

Indirect Taxes for Redistribution:
Should Necessity Goods be Favored?

Robin Boadway
Pierre Pestieau

CESIFO WORKING PAPER NO. 3667

CATEGORY 1: PUBLIC FINANCE

DECEMBER 2011

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Indirect Taxes for Redistribution: Should Necessity Goods be Favored?

Abstract

Atkinson and Stiglitz show that with weakly separability, differential commodity taxes are unnecessary given an optimal nonlinear income tax. Deaton showed that with an optimal linear progressive income tax, commodity taxes are superfluous under weakly separable and linear Engel curves. Using the latter case as an example, we derive two main results. If the income tax is less progressive than optimal, necessities should bear a lower tax rate than luxuries. If low-income households are income-constrained so cannot afford luxuries, it may be optimal to tax necessities at higher rates than luxuries, depending whether labor varies along the intensive or extensive margin.

JEL-Code: H210, H230.

Keywords: optimal income tax, Atkinson-Stiglitz Theorem, indirect taxes.

Robin Boadway
Department of Economics
Queen's University
Kingston / Ontario / Canada
boadwayr@econ.queensu.ca

Pierre Pestieau
Department of Economics
University of Liège
Liège / Belgium
p.pestieau@ulg.ac.be

December, 2011

1 Introduction

A key issue in tax policy concerns the use of differential commodity tax rates as part of the tax system. Most nations (a notable exception being the USA) now use value-added taxes (VATs) alongside direct taxes to raise the bulk of their revenues. Virtually all of them give preferential treatment to selected commodities using exemptions, zero-rating or reduced rates. In some cases, this reflects difficulties in taxing consumer services properly, as in the cases of housing and financial services. There are other instances where some goods are subject to higher tax rates because of externality arguments or as user charges. These include tobacco and alcohol products and petroleum. The case of interest to us is where differential commodity taxes are deployed as redistributive devices. Commodities that are taken to be relatively important for poorer taxpayers, such as food, children's clothing and footwear, and home heating, are often taxed at reduced rates. While there is some *prime facie* appeal to taxing more lightly goods with low elasticities of demand to improve redistributive outcomes, forceful arguments have been made that redistribution is better carried out by relying solely on progressive direct tax-transfer schemes, especially given the avoidable complexity preferential rates cause for a VAT system.

Recently, the Mirrlees Review (Mirrlees *et al* 2010) has taken this position, proposing that differential VAT rates be abolished by moving to uniform rates and adjusting the income tax system to maintain approximate distribution-neutrality. This is bound to be a controversial proposal politically since the cost of fully taxing necessities may be more salient to taxpayers than the relief given through income tax changes. Our focus is on the normative arguments for such reforms. The purpose of this paper is to shed some light on when it is reasonable to tax necessities preferentially for redistribution purposes. In doing so, we set aside arguments of administrative ease and political feasibility, and focus purely on tax reform principles based on optimal tax theory.

The theoretical basis for uniform commodity taxation originates with the Atkinson-Stiglitz Theorem (Atkinson and Stiglitz 1976). This states that, in a standard nonlinear income tax world à la Mirrlees (1971) with multiple commodities that can be taxed indirectly, if the government imposes an optimal nonlinear income tax, the commodity tax structure should be uniform if preferences are weakly separable in goods and leisure. Deaton (1979) showed that if the government were restricted to a linear income tax and set it optimally, the commodity tax structure should be uniform if preferences are separable and Engel curves for all goods were linear. More recently, Konishi (1995), Laroque (2005) and Kaplow (2006, 2008) have derived a potentially strong extension of the Atkinson-Stiglitz Theorem. Suppose preferences are weakly separable in goods and leisure, and start with a tax system that includes both differential commodity taxation and a non-optimal income tax. A Pareto-improving tax reform can be implemented that moves to uniform goods' taxation and adjusts all persons' income tax liabilities such that government budget balance is

maintained and incentive constraints are satisfied. Note the relevant point for our purposes that weak separability (with or without linear Engel curves) does not rule out very different income elasticities for different goods.

If preferences are not weakly separable, the theory can no longer recommend uniform taxation. The analogue of the well-known results of Corlett and Hague (1953) apply: goods that are relatively more complementary with leisure should bear correspondingly higher tax rates (Christiansen 1984; Edwards, Keen and Tuomala 1994; Nava, Schroyen and Marchand 1996). Other variations of the classical optimal tax model can also generate plausible arguments for differential taxation. Boadway, Marchand and Pestieau (1994) argue that indirect taxation can be desirable to the extent that it leads to less evasion than income taxation. Cremer and Gahvari (1995) show that, if consumer durables must be purchased before wage rates are known, a case can be made for preferential tax treatment of durables to offset the excessive precautionary saving to self-insure against wage uncertainty. Differential goods' taxation can also be called for a) if unobserved endowments of particular goods differ among individuals (Cremer, Pestieau and Rochat 2001), b) if preferences differ (Saez 2002a; Marchand, Pestieau and Racionero 2003; Blömquist and Christiansen 2008) or c) if needs for consumption for particular goods differ (Boadway and Pestieau 2003). And, Boadway and Gahvari (2006) show that when the time taken to consume goods is a substitute in utility for labor, a higher tax rate should be imposed on goods whose consumption is more time-intensive. While each of these studies provide a rationale for differential commodity taxes, they generally do not single out goods with low elasticities of demand for special treatment.¹

The proposals of the Mirrlees Review were informed by the Atkinson-Stiglitz Theorem and its extensions. The Review relied on consumer demand estimates that showed that for the most part, necessity goods that were favored by the UK VAT were not complementary with leisure, with notable exceptions such as child care. Based on that and setting aside arguments such as compliance costs and differences in preferences and needs, its proposal for a roughly distribution-neutral reform that moves to uniformity is supported by the Konishi-Laroque-Kaplow analysis.

The Mirrlees Review proposals effectively take the existing level of redistribution as given and seek to make it more efficient. Our analysis is concerned with policies to improve redistribution starting from some status quo. We study two forms of arguments for differential taxation of necessities. One argument applies when the income tax is not set optimally. We show, using a simple model with weakly separable preferences that satisfy the Deaton conditions, that when the

¹An exception is Cremer, Pestieau and Rochet (2001). They show that when the endowments of some goods differ among individuals and the tax authority cannot observe endowments, the tax rates on goods for whom endowments are identical should increase with the income elasticity of demand if cross-substitution effects are all zero (compensated demands depend only on own prices). In the likely event that cross-substitution effects are not zero, little can be said. We assume in this paper that endowments of all commodities are zero.

income tax system is not optimal to begin with, and when person-specific income tax adjustments are not feasible, giving preferential tax treatment to goods with relatively low income elasticities of demand can be welfare-improving. Second, we show that when low-income individuals are constrained to consume only necessity goods, paradoxically it may be optimal to tax necessities at a higher rate than luxuries when the income tax is set optimally. Thus, the tax treatment of necessities remains an open question.

Our basic approach essentially applies the Corlett and Hague (1953) tax reform methodology to a heterogeneous-household setting. The government can impose a linear progressive income tax, as in Sheshinski (1972), but since there is more than one consumer good, it can also deploy differential commodity taxes as in Deaton (1979). The utility function involves two goods and leisure, and takes an extreme form for transparency of interpretation. It is quasilinear in one of the goods, which implies that good has a zero income elasticity of demand so is a necessity, while the other has an income elasticity of demand greater than unity so is a luxury. The utility function satisfies Deaton's conditions, so if all consumers consumed both goods and the linear income tax were set optimally, uniform commodity taxation would be optimal. Our first main result shows that, starting with an arbitrary linear progressive tax and uniform commodity taxes, a revenue-neutral reform of the commodity tax structure that reduced the tax rate on the good with the low elasticity of demand and increased the other would be welfare-improving if the income tax were less progressive than is optimal, and vice versa. Our second result applies the same methodology to the case where low-income households are unable to afford any of the luxury good, and finds that the necessity should bear a differentially high tax rate when the income tax is optimal.

We extend the main analysis in three directions. First, we show that if labor supply varies along the extensive margin so that households choose only whether to participate, as in Diamond (1980) and Saez (2002b), the second result must be modified. There is now an efficiency argument against taxing the necessity since it discourages participation. Second, we extend the results to a nonlinear income tax setting using again the extensive-margin model. Unlike with a linear progressive tax, the meaning of the income tax being less progressive than optimal is ambiguous. We adopt a simple but appealing example where that ambiguity disappears. Finally, we allow government revenue requirements to be endogenous by assuming the government supplies a public good, a case originally considered by Atkinson and Stern (1974) in a simpler context. When the government sets the linear progressive tax optimally, not only should commodity taxes be uniform, but the Samuelson rule for the public good should be satisfied. That is, the marginal cost of public funds should be unity. If the income tax is less progressive than optimal, the rule for public goods provision depends on how the public good is financed. If financing comes from adjusting the lump-sum component of the income tax, the Samuelson condition continues to apply. However, if financing comes from changes in the income tax rate, as in Atkinson and Stern, public good supply

should be greater than that indicated by the Samuelson rule: the marginal cost of public funds should be less than unity. The opposite applies if the income tax is more progressive than optimal.

2 The Model

The economy is populated by a distribution of individuals who differ by their exogenous wage rate w_i . There are n_i persons of type w_i , with $i = 1, \dots, r$, where $\sum_{i=1}^r n_i = 1$ for simplicity (since the absolute population is irrelevant for the results). Individuals consume two goods, x_1 and x_2 , and supply labor, ℓ . All have the same preferences, which we take to be quasilinear in x_1 , or $x_1 + b(x_2) - h(\ell)$, where $b(x_2)$ is increasing and strictly concave and $h(\ell)$ is increasing and strictly convex. Let q_1 and q_2 be the consumer prices of the two goods. Since preferences are quasilinear in x_1 , all changes in disposable income go to x_1 , so $dx_1/dc = 1/q_1$ and $dx_2/dc = 0$, where c is disposable income, including labor income. The slopes of Engel curves are thus linear and the same for all households. The Deaton (1979) conditions are then satisfied, as is weak separability required for the Atkinson-Stiglitz Theorem (Atkinson and Stiglitz 1976). The proportion of income spent on good x_1 is $q_1 x_1/c$, which is increasing in c .² Thus, x_1 is a luxury, while x_2 is an extreme necessity.

Production is linear, and we normalize producer prices of both goods to be unity. All income comes from labor, with pre-tax labor income for a person of type w_i denoted by $y_i \equiv w_i \ell_i$. The government can impose a linear progressive income tax and commodity taxes on the two goods. Let t be the tax rate on income, θ_1 and θ_2 the tax rates on x_1 and x_2 , and a the lump-sum component of the linear tax system.³ The budget constraint of a type- w_i person is then:

$$(1 + \theta_1)x_{1i} + (1 + \theta_2)x_{2i} = (1 - t)w_i \ell_i + a \quad (1)$$

Using (1) to determine x_{1i} , consumer i solves the following problem:

$$\max_{\{x_{2i}, \ell_i\}} \frac{1 - t}{1 + \theta_1} w_i \ell_i + \frac{a}{1 + \theta_1} - \frac{1 + \theta_2}{1 + \theta_1} x_{2i} + b(x_{2i}) - h(\ell_i)$$

This yields as a solution the increasing functions $x_{2i}((1 + \theta_2)/(1 + \theta_1))$ and $\ell_i((1 - t)w_i/(1 + \theta_1))$ and the value function $v_i(\theta_1, \theta_2, t, a)$. Initially, we assume that all individuals consume strictly positive amounts of both goods. This implies that $x_{2i} = x_2$ for all i . Later, we allow for the possibility that for low-income persons, income is insufficient to purchase the desired amount of x_2 , or equivalently

²That is,

$$\frac{\partial(q_1 x_1/c)}{\partial c} = \frac{q_1}{c} \frac{\partial x_1}{\partial c} - \frac{q_1 x_1}{c^2} = \frac{1}{c} \left(1 - \frac{q_1 x_1}{c}\right) > 0$$

³The absolute level of the commodity taxes θ_1 and θ_2 do not matter, only their relative values. It is convenient for us to work with absolute tax rates θ_1 and θ_2 rather than relative ones θ_1/θ_2 .

a non-negativity constraint on the purchases of x_1 is binding. By the envelope theorem, we have:

$$v_{i\theta_1} = -\frac{x_{1i}}{1+\theta_1}; \quad v_{i\theta_2} = -\frac{x_{2i}}{1+\theta_1}; \quad v_{it} = -\frac{w_i\ell_i}{1+\theta_1} = -\frac{y_i}{1+\theta_1}; \quad v_{ia} = \frac{1}{1+\theta_1} \quad (2)$$

where subscripts on $v(\cdot)$ denote partial derivatives in an obvious way.

The government budget constraint is $\sum_i n_i(\theta_1 x_{1i} + \theta_2 x_{2i} + t w_i \ell_i) - a = 0$, where we assume with no loss of generality that the government needs no net revenue. Using the solution for x_{1i} from (1), this can be written as:

$$B(\theta_1, \theta_2, t, a) \equiv \sum_i n_i \left(\frac{\theta_2 - \theta_1}{1 + \theta_1} \cdot x_{2i} \left(\frac{1 + \theta_2}{1 + \theta_1} \right) + \frac{\theta_1 + t}{1 + \theta_1} \cdot w_i \ell_i \left(\frac{(1 - t) w_i}{1 + \theta_1} \right) \right) - \frac{a}{1 + \theta_1} = 0 \quad (3)$$

Differentiating $B(\cdot)$, we obtain:

$$B_{\theta_2} = \sum_i n_i \left(\frac{x_{2i}}{1 + \theta_1} + \frac{\theta_2 - \theta_1}{(1 + \theta_1)^2} x'_{2i} \right); \quad B_t = \sum_i n_i \left(\frac{w_i \ell_i}{1 + \theta_1} - \frac{\theta_1 + t}{(1 + \theta_1)^2} w_i^2 \ell'_i \right); \quad B_a = -\frac{1}{1 + \theta_1} \quad (4)$$

and

$$\begin{aligned} B_{\theta_1} &= \sum_i n_i \left(\frac{1 - t}{(1 + \theta_1)^2} w_i \ell_i - \frac{\theta_1 + t}{(1 + \theta_1)^2} \frac{1 - t}{1 + \theta_1} w_i^2 \ell'_i - \frac{1 + \theta_2}{(1 + \theta_1)^2} x_{2i} - \frac{\theta_2 - \theta_1}{(1 + \theta_1)^2} \frac{1 + \theta_2}{1 + \theta_1} x'_{2i} \right) + \frac{a}{(1 + \theta_1)^2} \\ &= \frac{1 - t}{1 + \theta_1} B_t - \frac{1 + \theta_2}{1 + \theta_1} B_{\theta_2} - \frac{a}{1 + \theta_1} B_a \end{aligned} \quad (5)$$

We assume in what follows that $B_t > 0$ and $B_{\theta_i} > 0$, so the tax revenues are increasing in all tax rates and we are on the rising part of the Laffer curve.

The objective function of the government is taken to be an additive social welfare function:

$$W(\theta_1, \theta_2, t, a) = \sum_i n_i u(v_i(\theta_1, \theta_2, t, a)) \quad (6)$$

We can interpret the function $u(\cdot)$ either as the individual's concave utility function or the planner's social utility function, or a combination of the two (Kaplow 2010). These are equivalent from an analytical perspective.

We consider two cases. In the first one, we set $\theta_1 = \theta_2 = 0$ and examine the social welfare properties of a given linear income tax system (t, a) when the government must balance its budget. This case includes that where some of the revenue to finance the transfer a is raised by a uniform commodity tax on x_1 and x_2 , since a uniform tax is equivalent to a proportional income tax and can be subsumed in the income tax rate t and lump-sum component a . We then perturb θ_1 and θ_2 holding t and a constant to investigate the case for differential commodity taxes, given the progressivity of the income tax. Formally, we increase θ_2 and allow θ_1 to change to balance the budget.

3 Uniform Commodity Taxation

Consider the effect of a change in the marginal income tax rate t starting at $\theta_1 = \theta_2 = 0$, with a adjusting to balance the budget, so that $da/dt = -B_t/B_a > 0$. Differentiating (6), we obtain:

$$\left. \frac{dW}{dt} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = \sum n_i u'_i \left(v_{it} + v_{ia} \left. \frac{da}{dt} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} \right) = \sum n_i u'_i \left(v_{it} - v_{ia} \frac{B_t}{B_a} \right)$$

where u'_i is the marginal social utility of a person with wage rate w_i . Using the envelope results (2), the properties of the government budget (4) and (5), and $-w^2 \ell'_i = \partial y_i / \partial t \equiv y_{it}$, this becomes:

$$\left. \frac{dW}{dt} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = \sum n_i u'_i \sum n_i y_i - \sum n_i y_i u'_i + \sum n_i u'_i \sum n_i t y_{it} = \underbrace{-\text{Cov}[u', y]}_{\text{equity (+)}} + \underbrace{t E[u'] E[y_t]}_{\text{efficiency (-)}} \quad (7)$$

These two terms have a familiar interpretation (Sheshinski 1972). The first term represents the equity benefits of increasing the tax rate, with the lump-sum transfer a increasing to balance the budget. The second term reflects the marginal efficiency cost of increasing the tax rate.

An increase in t increases the progressivity of the tax. The optimal linear progressive tax will apply when (7) equals zero so the equity and efficiency effects of a tax increase just offset one another. Denote the optimal linear tax rate by t^* . Then we have:

Lemma 1 *If the commodity tax structure is uniform, so $\theta_1 = \theta_2 = 0$, then for a given tax rate t ,*

$$t^* \begin{cases} \geq \\ \leq \end{cases} t \quad \text{as} \quad -\text{Cov}[u', y] \begin{cases} \geq \\ \leq \end{cases} -t E[u'] E[y_t] \quad (8)$$

As mentioned, the marginal tax rate t could include varying combinations of a tax on labor income and a uniform tax on commodities.

4 Differential Commodity Taxation

To investigate whether differential commodity taxation is desired, we undertake a tax reform exercise in the spirit of Corlett and Hague (1953). Holding the marginal income tax rate t and the lump-sum component a constant to maintain progressivity, we analyze the social welfare effect of a small increase in θ_2 , the tax on good x_2 , allowing θ_1 to adjust to balance the budget.

Starting from some arbitrary level of t and a , differentiation of $W(\theta_1, \theta_2, a, t)$ yields:

$$\left. \frac{dW}{d\theta_2} \right|_{t,a} = \sum n_i u'_i \left(v_{i\theta_2} + v_{i\theta_1} \frac{d\theta_1}{d\theta_2} \right) = \sum n_i u'_i \left(v_{i\theta_2} - v_{i\theta_1} \frac{B_{\theta_2}}{B_{\theta_1}} \right) = \frac{\sum n_i u'_i (v_{i\theta_2} B_{\theta_1} - v_{i\theta_1} B_{\theta_2})}{B_{\theta_1}}$$

Using (2), (4), (5) and $x_{2i} = x_2$, and evaluating this at $\theta_1 = \theta_2 = 0$, we obtain:

$$\left. \frac{B_{\theta_1}}{x_2} \frac{dW}{d\theta_2} \right|_{\substack{t,a \\ \theta_1=\theta_2=0}} = - \sum n_i u'_i \left(\sum n_i (1-t) y_i + t(1-t) y_{it} - x_2 + a \right) + \sum n_i u'_i x_{1i}$$

Using the individual budget constraint (1), this may be written:

$$\frac{B_{\theta_1}}{(1-t)x_2} \frac{dW}{d\theta_2} \Big|_{\substack{t,a \\ \theta_1=\theta_2=0}} = - \sum n_i u'_i \sum n_i (y_i + ty_{it}) + \sum n_i y_i u'_i = \text{Cov}[u', y] - tE[u']E[y_t] \quad (9)$$

Recall that we assume $B_{\theta_1} > 0$, so that an increase in the marginal tax rate increases revenue. Then, comparing (9) with (8) leads to the following proposition.

Proposition 1 *Starting from uniform commodity taxes ($\theta_1 = \theta_2 = 0$) and an arbitrary t and a , and assuming all individuals consume both goods:*

$$\frac{dW}{d\theta_2} \Big|_{\substack{t,a \\ \theta_1=\theta_2=0}} \geq 0 \quad \text{as} \quad t \geq t^* \quad (10)$$

The implication of Proposition 1 is that if the income tax rate is below the optimal one, so the tax system is less progressive than at the optimum, reducing the commodity tax rate on the good with the low income elasticity of demand (good x_2) below that of the good with the high income elasticity of demand (good x_1) starting from uniformity will be welfare-improving. Therein lies the argument for favoring necessities on redistributive grounds.

5 Demand for the Necessity Good Income-Constrained

So far, we have assumed that all individuals are able to choose an interior solution, which requires them to have enough income to purchase the common optimal level of x_2 . In this section, we explore the consequences of this not being the case for lower-income individuals, and show paradoxically that if the income tax is set optimally, the necessity good should be taxed at a *higher* rate than the luxury.

Individual preferences are again assumed to be quasilinear, $x_{1i} + b(x_{2i}) - h(\ell_i)$, and social utility from the planner's perspective is taken to be strictly concave. The individual maximizes utility subject to the usual budget constraint, (1), and a non-negativity constraint applied to good 1, $x_{1i} \geq 0$. (A non-negativity constraint on x_{2i} will never be binding as long as individuals have positive income.) This constraint will be assumed to be binding for low-wage persons, for whom income is insufficient to purchase the otherwise desired level of x_{2i} . Using (1) to eliminate x_{1i} , the Lagrange expression for a type- i individual is:

$$\mathcal{L}_i = \frac{1-t}{1+\theta_1} w_i \ell_i + \frac{a}{1+\theta_1} - \frac{1+\theta_2}{1+\theta_1} x_{2i} + b(x_{2i}) - h(\ell_i) + \mu_i \left(\frac{1-t}{1+\theta_1} w_i \ell_i + \frac{a}{1+\theta_1} - \frac{1+\theta_2}{1+\theta_1} x_{2i} \right)$$

where μ_i is the Lagrange multiplier on the non-negativity constraint. From the first-order conditions, we obtain the functions $x_{2i}((1+\mu_i)(1+\theta_2)/(1+\theta_1))$ and $\ell_i((1+\mu_i)(1-t)w_i/(1+\theta_1))$, where

$\mu_i x_{1i} = 0$. The indirect utility function is again $v_i(\theta_1, \theta_2, t, a)$, with the envelope properties

$$v_{i\theta_1} = -\frac{1 + \mu_i}{1 + \theta_1} x_{1i}, \quad v_{i\theta_2} = -\frac{1 + \mu_i}{1 + \theta_1} x_{2i}, \quad v_{it} = -\frac{1 + \mu_i}{1 + \theta_1} w_i \ell_i, \quad v_{ia} = \frac{1 + \mu_i}{1 + \theta_1} \quad (11)$$

There will be some cutoff wage, \tilde{w} , such that for all $w_i < \tilde{w}$, the constraint $x_{1i} \geq 0$ is binding, so $\mu_i > 0$. Moreover, the value of μ_i will be decreasing with the wage rate in this range. For those with $w_i > \tilde{w}$ the non-negativity constraint is not binding, so $\mu_i = 0$ as in the case considered above. For individuals for whom the non-negativity constraint is binding, their demand for x_2 will be lower and their supply of ℓ will be greater than for higher-wage persons who are not income-constrained.

Suppose the government chooses the optimal linear progressive income tax system, given the commodity tax rates θ_1 and θ_2 . The social welfare function is now $W(\cdot) = \sum_i n_i u(v_i(\theta_1, \theta_2, t, a))$, where $u(\cdot)$ is a strictly concave social utility function. Using the government budget constraint (3), the Lagrange expression for the optimal income tax problem can be written, using $y_i = w_i \ell_i$:

$$\begin{aligned} \mathcal{L} = & \sum_i n_i u(v_i(\theta_1, \theta_2, t, a)) \\ & + \lambda \sum_i n_i \left(\frac{\theta_2 - \theta_1}{1 + \theta_1} \cdot x_{2i} \left(\frac{(1 + \mu_i)(1 + \theta_2)}{1 + \theta_1} \right) + \frac{\theta_1 + t}{1 + \theta_1} \cdot y_i \left(\frac{(1 + \mu_i)(1 - t)w_i}{1 + \theta_1} \right) - \frac{a}{1 + \theta_1} \right) \end{aligned}$$

The first-order conditions on t and a evaluated at $\theta_1 = \theta_2 = 0$ are:

$$\sum_i n_i u'_i v_{it} + \lambda \sum_i n_i (y_i + t y_{it}) = 0; \quad \sum_i n_i u'_i v_{ia} - \lambda = 0 \quad (12)$$

Using the envelope theorem properties (11), this can be written:

$$-\sum_i n_i y_i u'_i (1 + \mu_i) + \sum_i n_i u'_i (1 + \mu_i) \sum_i n_i y_i + \sum_i n_i u'_i (1 + \mu_i) \sum_i n_i t y_{it} = 0$$

or,

$$-\text{Cov}[u'(1 + \mu), y] + t E[u'(1 + \mu)] E[y_i] = 0$$

This is the standard optimal tax expression analogous to (7) above involving an equity and an efficiency effect. Both are augmented by the multiplier μ_i arising from the non-negativity constraint that we assume is binding at the bottom end.

Let the value function for the optimal income tax problem be $\mathcal{W}(\theta_1, \theta_2)$. Applying the envelope theorem, the change in the value of social welfare from a change in the commodity tax rate θ_2 , evaluated at $\theta_1 = \theta_2 = 0$ is:

$$\left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\theta_1=0, \theta_2=0} = \left. \frac{\partial \mathcal{L}}{\partial \theta_2} \right|_{\theta_1=0, \theta_2=0} = \sum_i n_i u'_i v_{i\theta_2} + \lambda \sum_i n_i x_{2i}$$

Using the envelope results in (11), and the value of λ in (12), this becomes

$$\left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\theta_1=0, \theta_2=0} = -\sum_i n_i u'_i (1 + \mu_i) x_{2i} + \sum_i n_i u'_i (1 + \mu_i) \sum_i n_i x_{2i} = -\text{Cov}[u'(1 + \mu), x_2] \quad (13)$$

In the case where the non-negativity constraint $x_{1i} \geq 0$ is not binding, x_2 is constant, so the welfare change in (13) is zero as expected. However, when the non-negativity constraint is binding for some low-income persons, x_{2i} is increasing for them. Since $u'_i(1 + \mu_i)$ decreases with the wage rate while x_{2i} weakly increases, this expression is positive, implying that the tax on the necessity good x_2 should be *positive*. We summarize this in the following proposition.

Proposition 2 *If the optimal linear income tax is in place, the necessity good x_2 should be taxed at a higher rate than the luxury good x_1 if low-wage persons are income-constrained so that they purchase none of the luxury and a suboptimal amount of the necessity.*

Contrary to the basic case leading to Proposition 1, when low-income individuals are income-constrained from purchasing the luxury good x_1 , it is optimal to tax the necessity punitively when the income tax is optimal, apparently because high-wage persons consume more of the necessity. This somewhat surprising result seems to be of some practical relevance. In the real world, low-income persons devote much of their income to necessities, and consume virtually no luxury goods. This seems to cast some doubt on the relevance of the Atkinson-Stiglitz Theorem as a basis for advocating uniform commodity taxes, though as the following sections show the results of Proposition 2 are modified when we allow workers to respond to the tax-transfer system by changing labor supply along the extensive margin.

6 The Extensive-Margin Case

The above analysis extends to the case where each worker's labor market choice includes a participation decision. This case is useful to consider since it allows us to introduce nonlinear income taxes in a straightforward way in the next section. Consider the extreme case studied initially by Diamond (1980) and exploited further by Saez (2002b) where the only decision workers make is whether to participate in the labor market. Participation by a worker of skill-type i involves choosing to work in a type- i job for a fixed number of hours and earning a given income y_i , $i = 1, \dots, r$, where $y_i > y_{i-1}$. Workers of all skills who do not participate can engage in non-market or leisure pursuits, referred to as activity 0.⁴ Adopting the same preferences for goods as above, let the utility of a type- i participant take the quasilinear form $x_{1i} + b(x_{2i})$. Non-participants vary both by their skill-type and by their preference for leisure, denoted m . Let the utility of a non-participating person of skill-type i and preference-type m be given by $x_{10} + b(x_{20}) + m_i$, where m_i is distributed according to the distribution function $\Gamma_i(m_i)$, which could differ by skill level. The marginal participant of

⁴It would be fairly routine to extend the analysis to the case considered by Saez (2002b) where workers can also choose the job of a less-skilled worker.

skill-type i satisfies:

$$x_{1i} + b(x_{2i}) = x_{10} + b(x_{20}) + \hat{m}_i \quad (14)$$

All persons with $m_i < \hat{m}_i$, of whom there will be $n_i \Gamma_i(\hat{m}_i) = n_i \Gamma_i(x_{1i} + b(x_{2i}) - x_{10} - b(x_{20}))$, will participate in the labor market.

Assuming that a linear progressive income tax is in place, the budget constraint for a type- i participant is $(1 + \theta_1)x_{1i} + (1 + \theta_2)x_{2i} = (1 - t)y_i + a$. Using this, the participant chooses x_{2i} to maximize $((1 - t)y_i + a - (1 + \theta_2)x_{2i}) / (1 + \theta_1) + b(x_{2i})$. The solution gives the demand function $x_{2i}((1 + \theta_2) / (1 + \theta_1))$, which leads to the indirect utility function $v_i(\theta_1, \theta_2, t, a)$. The envelope theorem gives the following participation equilibrium condition:

$$v_{i\theta_1} = -\frac{x_{1i}}{1 + \theta_1}, \quad v_{i\theta_2} = -\frac{x_{2i}}{1 + \theta_1}, \quad v_{it} = -\frac{y_i}{1 + \theta_1}, \quad v_{ia} = \frac{1}{1 + \theta_1} \quad (15)$$

For non-participants, the same analysis applies except the transfer a is the only source of income, so $y_i = 0$. This implies that utility for a type- (i, m) non-participant can be written $(a - (1 + \theta_2)x_{20}) / (1 + \theta_1) + b(x_{20}) + m_i$, so the demand for x_2 is given by the function $x_{20}((1 + \theta_2) / (1 + \theta_1))$. We write indirect utility as $v_0(\theta_1, \theta_2, a) + m_i$, where $v_{0\theta_1} = -x_{10} / (1 + \theta_1)$, $v_{0\theta_2} = -x_{20} / (1 + \theta_1)$ and $v_{0a} = 1 / (1 + \theta_1)$. Assuming a is sufficiently large, we obtain $x_{2i} = x_{20} \equiv x_2$ for all i , whether they participate or not.

Given that x_2 is the same for all persons, and using the budget constraints for participants and non-participants, the participation equilibrium condition (14) reduces to $(1 - t)y_i / (1 + \theta_1) = \hat{m}_i$. The number of skill-type- i participants then becomes $n_i \Gamma_i((1 - t)y_i / (1 + \theta_1)) \equiv h_i((1 - t)y_i / (1 + \theta_1))$, where $h'_i(\cdot) > 0$. The responsiveness of participation to after-tax income depends upon the distribution of preferences for leisure in each group, something that we can be agnostic about.⁵ The total number of non-participants can be written $h_0 = 1 - \sum_{i \geq 1} h_i((1 - t)y_i / (1 + \theta_1))$.

The government budget constraint analogous to (3) is:

$$B(\theta_1, \theta_2, t, a) = \sum_{i \geq 1} h_i \left(\frac{(1 - t)y_i}{1 + \theta_1} \right) \frac{\theta_1 + t_1}{1 + \theta_1} y_i + \frac{\theta_2 - \theta_1}{1 + \theta_1} x_2 \left(\frac{1 + \theta_2}{1 + \theta_1} \right) - \frac{a}{1 + \theta_1} = 0 \quad (16)$$

Differentiation at $\theta_1 = \theta_2 = 0$ yields:

$$B_{\theta_1} = \sum_{i \geq 0} h_i x_{1i} - \sum_{i \geq 1} t(1 - t)y_i^2 h'_i; \quad B_{\theta_2} = x_2; \quad B_t = \sum_{i \geq 1} h_i y_i - \sum_{i \geq 1} h'_i t y_i^2; \quad B_a = -1 \quad (17)$$

We continue to assume that $B_t > 0$ and $B_{\theta_1}, B_{\theta_2} > 0$, so the government is operating on the increasing side of the Laffer curve.

⁵The elasticity of the participation function plays a critical role in defining the optimal tax system when the latter can be non-linear as we see in the next section. See Diamond (1980) and Saez (2002b).

Social utility for type- i participants will be written $u(v_i(\theta_1, \theta_2, t, a))$, where $u'(\cdot) > 0 > u''(\cdot)$ reflects the planner's aversion to inequality. For type- (i, m) non-participants, social utility is $u(v_0(\theta_1, \theta_2, a) + m_i)$.⁶ Social welfare can then be written as follows:

$$W(\theta_1, \theta_2, t, a) = \sum_{i \geq 1} h_i \left(\frac{(1-t)y_i}{1+\theta_1} \right) u(v_i(\theta_1, \theta_2, t, a)) + \sum_{i \geq 0} \int_{m_i > \hat{m}_i} u(v_0(\theta_1, \theta_2, a) + m_i) d\Gamma_i(m_i) \quad (18)$$

The first term is the sum of social utilities for participants, while the latter includes all non-participants. Non-participants differ by both skill-type and preference-type, and can include some skill-type-0's, none of whom work. As above, we consider two cases, the first with uniform commodity taxation and the second allowing the commodity tax on x_2 to vary.

6.1 Changing progressivity with uniform commodity taxation

Consider the effect of increasing the marginal income tax rate t with the lump-sum transfer a adjusting to balance the budget. The change in social welfare in (18) can be written:

$$\begin{aligned} \frac{dW}{dt} \Big|_{\substack{\theta_1=0 \\ \theta_2=0}} &= \sum_{i \geq 1} h_i \left(\frac{(1-t)y_i}{1+\theta_1} \right) u'(v_i(\theta_1, \theta_2, t, a)) \left(v_{it} + v_{ia} \frac{da}{dt} \right) \\ &\quad + \sum_{i \geq 0} \int_{m_i > \hat{m}_i} u'(v_0(\theta_1, \theta_2, a) + m_i) d\Gamma_i(m_i) v_{0a} \frac{da}{dt} \end{aligned}$$

Note that we can neglect changes in social welfare arising from marginal changes in participation, h'_i . Since marginal persons are indifferent between participating and not, the change in their utility, and thus in social welfare, is only second order in magnitude. Since $da/dt = -B_t/B_a$ from (16), $B_a = -1$ by (17), and $v_{ia} = v_{0a} = 1$ by the envelope theorem, this welfare change expression can be expanded as follows:

$$\begin{aligned} \frac{dW}{dt} &= - \sum_{i \geq 1} h_i u'_i y_i + \left(\sum_{i \geq 1} h_i u'_i + \sum_{i \geq 0} \int_{m_i > \hat{m}_i} u'_{i0} d\Gamma_i \right) \left(h_i y_i - \sum_{i \geq 1} h'_i t y_i^2 \right) \\ &= -E[u'y] + E[u']E[y] - tE[u']e[y^2h'] \end{aligned} \quad (19)$$

The first two terms give $-\text{Cov}[u', y]$ and is positive, while the last term represents the efficiency loss of an increase in the income tax rate. The optimal linear income tax rate is obtained when (19) is set to zero. The analogue of Lemma 1 applies.

⁶We assume that the planner attaches the same social weight to the value of leisure as do individuals. This is not essential for our main argument.

6.2 Differential commodity taxation

Following the same analysis as above, consider now an increase in θ_2 , adjusting θ_1 to balance the budget while holding t and a constant. Differentiating social welfare yields:

$$\begin{aligned} \left. \frac{dW}{d\theta_2} \right|_{t,a} &= \sum_{i \geq 1} h_i u' (v_i(\theta_1, \theta_2, t, a)) \left(v_{i\theta_2} + v_{i\theta_1} \frac{d\theta_1}{d\theta_2} \right) \\ &+ \left(\sum_{i \geq 0} \int_{m_i > \hat{m}_i} u' (v_0(\theta_1, \theta_2, a) + m_i) d\Gamma_i(m_i) \right) \left(v_{0\theta_2} + v_{0\theta_1} \frac{d\theta_1}{d\theta_2} \right) \end{aligned}$$

Using $d\theta_1/d\theta_2 = -B_{\theta_2}/B_{\theta_1}$, the properties of the budget (17), and the envelope theorem results, this reduces to the following when $\theta_1 = \theta_2 = 0$:

$$\left. \frac{B_{\theta_1}}{(1-t)x_2} \frac{dW}{d\theta_2} \right|_{\substack{t,a \\ \theta_1=\theta_2=0}} = E[u'y] - E[u']E[y] + tE[u']e[y^2h']$$

Comparing this with (19), we see that Proposition 1 applies here as well: if the linear income tax is less progressive than optimal, the necessity good x_2 should be subsidized (or taxed preferentially relative to x_1). While this analysis focuses solely on the participation decision, it is apparent that if both extensive and intensive responses are included, the results of Proposition 1 would continue to apply.

6.3 Low-income demand for the necessity income-constrained

Suppose now that the purchase of x_2 by some low-skilled workers as well as non-participants is income-constrained. Following the same analysis as above the individual problem must satisfy the constraint $x_{1i} \geq 0$ with the associated shadow price μ_i , $i = 0, 1, \dots, r$. The demand function for x_2 now becomes $x_{21}((1 + \mu_i)(1 + \theta - 2)/(1 + \theta_1))$, and the envelope conditions become:

$$v_{i\theta_1} = -\frac{1 + \mu_i}{1 + \theta_1} x_{1i}, \quad v_{i\theta_2} = -\frac{1 + \mu_i}{1 + \theta_1} x_{2i}, \quad v_{it} = -\frac{1 + \mu_i}{1 + \theta_1} y_i, \quad v_{ia} = \frac{1 + \mu_i}{1 + \theta_1}, \quad i = 0, \dots, r$$

where $y_0 = 0$ for non-participants. For low-income households μ_i is positive and decreasing in income, so their demands for x_2 are less than the unconstrained amounts and increasing in skill. The participation equilibrium condition (14) can now be written, using the household budget constraint, as:

$$\hat{m}_i = \frac{(1-t)y_i - (1+\theta_2)(x_{2i} - x_{20})}{1 + \theta_1} + b(x_{2i}) - b(x_{20})$$

so we can write the number of type- i participants and the number of non-participants as:

$$h_i \left(\frac{(1-t)y_i - (1+\theta_2)(x_{2i} - x_{20})}{1 + \theta_1} + b(x_{2i}) - b(x_{20}) \right), \quad h_0 = 1 - \sum_{i \geq 1} h_i(\cdot) \quad (20)$$

where x_{2i} and x_{20} are given by the above demand functions.

Assume the government chooses the optimal linear income tax system. Using the expression for social welfare (18) and the government budget (16) revised to incorporate the income-constrained variables, the Lagrangian expression can be written:

$$\begin{aligned} \mathcal{L} = & \sum_{i \geq 1} h_i(\cdot) u(v_i(\theta_1, \theta_2, t, a)) + \sum_{i \geq 0} \int_{m_i > \hat{m}_i} u(v_0(\theta_1, \theta_2, a) + m_i) d\Gamma_i(m_i) \\ & + \lambda \sum_{i \geq 0} h_i(\cdot) \left(\frac{\theta_1 + t_1}{1 + \theta_1} y_i + \frac{\theta_2 - \theta_1}{1 + \theta_1} x_2 \left(\frac{(1 + \mu_i)(1 + \theta_2)}{1 + \theta_1} \right) - \frac{a}{1 + \theta_1} \right) \end{aligned} \quad (21)$$

where $h_i(\cdot)$ is given by (20). Taking the first-order conditions with respect to t and a and evaluating them at $\theta_1 = \theta_2 = 0$, we obtain (19) set equal to zero. This is the standard optimal linear income tax formula in this context.

As before, let the value function for this optimal tax problem be $\mathcal{W}(\theta_1, \theta_2)$. To study the effect of introducing differential commodity taxes, we can evaluate a change in θ_2 using the envelope theorem. Differentiating the Lagrangian expression, we obtain the following starting at $\theta_1 = \theta_2 = 0$:

$$\left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = \sum_{i \geq 1} h_i u'_i v_{i\theta_2} + \sum_{i \geq 0} \int_{m_i > \hat{m}_i} u'_{i0} d\Gamma_i v_{0\theta_2} + \lambda \left(\sum_{i \geq 0} h_i x_{2i} + \sum_{i \geq 1} \frac{dh_i}{d\theta_2} t y_i \right)$$

Using $v_{i\theta_2} = -(1 + \mu_i)x_{2i}$ when $\theta_1 = \theta_2 = 0$ and the first-order condition on a obtained from (21), this reduces to:

$$\begin{aligned} \left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} &= -E[u'(1 + \mu)x_2] + E[u'(1 + \mu)]E[x_2] + E[u'(1 + \mu)]t \sum_{i \geq 1} \frac{dh_i}{d\theta_2} y_i \\ &= -\text{Cov}[u'(1 + \mu), x_2] + E[u'(1 + \mu)]t \sum_{i \geq 1} \frac{dh_i}{d\theta_2} y_i \end{aligned} \quad (22)$$

This is analogous to (13) for the intensive-margin case with the addition of the second term. The first term is positive when the non-negativity constraint applies on x_{1i} for some households, and tends to favor a tax on the necessity. This can be thought of as an equity effect.

The second term arise because the tax on the necessity changes the number of persons who participate in the labor market, which is like an efficiency effect. Differentiating (20) with respect to θ_2 , taking into account the effect on x_{2i} and x_{20} , yields:

$$\left. \frac{dh_i}{d\theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = h'_i \cdot \left(x_{20} - x_{2i} + (b'(x_{2i}) - 1)x'_{2i} - (b'(x_{20}) - 1)x'_{20} \right) \quad (23)$$

The sign of (23) is generally ambiguous since it depends on the relative magnitude of x'_{2i} and x'_{20} . However, a sufficient condition for it to be negative is that $x'_{2i} \leq x'_{20}$, or that the demand function

for x_2 be concave.⁷ If h_i is decreasing in θ_2 , so it becomes less attractive to participate in the labor market because that involves higher demand for the taxed commodity x_2 , the efficiency effect will counter the equity effect and the welfare effect of increasing the tax on the necessity will be ambiguous. The following proposition summarizes.

Proposition 3 *Assume a) labor supply varies along the extensive margin so individuals can choose whether to work for a given income, and b) the government imposes a linear progressive income tax on all households whether they participate or not.*

1. *Suppose the non-negativity constraint on x_1 is slack for all skill levels. If the income tax is less progressive than optimal, social welfare can be improved by taxing the necessity good x_2 at a lower rate relative to the luxury good, and vice versa.*
2. *Suppose the non-negativity constraint on x_1 is binding for low skill levels. If the linear progressive tax is optimal, taxing the necessity good will have a beneficial equity effect, but an adverse efficiency effect, and the overall effect will be ambiguous.*

This result that Proposition 2 no longer necessarily applies when households decide whether to participate could be extended to allow participants to vary their labor supply along the intensive margin. There would still be an equity effect favoring a tax on necessities, and an efficiency effect working in the opposite direction reflecting the fact that the tax on the necessity might discourage labor market participation. We next investigate whether these results continue to apply when we allow nonlinear taxes.

7 Nonlinear Income Taxation

So far we have assumed that the government deploys a linear progressive income tax as well as any commodity taxes. We turn now to an example where the government can use a nonlinear income tax. Whereas with a linear progressive income tax, the meaning of greater progressivity is unambiguous, that is not the case with nonlinear taxes: the tax system has many parameters, and there are many different ways of increasing progressivity. We shall present a specific, but intuitively appealing, example in which increasing progressivity involves a one-dimensional policy change. For this example, results corresponding to the previous section apply.

Our example involves the extensive-margin case just considered but with the government assumed to be able to set skill-specific taxes, as in Diamond (1980) and Saez (2002b). Let t_i for $i = 1, \dots, r$ be the tax on workers of skill i who choose to work, while c_0 is the transfer to non-participants. The budget constraint for type- i participants is now $(1 - \theta_1)x_{1i} + (1 + \theta_2)x_{2i} = y_i - t_i$.

⁷This follows because, $x_{20} < x_{2i}$, and $b'(x_{20}) - 1 > b'(x_{2i}) - 1 > 0$.

Using this budget constraint, the participant maximizes utility $(y_i - t_i - (1 + \theta_2)x_{2i}) / (1 + \theta_1) + b(x_{2i})$, resulting in the demand function $x_{2i}((1 + \theta_2) / (1 + \theta_1))$. The indirect utility function is $v_i(\theta_1, \theta_2, t_i)$. Similarly, non-participants maximize $(c_0 - (1 + \theta)x_{20}) / (1 + \theta_1) + b(x_{20})$, giving the demand function $x_{20}((1 + \theta_2) / (1 + \theta_1))$, and the indirect utility function $v_0(\theta_1, \theta_2, c_0)$. Applying the envelope theorem to these indirect utility functions gives:

$$v_{i\theta_1} = -\frac{x_{1i}}{1 + \theta_1}, \quad v_{i\theta_2} = -\frac{x_{2i}}{1 + \theta_1}, \quad v_{it} = -\frac{1}{1 + \theta_1}, \quad v_{0\theta_1} = -\frac{x_{10}}{1 + \theta_1}, \quad v_{0\theta_2} = -\frac{x_{20}}{1 + \theta_1}, \quad v_{0c} = \frac{1}{1 + \theta_1}$$

As above, assuming that the non-negativity constraint on x_1 is not binding for low-skill workers, all individuals demand the same x_2 :

$$x_{2i}\left(\frac{1 + \theta_2}{1 + \theta_1}\right) = x_{20}\left(\frac{1 + \theta_2}{1 + \theta_1}\right) \equiv x_2\left(\frac{1 + \theta_2}{1 + \theta_1}\right)$$

Non-participants obtain a utility of leisure m_i that follows the distribution function $\Gamma_i(m_i)$ as above. The number of workers of skill i is $n_i\Gamma_i(\hat{m}_i) = h_i(x_{1i} + b(x_2) - x_{10} - b(x_2)) = h_i((y_i - t_i - c_0) / (1 + \theta_1))$ since $x_{1i} - x_{10} = (y_i - t_i - c_0) / (1 + \theta_1)$ using the household budget constraints.

Suppose we take as given both the commodity taxes θ_1 and θ_2 and the amount of redistribution to the non-participants c_0 , and assume the government finances the latter by choosing the optimal set of skill-specific taxes t_i ($i > 0$). The government budget constraint is:

$$\sum_{j>0} h_j\theta_1x_{1j} + \theta_2x_2 + \sum_{j>0} h_jt_j - \left(1 - \sum_{j>0} h_j\right)c_0 = 0$$

Using the participants' and non-participants' budget constraints to eliminate x_{1i} and x_{10} from the government budget constraint, the following Lagrangian expression characterizes the problem of the government:

$$\begin{aligned} \mathcal{L} = & \sum_{j>0} h_j \left(\frac{y_j - t_j - c_0}{1 + \theta_1} \right) u(v_j(\theta_1, \theta_2, t_j)) + \sum_{j>0} \int_{m_j > \hat{m}_j} u(v_0(\theta_1, \theta_2, c_0) + m_j) d\Gamma_j(m_j) \\ & + \lambda \left(\sum_{j>0} h_j \left(\frac{y_j - t_j - c_0}{1 + \theta_1} \right) \frac{t_j + \theta_1 y_j}{1 + \theta_1} + \frac{\theta_2 - \theta_1}{1 + \theta_1} x_2 \left(\frac{1 + \theta_2}{1 + \theta_1} \right) - \left(1 - \sum_{j>0} h_j \left(\frac{y_j - t_j - c_0}{1 + \theta_1} \right) \right) \frac{c_0}{1 + \theta_1} \right) \end{aligned}$$

where $u(\cdot)$ is again the social utility function.

Uniform commodity taxes

Begin with the case of uniform commodity taxes. The first-order conditions on t_i evaluated at $\theta_1 = \theta_2 = 0$ are as follows (again ignoring any change in welfare for those whose participation choice changes):

$$h_i(\cdot)u'(v_i)v_{it} + \lambda(h_i(\cdot) - (t_i + c_0)h'_i(\cdot)) = 0 \quad i > 0$$

Using $v_{it} = -1$ and defining the value of marginal social utility of a type- i worker in terms of government revenue as $g_i \equiv u'(v_i) / \lambda$, this can be written:

$$(1 - g_i)h_i(\cdot) = (t_i + c_0)h'_i(\cdot) \quad i > 0 \quad (24)$$

Following Saez (2002b), (24) has a standard interpretation. The left-hand side represents the social gain from transferring a unit of income from type- i workers to the government, since g_i is the value of income in the hands of type- i 's in terms of government revenue. The right-hand side is the loss in revenue from those persons choosing to opt out of the labor market as a result of an increase in t_i . In an optimum, these are equal, leading to optimal tax rates satisfying:

$$\frac{t_i + c_0}{y_i - t_i - c_0} = \frac{1 - g_i}{\eta_i} \quad i > 0$$

where $\eta_i = h'_i(c_i - c_0)/h_i$ is the elasticity of participation. Let the value function for this government problem be given by $\mathcal{W}(c_0, \theta_1, \theta_2)$.

Consider now the effect of increasing the transfer c_0 to all non-participants starting in the above optimal income tax outcome with $\theta_1 = \theta_2 = 0$. We interpret this as equivalent to increasing the progressivity of the tax since it will require an increase in the taxes paid by those in the labor force. Using the envelope theorem, we obtain:

$$\frac{\partial \mathcal{W}}{\partial c_0} = \frac{\partial \mathcal{L}}{\partial c_0} = \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) v_{0c} d\Gamma_j(m_j) - \lambda \left(h_0 + \sum_{j > 0} (t_j + c_0) h'_j \right)$$

Using $v_{0c} = 1$ and defining the marginal social value of income of non-participants as $g_0 \equiv \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) d\Gamma_j(m_j) / (h_0 \lambda)$, this may be written:

$$\frac{1}{\lambda} \frac{\partial \mathcal{W}}{\partial c_0} = (g_0 - 1) h_0 - \sum_{j > 0} (t_j + c_0) h'_j \quad \stackrel{\geq}{\leq} 0 \quad \text{as} \quad c_0^* \stackrel{\geq}{\leq} c_0 \quad (25)$$

where c_0^* is the optimal transfer to those not participating in the labor market. This is intuitive. The first term in (25) represents the equity effect of an increase in c_0 : the net effect on social welfare of transferring a unit of income from the government to each type-0. The second term is the efficiency cost to the government from the increase in the number of non-participants. In an optimum these two are equal, and using (24) we obtain $\sum_{j \geq 0} g_j h_j = 1$ as in Saez (2002b).

Differential commodity taxes

Suppose now that, starting from some given value of c_0 — not necessarily the optimal one — we increase the tax rate θ_2 on good x_2 . Applying the envelope theorem to $\mathcal{W}(c_0, \theta_1, \theta_2)$ and evaluating the result at $\theta_1 = \theta_2 = 0$, we obtain:

$$\left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = \sum_{j > 0} h_j u'(v_j) v_{j\theta} + \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) v_{0\theta} d\Gamma_j(m_j) + \lambda x_2$$

Using $v_{j\theta} = v_{0\theta} = -x_2$ and the definitions of g_j and g_0 , this may be rewritten:

$$\frac{1}{\lambda x_2} \frac{\partial \mathcal{W}}{\partial \theta} = - \sum_{j > 0} h_j g_j - g_0 h_0 + 1$$

From the first-order conditions on t_i , (24), we obtain $\sum_{j>0} h_j g_j = -\sum_{j>0} (t_j + c_0) h'_j - h_0 + 1$, so, drawing on (25), we obtain:

$$\frac{1}{\lambda x_2} \frac{\partial \mathcal{W}}{\partial \theta_2} = \sum_{j>0} (t_j + c_0) h'_j + (1 - g_0) h_0 \quad \geq 0 \quad \text{as} \quad c_0 \geq c_0^*$$

Therefore, the analogue of Proposition 1 applies here as well: if the progressivity of the tax is less than optimal so $c_0 < c_0^*$, social welfare would be improved by subsidizing the necessity good x_2 .

Low-income demand for the necessity income-constrained

Next, suppose lower-wage individuals are income-constrained, and are unable to buy the standard amount of x_2 . Those labor market participants who are constrained maximize utility $x_{1i} + b(x_{2i})$ subject the budget $(1 + \theta_1)x_{1i} + (1 + \theta_2)x_{2i} = y_i - t_i$ and the non-negativity constraint $x_{1i} \geq 0$. As above, this yield the demand function $x_{2i}((1 + \mu_i)(1 + \theta_2)/(1 + \theta_1))$ and the indirect utility $v_i(\theta_1, \theta_2, t_i)$, with $v_{i\theta_1} = -(1 + \mu_i)x_{1i}/(1 + \theta_1)$, $v_{i\theta_2} = -(1 + \mu_i)x_{2i}/(1 + \theta_1)$ and $v_{it} = -(1 + \mu_i)/(1 + \theta_1)$.

Non-participants maximize $x_{10} + b(x_{20})$ subject the budget $(1 + \theta_1)x_{10} + (1 + \theta_2)x_{20} = c_0$ and the non-negativity constraint $x_{10} \geq 0$. Assume that non-participants are income-constrained, so $x_{10} = 0$ and $\mu_0 > 1$. Their demand and indirect utility functions are $x_{10}((1 + \mu_0)(1 + \theta_2)/(1 + \theta_1))$, $x_{20}((1 + \mu_0)(1 + \theta_2)/(1 + \theta_1))$ and $v_0(\theta_1, \theta_2, c_0)$, with $v_{0\theta_1} = -(1 + \mu_0)x_{10}/(1 + \theta_1)$, $v_{0\theta_2} = -(1 + \mu_0)x_{20}/(1 + \theta_1)$ and $v_{0c} = (1 + \mu_0)/(1 + \theta_1)$.

The participation condition for the marginal participant is $x_{1i} + b(x_{2i}) = x_{10} + b(x_{20}) + \hat{m}_i$. Using the individual budget constraint, the number of persons of type i who do participate in the market (those with $m_i < \hat{m}_i$) is given by:

$$h_i \left(x_{1i} + b(x_{2i}) - x_{10} - b(x_{20}) \right) = h_i \left(\frac{y_i - t_i - c_0 - (1 + \theta_2)(x_{2i} - x_{20})}{1 + \theta_1} + b(x_{2i}) - b(x_{20}) \right) \quad (26)$$

Suppose we take θ_1 and θ_2 as given and choose the optimal nonlinear tax-transfer system, $(c_0, t_1, \dots, t_i, \dots)$. The Lagrange function is:

$$\begin{aligned} \mathcal{L} = & \sum_{j>0} h_j(\cdot) u(v_j(\theta_1, \theta_2, t_j)) + \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u(v_0(\theta_1, \theta_2, c_0) + m_j) d\Gamma_j(m_j) \\ & + \lambda \left(\sum_{j \geq 0} h_j(\cdot) \left(\frac{t_j + \theta_1 y_j}{1 + \theta_1} + \frac{\theta_2 - \theta_1}{1 + \theta_1} x_{2j} \left(\frac{(1 + \mu_j)(1 + \theta_2)}{1 + \theta_1} \right) \right) - \left(1 - \sum_{j>0} h_j(\cdot) \right) \frac{c_0}{1 + \theta_1} \right) \end{aligned}$$

The first-order conditions on t_i and c_0 evaluated at $\theta_1 = \theta_2 = 0$ are as follows (again ignoring any change in welfare for those whose participation choice changes):

$$\begin{aligned} h_i u'(v_i) v_{it} + \lambda (h_i - (t_i + c_0) h'_i) &= 0 \quad i > 0 \\ \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) d\Gamma_j(m_j) v_{0c} + \lambda \left(-1 + \sum_{j>0} h_j - \sum_{j>0} h'_j \cdot (t_j + c_0) \right) &= 0 \end{aligned}$$

Using $v_{it} = -(1 + \mu_i)$, $v_{oc} = 1 + \mu_0$, these may be written in the well-known form:

$$(1 - g_i)h_i - (t_i + c_0)h'_i = 0, \quad i > 0 \quad \text{and} \quad g_0h_0 + \sum_{j>0} g_jh_j = 1 \quad (27)$$

where g_i and g_0 are now defined to include the shadow price of the non-negativity constraint as follows:

$$g_i \equiv \frac{u'(v_i)(1 + \mu_i)}{\lambda}, \quad g_0 \equiv \frac{\sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) d\Gamma_j(m_j)(1 + \mu_j)}{h_0 \lambda}$$

Thus, the form of the optimal nonlinear tax is comparable to the standard case where the non-negativity constraints on x_{1i} are not binding.

Let the value function from this optimal tax problem be denoted as $\mathcal{W}(\theta_1, \theta_2)$. Applying the envelope theorem, the effect of a change in θ_2 on social welfare evaluated at $\theta_1 = \theta_2 = 0$ is:

$$\left. \frac{\partial \mathcal{W}}{\partial \theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} = \sum_{j>0} h_j u'(v_j) v_{j\theta_2} + \sum_{j \geq 0} \int_{m_j > \hat{m}_j} u'(v_0 + m_j) d\Gamma_j v_{0\theta_2} + \lambda \left(\sum_{j>0} \frac{dh_j}{d\theta_2} \cdot (t_j + c_0) + \sum_{j \geq 0} h_j x_{2j} \right)$$

Using $v_{j\theta_2} = -(1 + \mu_j)x_{2j}$, $v_{0\theta_2} = -(1 + \mu_0)x_{20}$ and $\lambda = E[u'(1 + \mu)]$ from (27) and the above definitions of g_j and g_0 , this may be written:

$$\begin{aligned} \left. \frac{d\mathcal{W}}{d\theta_2} \right|_{\substack{\theta_1=0 \\ \theta_2=0}} &= -E[u'(1 + \mu)x_2] + E[u'(1 + \mu)]E[x_2] + E[u'(1 + \mu)] \sum \frac{dh_j}{d\theta_2} \cdot (t_j + c_0) \\ &= -\text{Cov}[u'(1 + \mu), x_2] + E[u'(1 + \mu)] \sum \frac{dh_j}{d\theta_2} \cdot (t_j + c_0) \end{aligned} \quad (28)$$

This expression (28) is similar to (22) for the linear tax case. The first term involving the covariance is an equity term, which is positive. It favors a positive tax rate on the necessity good, of which more is consumed by the high-income persons. The second term is an efficiency effect that reflects the fact that imposing a tax on x_2 will affect participation, since non-participants will consume less of the taxed good.

Differentiating (26) with respect to θ_2 yields (23) as above. Therefore, under weak conditions, $dh_j/d\theta_2 < 0$, the second term in (28) will be negative. Depending on the balance between the equity and efficiency terms, it will be welfare-improving to impose a positive or a negative tax on the necessity when the income tax is set optimally. Thus, Proposition 3 applies with nonlinear taxes as well as linear ones. The Atkinson-Stiglitz Theorem does not apply when non-negativity constraints on goods with high-income elasticities of demand are binding for some households.

8 Public Goods and the Marginal Cost of Public Funds

So far, we have taken government revenues to be fixed. Suppose that revenues are used to finance a public good, denoted g , whose level is chosen by the government. Two well-known results in

this context are relevant. First, suppose preferences are weakly separable in the public good and all private commodities including leisure. Then, analogous to the Atkinson-Stiglitz Theorem, if an optimal nonlinear income tax is in place, the optimal level of public goods satisfies the Samuelson conditions (Christiansen 1981; Tuomala 1990; Boadway and Keen 1993). Moreover, Kaplow (1996, 2008) argues that the Samuelson conditions should apply even if the income tax is not set optimally. His argument, analogous to the Konishi-Laroque-Kaplow argument above, is that if the level of public goods does not satisfy the Samuelson conditions, a Pareto improving reform can be achieved by moving g toward that satisfying the Samuelson conditions and financing the change by adjusting the income tax. Second, suppose g is financed by a proportional income tax. Then, Atkinson and Stern (1974) show that the level of public goods should be such that the sum of marginal benefits should equal the marginal cost multiplied by the marginal cost of public funds, where the latter exceeds unity.

In this section, we investigate the decision rule for public goods when the government uses a linear progressive tax. We show that in our example where the Deaton conditions are satisfied, the Samuelson rule should apply if the optimal linear progressive tax is deployed. However, if the tax is less progressive than optimal, the public good decision rule depends on how marginal changes in the public good are financed. If they are financed using changes in the lump-sum component a , the Samuelson rule should apply regardless of whether t is optimal. If changes in t are used, the level of public goods should exceed the Samuelson level, and vice versa. That is, the marginal cost of public funds should be less than or greater than unity as the progressivity of the income tax is greater or less than the optimal progressivity.

We assume in what follows that $\theta_1 = \theta_2 = 0$, so no commodity tax is in place. Following the basic model, let utility now be $x_1 + b(x_2) + d(g) - h(\ell)$, or using the budget constraint, $(1-t)w_i\ell_i + a - x_{2i} + b(x_{2i}) + d(g) - h(\ell_i)$, where $d(g)$ is the benefit from the public good. The indirect utility function can be written $v_i(t, a, g)$, where the envelope condition on g is $v_{ig} = d'$. The government budget constraint now becomes

$$B(t, a, g) \equiv \sum_i n_i \left(tw_i\ell_i((1-t)w_i) - a \right) - pg = 0 \quad (29)$$

where the marginal cost of the public good is p , a constant. Eqs. (4) continue to apply for B_t and B_a with the commodity taxes set to zero, as well as $B_g = -p$.

Consider first the welfare effect of changing the income tax system. Social welfare is given by (6) with indirect utility augmented to include g . Taking g as given, the effect of a change in the tax rate t on social welfare, with a varying to balance the budget, yields (7) for $dW/dt|g$.

Consider now the effect of a change in g . Three cases are of interest, one in which the increment in g is financed by changes in the lump-sum transfer a , a second in which changes in t are used, and a third when a non-negativity constraint applies to the luxury good.

8.1 Effect of a change in g financed by a change in a

The change in social welfare of a small change in the public good g financed by a change in a holding the tax rate t constant is:

$$\left. \frac{dW}{dg} \right|_t = \sum n_i u'_i \left(v_{ig} + v_{ia} \frac{da}{dg} \right) = \sum n_i u'_i \left(v_{ig} - v_{ia} \frac{B_g}{B_a} \right)$$

Using the envelope results and the properties of the budget constraint, this reduces to:

$$\left. \frac{dW}{dg} \right|_t = \sum n_i u'_i (d' - p) \quad (30)$$

Not surprisingly, changes in g financed by a lump-sum tax on all persons can be evaluated by a comparison of the sum of private benefits less the marginal cost, regardless of whether t is optimal. Optimal g satisfies $d' = p$, which is the analog of the Samuelson condition in this context. Let g^* be the level of g that satisfies the Samuelson condition, or $d'(g^*) = p$. (Recall that the total population is normalized to unity.)

8.2 Effect of a change in g financed by a change in t

Suppose now, following Atkinson and Stern (1974), that changes in g are financed by changes in the tax rate t . Differentiating $W(\cdot)$, we obtain

$$\left. \frac{dW}{dg} \right|_a = \sum n_i u'_i \left(v_{ig} + v_{it} \frac{dt}{dg} \right) = \sum n_i u'_i \left(v_{ig} - v_{it} \frac{B_g}{B_t} \right)$$

Rewrite this as follows

$$\left. \frac{dW}{dg} \right|_a = \sum n_i u'_i \left(v_{ig} - v_{ia} \frac{B_g}{B_a} \right) + \frac{B_g}{B_t} \sum n_i u'_i \left(v_{ia} \frac{B_t}{B_a} - v_{it} \right) = \left. \frac{dW}{dg} \right|_t - \frac{B_g}{B_t} \left. \frac{dW}{dt} \right|_g$$

Since $dW/dt \geq 0$ as $t^* \geq t$, we obtain

$$\left. \frac{dW}{dg} \right|_a \geq \left. \frac{dW}{dg} \right|_t \quad \text{as} \quad t^* \geq t \quad (31)$$

Together with (30), (31) implies that if the linear income tax is less progressive than optimal, g should exceed g^* when increments of g are financed by t . This is intuitive since increases in g indirectly move the income tax system closer to the optimum.

8.3 Non-negative constraint on x_1 binding

Finally, suppose $x_{1i} \geq 0$ is binding for at least some low-income households, and let the government set the income tax rate optimally. The government chooses t and a to maximize $\sum n_i u(v_i(t, a, g))$

subject to (29), where now the labor supply function is $\ell((1 + \mu_i)(1 - t)w_i)$. The first-order conditions on t and a are (12) as before, and the same interpretation applies.

The first-order condition on g can be written as follows, using $\lambda = \sum_i n_i u'_i \cdot (1 + \mu_i)$ from (12):

$$\sum_i n_i u'_i v_{ig} - \lambda p = \sum_i n_i u'_i d'(g) - \sum_i n_i u'_i \cdot (1 + \mu_i) p = 0$$

Therefore, the optimal choice of g satisfies:

$$d'(g) = \frac{\sum_i n_i u'_i \cdot (1 + \mu_i)}{\sum_i n_i u'_i} p > p$$

The coefficient on p is like a marginal cost of public funds. It reflects the fact that an increase in revenues required to finance g will increase the distortion on labor supply arising from the binding non-negativity constraint ($\mu_i > 0$ for low-wage persons). This applies even if the income tax is set optimally. Of course, if this constraint is not binding for any individuals, the marginal cost of public funds is unity.

These results can be summarized in the following proposition.

Proposition 4 *Start from uniform commodity taxes ($\theta_1 = \theta_2 = 0$) and a given income tax rate t .*

1. *If changes in g are financed by changes in a , the optimal choice of g is g^* where g^* satisfies the Samuelson condition $d'(g^*) = p$ regardless of whether $t \gtrless t^*$.*
2. *If changes in g are financed by changes in t , the optimal choice of g satisfies $g \gtrless g^*$ as $t^* \gtrless t$.*
3. *If $x_{1i} \geq 0$ is binding for some households the optimal choice of g satisfies $g < g^*$ if $t = t^*$.*

These results have some implications for the debate over the marginal cost of public funds.⁸ In our simple example with linear progressive taxes when the Deaton conditions are satisfied, the marginal cost of raising revenues using the lump-sum component of the income tax is unity, regardless of whether the linear tax is optimal. However, the marginal cost of raising revenues using changes in the tax rate will deviate from unity if the tax is not optimal or if the non-negativity constraint on x_2 is binding for low-income households.

⁸There is some debate about what is meant by the marginal cost of public funds. Jacobs (2010) has argued that whenever income taxes are set optimally, the marginal cost of public funds is unity. He interprets the marginal cost of public funds as the opportunity cost of raising an increment of revenue. However, a unit marginal cost of public funds in his sense need not imply that the Samuelson conditions will be satisfied. For example, increasing the public good above the Samuelson level might relax the incentive constraint. Other authors assume that the marginal cost of public funds will exceed one if the optimal level of the public good is less than the Samuelson level.

9 Concluding Remarks

Our analysis has been conducted using fairly strong assumptions. The quasilinear utility function is the simplest one that both satisfies the Deaton assumptions and makes a stark distinction between a necessity good and a luxury good. In this setting and assuming labor varies along the intensive margin, if an optimal linear progressive tax is implemented and non-negativity constraints are not binding, preferential commodity tax treatment should not be given to goods with low income elasticities of demand. Optimal redistribution can be achieved solely through the income tax. However, if the progressivity of the income tax is less than optimal, for example, to mitigate tax evasion or because of political constraints, social welfare can be improved by differentiating the commodity tax structure in favor of the necessity good. On the other hand, if the non-negativity constraint on the luxury good is binding for low-wage persons, a differential tax on necessities improves social welfare if the optimal income tax is in place.

When labor varies along the extensive margin, the case for taxing necessities when a non-negativity constraint is binding must be modified. There is now an efficiency argument against the differential taxation of necessities because that discourages participation. These results also apply in the extensive-margin case with non-linear income taxes. If the government is constrained from making the optimal transfer to non-participants, preferential tax treatment of necessities is welfare-improving as long as low-wage persons do not face a binding non-negative constraint on their purchases of the luxury good. If such a constraint is binding, there is an equity-efficiency trade-off involved in taxing necessities.

The results extend to the case where the government finances a public good with the tax revenues. As long as the Deaton conditions apply, if the optimal linear income tax is in place, the optimal level of the public good satisfies the Samuelson conditions even though distortionary taxation is used to finance the public good. If the income tax is less progressive than optimal, the Samuelson condition should still apply if public goods increments are financed by changes in the lump-sum component of the linear progressive income tax. However, if public good increases are financed by an increase in the tax rate, the provision of the public good should exceed that satisfying the Samuelson condition. These results contrast with those inspired by Atkinson and Stern (1974) as well as Browning (1976), who argued that the optimal provision of public goods financed by a tax on labor income should satisfy a modified Samuelson condition whereby the sum of marginal benefits should be equated to the marginal cost augmented by a marginal cost of public funds that generally exceeds unity. In our context, we found that the marginal cost of public funds should exceed unity if the non-negativity constraint on x_1 is binding for some households.

Our use of a simple model illustrates these points in the clearest way. Extending the analysis to more general utility functions or to more general nonlinear income tax schedules would make

the model more realistic, but would also make the analysis more complicated and less transparent. One might expect that as long as the income tax is less progressive than optimal, it will be useful to augment it with commodity taxes that favor necessities. And, if the optimal income tax is in place, it will be useful to tax the necessity if low-income persons purchase only that good except to the extent that the taxation of necessities discourages extensive-margin labor supply. If the relevant separability conditions do not apply, there is already an a priori reason for differentiating the commodity tax system in favor of goods that are substitutable with leisure when an optimal income tax is in place.

References

- [1] Atkinson, Anthony B. and Nicholas Stern (1974), ‘Pigou, Taxation and Public Goods’, *Review of Economic Studies* **41**, 119–28.
- [2] Atkinson, Anthony B. and Joseph E. Stiglitz (1976), ‘The Design of Tax Structure: Direct vs. Indirect Taxation’, *Journal of Public Economics* **6**, 55–75.
- [3] Blömquist, Sören and Vidar Christiansen (2008), ‘Taxation and Heterogeneous Preferences’, *FinanzArchiv* **64**, 218–44.
- [4] Boadway, Robin and Firouz Gahvari (2006), ‘Optimal Taxation with Consumption Time as a Leisure or Labor Substitute’, *Journal of Public Economics* **90**, 1851–78.
- [5] Boadway, Robin W. and Michael Keen (1993), ‘Public Goods, Self-Selection and Optimal Income Taxation’, *International Economic Review* **34**, 463–78.
- [6] Boadway, Robin, Maurice Marchand and Pierre Pestieau (1994), ‘Towards a Theory of the Direct-Indirect Tax Mix’, *Journal of Public Economics* **55**, 71–88.
- [7] Boadway, Robin and Pierre Pestieau (2003), ‘Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem’, in Richard Arnott, Bruce Greenwald, Ravi Kanbur and Barry Nalebuff (eds.), *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz* (Cambridge, MA: MIT Press), 387–403.
- [8] Browning, Edgar K. (1976), ‘The Marginal Cost of Public Funds’, *Journal of Political Economy* **84**, 283–98.
- [9] Christiansen, Vidar (1981), ‘Evaluation of Public Projects under Optimal Taxation’, *Review of Economic Studies* **48**, 447–57.
- [10] Christiansen, Vidar (1984), ‘Which Commodity Taxes Should Supplement the Income Tax?’, *Journal of Public Economics* **24**, 195–220.
- [11] Corlett, W.J. and D.C. Hague (1953), ‘Complementarity and the Excess Burden of Taxation’, *Review of Economic Studies* **21**, 21–30.
- [12] Cremer, Helmuth and Firouz Gahvari (1995), ‘Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy’, *Economic Journal* **105**, 1165–79.
- [13] Cremer, Helmuth, Pierre Pestieau and Jean-Charles Rochet (2001), ‘Direct Versus Indirect Taxation: The Design of the Tax Structure Revisited’, *International Economic Review* **42**, 781–99.

- [14] Deaton, Angus (1979), ‘Optimally Uniform Commodity Taxes’, *Economics Letters* **2**, 357–61.
- [15] Diamond, Peter (1980), ‘Income Taxation with Fixed Hours of Work’, *Journal of Public Economics* **13**, 101–10.
- [16] Edwards, Jeremy, Michael Keen and Matti Tuomala (1994), ‘Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide’, *Finanzarchiv* **51**, 472–97.
- [17] Jacobs, Bas (2010), ‘The Marginal Cost of Public Funds is One’, CESifo Working Paper No. 3250, Munich, Germany.
- [18] Kaplow, Louis (1996), ‘The Optimal Supply of Public Goods and the Distortionary Cost of Taxation’, *National Tax Journal* **49**, 513–33.
- [19] Kaplow, Louis (2006), ‘On the Desirability of Commodity Taxation Even When Income Taxation Is Not Optimal’, *Journal of Public Economics* **90**, 1235–50.
- [20] Kaplow, Louis (2008), *The Theory of Taxation and Public Economics*, (Princeton: Princeton University Press).
- [21] Kaplow, Louis (2010), ‘Concavity of Utility, Concavity of Welfare, and Redistribution of Income’ *International Tax Public Finance* **17**, 25–42.
- [22] Konishi, Hideo (1995), ‘A Pareto-Improving Commodity Tax Reform under a Smooth Nonlinear Income Tax’, *Journal of Public Economics* **56**, 413–46.
- [23] Laroque, Guy (2005), ‘Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof’, *Economics Letters* **87**, 141–44.
- [24] Marchand, Maurice, Pierre Pestieau and María Racionero (2003), ‘Optimal Redistribution When Different Workers Are Indistinguishable’, *Canadian Journal of Economics* **36**, 911–22.
- [25] Mirrlees, Sir James, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles and James Poterba (2010), *Tax by Design: The Mirrlees Review* (London: Institute for Fiscal Studies).
- [26] Nava, Mario, Fred Schroyen and Maurice Marchand (1996), ‘Optimal Fiscal and Public Expenditure Policy in a Two-Class Economy’, *Journal of Public Economics* **61**, 119–37.
- [27] Saez, Emmanuel (2002a), ‘The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes’, *Journal of Public Economics* **83**, 217–30.

- [28] Saez, Emmanuel (2002b), ‘Optimal Income Transfer Programs: Intensive vs Extensive Labor Supply Responses’, *Quarterly Journal of Economics* **117**, 1039–73.
- [29] Sheshinski, Eytan (1972), ‘The Optimal Linear Income Tax’, *Review of Economic Studies* **39**, 297-302.
- [30] Tuomala, Matti (1990), *Optimal Income Tax and Redistribution* (Oxford: Clarendon Press).