right)\right)\Delta z_{n_1}, \Delta\bar{T}_{n_1}\} + \sum_i t_i x_{n_1}^i > s_{n_1} > 0$ ($\mu_{n_1}^T$ or $\mu_{n_1}^T + \left(1-T'\left(z_{n_1}\right)\right)\Delta z_{n_1}$ is negative).

Derivation of (10.2) and (10.2)': The second possibility for (B.1) to hold is the following.

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

For m such that $g_m>1$, $(1-g_{n_1})\frac{1}{v_y^{n_1}}\tau_x^{n_1}>\ (1-g_m)\frac{1}{v_y^m}\tau_x^m$, $(1-g_{n_1})\frac{1}{v_y^{n_1}}\tau_x^{n_1}-(1-g_m)\frac{\tau_x^m}{v_y^m}S_{mn_1}>0$.

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 (10.2) as a necessary condition:

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

Under A2:

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

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

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

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

$$-\iota(\eta)\frac{\Delta\eta}{\eta}[\Psi''(v_n)/\eta\frac{-v_l^nl_n}{n}-\Psi'(v_n)/\eta\frac{v_{yl}^n}{v_y^nn}l_n]$$

For $\alpha^m=1-\frac{v_{yl}^mv_y^m/v_l^m}{v_{yy}^m}$, $\frac{n}{l_n}\frac{\partial l_n}{\partial n}=\epsilon_{ln}$, $\epsilon_{zn}\equiv\frac{n}{z_n}\frac{\partial z_n}{\partial n}=1+\epsilon_{ln}$, since $(1-\alpha^n)(1-T'(z_n))n=-\frac{v_{yl}^n}{v_{yy}^n}$, so $-\frac{\tau_x^n}{v_y^n}\frac{v_{yl}^n}{v_y^nn}l_n(\epsilon_{ln}+1-g_n)=\frac{\tau_x^n}{v_y^n}\frac{v_{yy}^n}{v_y^nn}(1-T'(z_n))z_n(1-\alpha^n)(\epsilon_{zn}-g_n)$. 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:

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

For $(1-T'(z_n))z_n=y_n\frac{1-T'(z_n)}{1-T(z_n)/z_n}$, the above formula is (10.2)'. From (A.3), $\alpha^n=\frac{-\epsilon_l^l}{\epsilon_{zT'}^*}\left(\frac{1-T'(z_n)}{1-T(z_n)/z_n}y_n(-v_{yy}^n)/v_y^n\right)^{-1}$.

QED

References

Atkinson, A., Stiglitz, J. 1976. The design of tax structure: direct versus indirect taxation. Journal of Public Economics, 6, 55–75.

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.

Edwards, J., Keen, M., Tuomala, M., 1994. Income tax, commodity taxes and public good provision: a brief guide. FinanzArchiv 51, 472–487.

Jacobs, B., Boadway, R., 2014. Optimal linear commodity taxation under optimal nonlinear income taxation. Journal of Public Economics, 117, 201–210.

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

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.

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

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

Saez, E., (2001), Using elasticities to derive optimal income tax rates. Review of Economic Studies, 68, 205–229.

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