

Jacobs, Bas; Boadway, Robin

Working Paper

Optimal linear commodity taxation under optimal non-linear income taxation

CESifo Working Paper, No. 4142

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Jacobs, Bas; Boadway, Robin (2013) : Optimal linear commodity taxation under optimal non-linear income taxation, CESifo Working Paper, No. 4142, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/70266>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

Optimal Linear Commodity Taxation under Optimal Non-Linear Income Taxation

Bas Jacobs
Robin Boadway

CESIFO WORKING PAPER NO. 4142
CATEGORY 1: PUBLIC FINANCE
FEBRUARY 2013

Presented at CESifo Area Conference on Public Sector Economics, April 2012

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

Optimal Linear Commodity Taxation under Optimal Non-Linear Income Taxation

Abstract

This paper analyzes optimal linear commodity taxes joint with non-linear income taxes. We provide optimal tax rules based on empirically observable elasticities. We demonstrate that commodities should be taxed/subsidized if doing so boosts labor supply. The critical role of commodity taxation is to alleviate distortions on labor supply caused by income taxation. In addition, we extend the standard formula for optimal non-linear income taxation for the presence of optimal linear commodity taxes. We correct parts of the literature that suggest that the optimal tax rules for commodity taxes derived by Atkinson and Stiglitz (1976, 1980) apply as well to linear tax commodity taxes. We show that the optimal second-best allocation cannot be implemented with linear commodity taxes and non-linear income taxes.

JEL-Code: H200.

Keywords: Atkinson-Stiglitz Theorem, optimal non-linear income taxation, optimal linear and non-linear indirect taxation.

Bas Jacobs
Erasmus School of Economics
Erasmus University Rotterdam
Rotterdam / The Netherlands
bjacobs@ese.eur.nl

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

We are grateful to comments and suggestions by Vidar Christiansen, Jukka Pirttilä, Eric Stephens, and seminar participants at the University of Antwerpen; Erasmus University Rotterdam; CREST Paris; CESifo Public Sector Area Conference, Munich, April 12-14, 2012; Taxation Theory Conference, Nashville Tennessee, May 18-20, 2012. The usual disclaimer applies.

1 Introduction

The Atkinson-Stiglitz theorem is one of the cornerstones in normative public economics. Atkinson and Stiglitz (1976) studied an economy in which households with different wage rates consumed a vector of goods financed by labor income. They demonstrated that when the government can optimize a fully non-linear income tax, indirect tax instruments are superfluous as long as households have homogeneous preferences that are weakly separable between leisure and commodities.¹

When preferences are not separable, differential commodity taxes should be deployed, but the analysis of the optimal tax structure remains opaque.² There is a presumption following the classical textbook treatment in Atkinson and Stiglitz (1980), and subsequently reported in Myles (1995) and Salanié (2011), that the optimal second-best allocations derived by these authors can be decentralized using non-linear income and linear commodity taxes.³ However, as long as quantities of taxed goods are used as control variables, as is done by all authors mentioned, it is implicitly assumed that the government can employ non-linear commodity taxes. We demonstrate that the commodity tax rates implementing the optimal second-best allocation are necessarily different among households, and can therefore not be supported by linear commodity taxes.

Although it is almost 40 years since the seminal contribution of Atkinson and Stiglitz (1976), it remains unclear how *linear* commodity taxes should be set, and how the optimal non-linear income tax should be modified when the government can only optimize linear commodity taxes. This is of theoretical interest, but it is also of practical importance. While the characterization of a fully non-linear system of income and commodity taxes is instructive, it is of limited policy relevance. Since most commodity transactions are anonymous, only linear commodity taxes can practically be levied for many commodities, as emphasized by Guesnerie (1995).

The main purpose of this paper will be to provide a thorough analysis of the optimal income and commodity tax system when preferences are not weakly separable. We follow Mirrlees (1976) who analyzed the same problem as Atkinson and Stiglitz (1976), but in addition allowed for i) general preference structures that depend on ability, and ii) commodities that are subject to either linear or non-linear taxes. We restrict the utility function to be identical across agents to obtain a direct correspondence with the Atkinson-Stiglitz analysis.

We synthesize existing results, correct some misconceptions, and provide some new results and intuitions. We derive expressions for both optimal non-linear and linear commodity taxes

¹Using the same preference assumptions, and starting with a differential commodity tax system and an arbitrary non-linear income tax, Laroque (2005) and Kaplow (2006) demonstrated that eliminating all commodity-tax differentiation, while adjusting the non-linear income-tax schedule such that distributional effects are neutralized, yields a Pareto improvement. When utility does depend on ability, Mirrlees (1976), Saez (2002), and Boadway and Pestieau (2002) demonstrate that the Atkinson-Stiglitz theorem generally breaks down. Similarly, Cremer, Pestieau, and Rochet (2001) demonstrate that the Atkinson-Stiglitz theorem also breaks down when individuals differ in unobserved endowments, besides skills.

²Browning and Meghir (1991), Gordon and Kopczuk (2010), Pirttilä and Suoniemi (2010), and Crawford, Keen, and Smith (2010) all empirically reject weak separability in the utility function, vindicating the relevance of analyzing non-separable preferences in optimal-tax studies.

³In their original paper, Atkinson and Stiglitz (1976, p. 67) explicitly note that they ‘allow for the possibility that the tax rate on commodities may be a function of the level of consumption’. That is, they assume that commodity taxes are in fact non-linear.

under optimal non-linear income taxation. In addition, we derive optimal commodity-tax rules that are only dependent on empirically measurable elasticities. Our results will confirm that when preferences are not weakly separable, commodity taxes should generally be deployed to reduce the distortions associated with redistribution, that is, to improve the efficiency of redistribution. In particular, when the government can employ non-linear commodity taxes there should be a marginal tax (subsidy) on a commodity when this commodity is more complementary to leisure (work) than the numéraire commodity is. In that case, the marginal rate of substitution between this commodity and the numéraire commodity decreases (increases) with labor effort. Non-uniform commodity taxes then help to boost labor supply and thereby offset some of the distortions created by the income tax. Such a policy, however, comes at the cost of distorting commodity demands. This finding is the non-linear counterpart of the classic result of Corlett and Hague (1953) derived for a representative household facing linear taxation.

We show how this intuition applies also in the more restrictive but realistic case where only linear commodity taxes can be implemented. However, the specific optimal commodity tax rules differ in the two cases. We demonstrate that the formulae for optimal commodity taxes very much resemble those from the Ramsey framework with homogeneous agents. In particular, we derive that the ‘index of discouragement’ of taxing commodities is directly related to the benefits of commodity taxes in reducing labor-market distortions. Optimal linear commodity taxes are shown to depend in a complex way on whether the taxed commodities are indeed relative more (or less) complementary to leisure than the numéraire commodity is. We establish that with weakly separable preferences all commodities are equally complementary to leisure as the numéraire commodity, and that commodity taxation is indeed superfluous. Relating weak separability of the utility function to the relative complementarity of commodities with leisure thus sheds a new light on the intuition behind the Atkinson-Stiglitz theorem.

We also provide a characterization of the optimal non-linear income tax system in the presence of optimal non-linear and optimal linear commodity taxes. Optimal non-linear income taxes under optimal non-linear commodity taxes do not fundamentally change compared to Mirrlees (1971), Diamond (1998), and Saez (2001). However, we show how these standard representations must be revised under linear commodity taxation to include a term due to commodity taxes. That term cannot be unambiguously signed, hence it cannot be determined in general whether the optimal use of linear commodity taxation increases or decreases optimal marginal income tax rates. In addition, the zero marginal tax rates at the endpoints of the skill distribution in the absence of bunching reported in Sadka (1976) and Seade (1977) no longer apply. Instead, as in Edwards, Keen, and Tuomala (1994) and Nava, Schroyen, and Marchand (1996), the ‘effective marginal tax rates’ on income, which include the impact of indirect taxes on labor effort, should be zero at the endpoints of the skill distribution.

We can derive more specific results when we impose more structure on the utility function. If we assume that utility is weakly separable between taxed commodities on the one hand, and labor and the untaxed numéraire commodity on the other, and sub-utility from taxed commodities is homothetic, then commodity taxes are optimally uniform but non-zero. Optimal income taxes unambiguously increase as a result of using commodity taxation, for a given desire to redistribute income. Commodity taxes are always positive (negative) if taxed commodities

are more complementary to leisure than the untaxed numéraire commodity. And, marginal income tax rates at the endpoints are unambiguously positive. These results also apply to a setting with only two commodities, i.e. one taxed and one untaxed commodity, as in Nava, Schroyen, and Marchand (1996).

The approach we adopt in this paper to analyze optimal non-linear income and linear commodity taxation and some of the results we derive have precursors in the literature. Most notable are the contributions by Christiansen (1984), Edwards, Keen, and Tuomala (1994), Guesnerie (1995), Nava, Schroyen, and Marchand (1996) and Saez (2002). Christiansen (1984) explores the desirability of introducing small linear commodity taxes alongside the optimal non-linear income tax. Small taxes on goods that are relatively more complementary with leisure than the untaxed numéraire commodity or small subsidies on those commodities that are relatively more substitutable for leisure will improve social welfare. Saez (2002) follows the approach of Christiansen (1984) to analyze the implications of heterogeneous preferences as in Mirrlees (1976) and demonstrates that goods for which high-income earners have a stronger taste should be taxed more.

The studies of Edwards, Keen, and Tuomala (1994) and Nava, Schroyen, and Marchand (1996) are closest to ours. They adapt the two-type Stiglitz (1982) model to the case with multiple goods and leisure to analyze non-linear income taxes with linear commodity taxes. They show that the usefulness of commodity taxes lies in their ability to weaken the incentive constraints. Imposing a higher commodity tax rate on the good most complementary with leisure makes it more difficult for high-skilled workers to mimic low-skilled ones.⁴ They also demonstrate that the zero marginal tax at the top result only apply to the total effective marginal tax rate, and that the optimal income tax at the end points of the earnings distribution is non-zero. We demonstrate that the intuition for the non-zero tax at the top is that the optimal use of commodity taxes always lowers the effective marginal tax burden on work effort. In order to ensure a zero net distortion on labor supply at the top, the optimal income tax must adjust to offset the impact of commodity taxation at the endpoints. Since commodity taxes cannot be differentiated by skill type, there is – by definition – a non-zero distortion on commodity demands at the end points. In the special 2-commodity case analyzed by Nava, Schroyen, and Marchand (1996) we can show that the marginal tax rates at the endpoints are always positive.

More generally, by adopting the continuous-type Mirrlees (1971) framework, we are able to provide a complete characterization of the consequences of non-separable preferences for the structure of commodity taxes and the pattern of the optimal income tax rates over the full range of incomes. This allows us to show how optimal commodity taxes can be depicted in terms of measurable elasticities, and how the familiar *ABC*-type non-linear income tax formula of Diamond (1998) can be adjusted for linear commodity taxes.

The next section sets up the model. The following two sections take up in sequence the cases where commodity taxes can be non-linear and linear. In each case, optimal income taxes and optimal commodity taxes are derived and explained. The last section concludes.

⁴Guesnerie (1995) finds the same result.

2 Model

We follow the setup of Atkinson and Stiglitz (1976). There is a continuum of individuals with mass one distributed by their skill level n according to $F(n)$ with density $f(n)$, for $n \in \mathcal{N} \equiv [\underline{n}, \bar{n}]$, where $\underline{n} > 0$ and $\bar{n} \leq \infty$. The skill level n measures the number of efficiency units of labor of each individual. Assuming that all workers are perfect substitutes in production and that the wage rate per efficiency unit of labor is normalized to one, n corresponds to the wage rate per unit of time worked for a worker of skill n .

Individuals have an identical, strictly concave and twice differentiable utility function given by $u(c, x^1, x^2, \dots, x^I, \ell)$, where c is a numéraire commodity, x^1, x^2, \dots, x^I are other commodities $i = 1, \dots, I$, and ℓ is labor supply. For a type- n individual, we rewrite utility as:

$$u_n \equiv u(c_n, \mathbf{x}_n, \ell_n), \quad \forall n \in \mathcal{N}, \quad (1)$$

where $\mathbf{x}_n = (x_n^1, x_n^2, \dots, x_n^I)$ is the vector of commodities consumed by a type- n individual. The level of utility u_n will serve as a state variable in the analysis to follow. Individuals of skill level n whose labor supply is ℓ_n produce output of $z_n \equiv n\ell_n$.

Producer prices for all commodities are constant and normalized to unity. Suppose the government requires an exogenous amount of resources R . Then, the resource constraint of the economy is given by

$$\int_{\mathcal{N}} \left(z_n - c_n - \sum_{i=1, \dots, I} x_n^i - R \right) dF(n) = 0. \quad (2)$$

Satisfaction of this resource constraint and all the household budget constraints implies that the government budget constraint will hold by Walras' law.

The government is assumed to maximize a sum of concave social utilities, $\Psi(u_n)$:

$$\int_{\mathcal{N}} \Psi(u_n) dF(n), \quad \Psi'(u_n) > 0, \quad \Psi''(u_n) \leq 0. \quad (3)$$

Diminishing private marginal utility of income and non-increasing marginal social welfare $\Psi'(u_n)$ yields a social preference for redistribution. The government is utilitarian if $\Psi'(u_n) = 1$ for all n , and is Rawlsian if $\Psi(u_n) = 0$ for all skill-types n except \underline{n} . If the government could observe individuals' skills, it could implement the first-best outcome by maximizing social welfare (3) subject only to the resource constraint (2). This solution could be supported by a set of skill-specific lump-sum taxes.

In what follows we assume that skills cannot be observed. We begin first by supposing, following Atkinson and Stiglitz (1976), that the government can observe both individual incomes and the quantities of all commodities purchased by each individual. We characterize the second-best planning solution for this case using the revelation principle, and show how the second-best optimum can be decentralized using non-linear income and commodity taxes. As long as preferences are weakly separable in commodities and labor, commodity taxes can be dispensed with entirely. This is of course well known. However, our contribution lies in demonstrating that with non-separable preferences, non-linear commodity taxes are required in order to achieve the second-best optimum, contrary to the impression left by the literature. We then turn to

the analysis of the case where the government can observe individuals' incomes, but can only observe anonymous commodity sales, so that only linear commodity taxes are feasible.

3 Optimal non-linear income and commodity taxation

Since the government can observe income and commodity demands but not skill levels, it cannot rule out individuals of one skill-type choosing a commodity-income bundle intended for another skill-type. To preclude that, the following incentive constraint must be satisfied, cf. Atkinson and Stiglitz (1976):

$$u_n = u(c_n, \mathbf{x}_n, z_n/n) \geq u(c_{n'}, \mathbf{x}_{n'}, z_{n'}/n), \quad \forall n, n' \neq n \in \mathcal{N}. \quad (4)$$

Equivalently, $u_n = \max_{n'} u(c_{n'}, \mathbf{x}_{n'}, z_{n'}/n)$. Adopting the first-order approach, we can apply the envelope theorem with respect to n to obtain the first-order incentive compatibility constraint:

$$\dot{u}_n = -\frac{\ell_n u_\ell(c_n, \mathbf{x}_n, \ell_n)}{n}, \quad \forall n, \quad (5)$$

where the dot denotes a derivative with respect to n . The first-order approach is valid for characterizing the second-best optimum if the Spence-Mirrlees and monotonicity conditions are met. Lemma 1 provides these familiar conditions.

Lemma 1 *Let $U(c_n, \mathbf{X}_n, n) \equiv u(c_n, \mathbf{x}_n, z_n/n)$, where $\mathbf{X} \equiv (x_1, x_2, \dots, x_I, z)$, then the following Spence-Mirrlees and monotonicity conditions must hold at the optimal allocation:*

$$\frac{d(U_{\mathbf{X}}/U_c)}{dn} > 0, \quad \frac{d\mathbf{X}_n}{dn} \geq 0. \quad (6)$$

Proof. See Mirrlees (1976, 334–335). ■

When we analyze non-linear commodity taxes, we assume that Lemma 1 holds. However, when we analyze linear commodity taxes, then we need to assume that the Spence-Mirrlees condition and the monotonicity condition apply to gross labor earnings z only.

The second-best allocation is obtained by maximizing social welfare (3) subject to the resource constraint (2) and the incentive constraint (5). We use ℓ_n and x_n^i as controls, and u_n as a state variable. The numéraire c_n can be treated as endogenous by inverting the utility function (1) to obtain the function $c_n(\mathbf{x}_n, \ell_n, u_n)$. The Lagrangian can then be written as:

$$\begin{aligned} \mathcal{L} \equiv & \int_{\mathcal{N}} \left(\Psi(u_n) + \eta \left(n\ell_n - c_n(\mathbf{x}_n, \ell_n, u_n) - \sum_{i=1, \dots, I} x_n^i - R \right) \right) f(n) dn \\ & + \int_{\mathcal{N}} \theta_n \left(\dot{u}_n + \frac{\ell_n u_\ell(\cdot)}{n} \right) dn, \end{aligned} \quad (7)$$

where η is the Lagrange multiplier on the economy's resource constraint (2), and θ_n is the co-state variable on the incentive constraint (5). Applying partial integration to the term

$\int_{\mathcal{N}} \theta_n \dot{u}_n dn$, the Lagrangian can be rewritten as:

$$\begin{aligned} \mathcal{L} \equiv & \int_{\mathcal{N}} \left(\Psi(u_n) + \eta \left(n\ell_n - c_n(\mathbf{x}_n, \ell_n, u_n) - \sum_{i=1, \dots, I} x_n^i - R \right) \right) f(n) dn \\ & + \int_{\mathcal{N}} \left(\theta_n \frac{\ell_n u_{\ell}(c_n(\mathbf{x}_n, \ell_n, u_n), \mathbf{x}_n, \ell_n)}{n} - u_n \dot{\theta}_n \right) dn + \theta_{\bar{n}} u_{\bar{n}} - \theta_{\underline{n}} u_{\underline{n}}. \end{aligned} \quad (8)$$

The first-order conditions with respect to ℓ_n , x_n^i , and u_n are given by

$$\frac{\partial \mathcal{L}}{\partial \ell_n} = \eta \left(n - \frac{\partial c_n}{\partial \ell_n} \right) f(n) + \frac{\theta_n u_{\ell}}{n} \left(1 + \frac{\ell_n u_{\ell \ell}}{u_{\ell}} + \frac{\ell_n u_{\ell c}}{u_{\ell}} \frac{\partial c_n}{\partial \ell_n} \right) = 0, \quad \forall n, \quad (9)$$

$$\frac{\partial \mathcal{L}}{\partial x_n^i} = -\eta \left(1 + \frac{\partial c_n}{\partial x_n^i} \right) f(n) + \frac{\theta_n \ell_n}{n} \left(u_{\ell c} \frac{\partial c_n}{\partial x_n^i} + u_{\ell x^i} \right) = 0, \quad \forall i, n, \quad (10)$$

$$\frac{\partial \mathcal{L}}{\partial u_n} = \left(\Psi' - \eta \frac{\partial c_n}{\partial u_n} \right) f(n) + \frac{\theta_n u_{\ell}}{n} \frac{\ell_n u_{\ell c}}{u_{\ell}} \frac{\partial c_n}{\partial u_n} - \dot{\theta}_n = 0, \quad \forall n \neq \bar{n}, \underline{n}, \quad (11)$$

$$\frac{\partial \mathcal{L}}{\partial u_{\bar{n}}} = -\theta_{\bar{n}} = 0, \quad \frac{\partial \mathcal{L}}{\partial u_{\underline{n}}} = \theta_{\underline{n}} = 0. \quad (12)$$

These first-order conditions correspond to the characterization of the second-best optimum in Atkinson and Stiglitz (1976).

We next turn to how the optimum can be decentralized by non-linear income and commodity taxation. Let the income tax function be given by $T(z_n)$, and the commodity tax function for commodity x^i be $t_i(x_n^i)$, where derivatives are assumed to be continuous and are denoted $T'(z_n)$ and $t'_i(x_n^i)$. Commodity c remains untaxed since one commodity tax is redundant. In principle, we should have written a single tax function in terms of income and all commodity demands, since the tax imposed on, say, income, generally depends on the level of commodity demands. However, it is more instructive to define separate tax functions for income and each commodity demand, as in the literature.⁵ Assuming separate tax schedules, the household budget constraint is given by

$$c_n + \sum_{i=1, \dots, I} (x_n^i + t_i(x_n^i)) = z_n - T(z_n), \quad \forall n. \quad (13)$$

Maximizing utility $u(c_n, \mathbf{x}_n, z_n/n)$ subject to the household budget constraint yields the following necessary first-order conditions:

$$-\frac{u_{\ell}}{u_c} = (1 - T'(z_n)) n, \quad \frac{u_{x^i}}{u_c} = 1 + t'_i(x_n^i), \quad \forall n, i. \quad (14)$$

The marginal willingness to supply labor increases with the wage rate and decreases with the marginal tax rate on earnings. In addition, the marginal tax rate on commodity x^i distorts the consumption choice away from x^i towards c if the marginal tax is positive (and vice versa if it is negative). These will be useful in what follows. We characterize first the optimal income tax and then optimal non-linear commodity taxes.

⁵A formal proof that these tax schedules indeed implement the optimal second-best allocation is generally missing in the literature. However, recently Renes and Zoutman (2012) have demonstrated that, in economies with a one-dimensional type-space and no externalities, the Spence-Mirrlees and monotonicity conditions of Lemma 1 are sufficient conditions to implement the second-best optimal allocation using separate tax schedules on labor income and commodity demands. Consequently, our separate tax schedules are indeed implementable.

3.1 Optimal non-linear income taxation

The structure of marginal tax rates under the optimal non-linear income tax is given in Proposition 1.

Proposition 1 *The optimal non-linear marginal income tax schedule under optimal non-linear commodity taxes is given by*

$$\frac{T'(z_n)}{1 - T'(z_n)} = \frac{u_c \theta_n}{\eta n f(n)} \left(1 + \frac{1}{\varepsilon_n^*}\right), \quad \forall n, \quad (15)$$

where

$$\varepsilon_n^* \equiv \left(\frac{\partial \ln(-u_\ell/u_c)}{\partial \ln \ell_n}\right)^{-1} = \left(\frac{\partial \ln(1 - T'(z_n)n)}{\partial \ln \ell_n}\right)^{-1} > 0, \quad \forall n, \quad (16)$$

and

$$\frac{\theta_n}{\eta} = \int_n^{\bar{n}} \left(\frac{1}{u_c} - \frac{\Psi'(\cdot)}{\eta}\right) \exp\left[-\int_n^m \left(\frac{\partial \ln u_c}{\partial \ln \ell_s}\right) \frac{ds}{s}\right] f(m) dm > 0, \quad \forall n \neq \bar{n}, \underline{n}. \quad (17)$$

Proof. Recall that $c_n(\mathbf{x}_n, \ell_n, u_n)$ is obtained by inverting the utility function (1). Totally differentiating (1), we obtain $\partial c_n / \partial \ell_n = -u_\ell / u_c = (1 - T')n$ using (14). Note also that $\ell u_{\ell\ell} / u_\ell = \partial \ln u_\ell / \partial \ln \ell$ and $\ell u_{\ell c} / u_c = \partial \ln u_c / \partial \ln \ell$. Substituting all of these into (9), and simplifying yields (15). Similarly, the first-order condition on u_n can be rewritten by totally differentiating utility (1) to obtain: $\partial c_n / \partial u_n = 1 / u_c$, and $\ell u_{\ell c} / u_c = \partial \ln u_c / \partial \ln \ell$. Substituting these into (11) yields a linear differential equation in θ , which can be integrated, using a transversality condition from (12), to find (17). ■

The interpretation of the optimal income tax structure in Proposition 1 is familiar. The variable ε_n^* defined in (16) is the compensated elasticity of labor supply with respect to the net marginal wage rate $(1 - T'(z_n))n$ at skill level n . In what follows, an asterisk will be used to indicate compensated functions. The ratio of the shadow value of the state variable θ_n to the shadow price of government revenue η in (17) equals the marginal value in terms of government revenue of redistributing a marginal unit of revenue from all individuals above skill level n to the government. Substituting (17) into (15) gives the standard expression for the optimal non-linear income tax found in Mirrlees (1971), and discussed further in Seade (1977), Tuomala (1984), Diamond (1998), Saez (2001), or Kaplow (2008). There is no need to discuss it further here.

3.2 Optimal non-linear commodity taxation

Using the same techniques as above, we derive optimal non-linear marginal commodity tax rates reported in Proposition 2.⁶

Proposition 2 *The optimal non-linear commodity taxes under optimal non-linear income taxes*

⁶It might be wondered why the marginal tax rate is written as $t'_i(x_n^i)/(1 + t'_i(x_n^i))$ rather than simply $t'_i(x_n^i)$. The reason is that $t'_i(x_n^i)$ is the ad valorem marginal tax rate in terms of pre-tax (producer) prices, while $t'_i(x_n^i)/(1 + t'_i(x_n^i))$ is defined in terms of after-tax (consumer) prices. By the same token, the marginal income tax rate $T'(z_n)/(1 - T'(z_n))$ in (16) is defined in terms of after-tax income, whereas $T'(z_n)$ is defined in terms of pre-tax income.

are given by

$$\frac{t'_i(x_n^i)}{1+t'_i(x_n^i)} = -\frac{u_c\theta_n}{\eta f(n)n} \frac{\partial \ln(u_{x^i}/u_c)}{\partial \ln \ell_n}, \quad \forall n, i, \quad (18)$$

Proof. Totally differentiating u_n in (1) and using (14), we obtain $\partial c_n/\partial x_n^i = -u_{x^i}/u_c = -(1+t'_i(x_n^i))$. Substituting this into (10) and rearranging yields (18). ■

Since $\theta_n > 0$ by (17), the optimal marginal commodity tax rate on a type- n individual's consumption of x^i has the opposite sign of $\partial \ln(u_{x^i}/u_c)/\partial \ln \ell_n$. If the willingness to pay for x^i in terms of the numéraire decreases (increases) with ℓ , a marginal tax (subsidy) on that commodity boosts labor supply. This reduces the distortion on labor supply created by the non-linear income tax, as stressed by Christiansen (1984). Hence, marginal commodity tax differentiation is desirable as long as $\partial \ln(u_{x^i}/u_c)/\partial \ln \ell_n \neq 0$. (Recall that the tax on c is zero.) Naturally, an indirect tax different from zero distorts the optimal choice of consumption goods. Starting from an equilibrium without indirect taxes, the distortion in consumption choices due to the indirect tax is second-order, whereas the reduction of distortions in labor supply is first-order. At the optimum, the marginal reduction in labor-market distortions equals the marginal increase in goods-market distortions.

Optimal marginal commodity tax rates in (18) can be interpreted as the analogue of the Corlett-Hague rule for the case of non-linear taxation. This is so despite the fact that the Corlett-Hague rule applies to an economy with identical households and linear taxes, while (18) applies with heterogenous households and non-linear taxes. The next proposition substantiates this claim.

Proposition 3 *Let $\rho_{i\ell} \equiv -u_{x^i\ell}u/(u_\ell u_{x^i})$ denote the Hicksian partial elasticity of complementarity between x^i and ℓ , and let $\rho_{c\ell} \equiv -u_{c\ell}u/(u_\ell u_c)$ denote the Hicksian partial elasticity of complementarity between c and ℓ , then the optimal non-linear marginal commodity tax rates are given by:*

$$\frac{t'_i(x_n^i)}{1+t'_i(x_n^i)} = -\frac{u_c\theta_n}{\eta f(n)n} \omega_\ell (\rho_{i\ell} - \rho_{c\ell}), \quad \forall n, i, \quad \forall n, i, \quad (19)$$

where $\omega_\ell \equiv -u_\ell \ell_n/u$ is the share of labor in utility.

Proof. Substitute the definitions for $\rho_{i\ell}$, $\rho_{c\ell}$, and ω_ℓ in (19) and rearrange to establish equivalence with (18). ■

The relationship with the Corlett-Hague rule follows by noting that if $\rho_{i\ell} > \rho_{c\ell}$, x^i is more complementary with labor than c is. Thus, a marginal subsidy on x^i serves to offset part of the distortion of the income tax on labor supply.

Following Edwards, Keen, and Tuomala (1994), Guesnerie (1995), and Nava, Schroyen, and Marchand (1996), the role of commodity taxes can be interpreted in terms of relaxing the incentive constraints. If $\partial(\ln u_{x^i}/u_c)/\partial \ln \ell_n < 0$, a higher-skilled individual who is mimicking a lower-skilled individual has the same disposable income and derives a higher net benefit from x^i so consumes more of it, since he supplies less labor. Increasing the tax on x^i and redistributing the revenue so as to keep the utility of the mimicked individual unchanged makes the mimicking individual worse off, so this policy relaxes the incentive constraint. The government can then increase redistribution through the income tax. The reverse argument holds true for

$\partial(\ln u_{x^i}/u_c)/\partial \ln \ell_n > 0$. Note finally that all marginal commodity taxes are proportional to the multiplier θ on the incentive compatibility constraint.

The next proposition demonstrates that commodity taxes are used primarily for efficiency reasons and not for direct redistribution.⁷

Proposition 4 *The optimal commodity taxes can be expressed as a function of the optimal income taxes as follows:*

$$\frac{t'_i(x_n^i)}{1+t'_i(x_n^i)} \left(\frac{\partial \ln(u_{x^i}/u_c)}{\partial \ln \ell_n} \right)^{-1} = -\frac{T'(z_n)}{1-T'(z_n)} \left(1 + \frac{1}{\varepsilon_n^*} \right)^{-1}, \quad \forall n, i. \quad (20)$$

Proof. Substitute (15) into (18) to obtain (20). ■

Proposition 5 indicates that the use of commodity taxes is related one-to-one with the use of income taxes. The optimal structure of commodity and income taxes follows a sort of non-linear Ramsey rule for income and commodity taxes. It is not directly dependent on the particular social welfare function used, or the distribution of skills in the population (Mirrlees, 1976). Naturally, commodity taxation is indirectly dependent on redistribution as marginal income tax rates increase with stronger redistributive tastes.

Two important special results follow from the above characterization of optimal skill-specific marginal commodity tax rates: the Atkinson-Stiglitz theorem and the general impossibility of optimal linear commodity taxes. Consider these in turn.

3.2.1 Atkinson-Stiglitz theorem

The well-known Atkinson-Stiglitz theorem follows immediately from (18).

Proposition 5 *Non-linear commodity taxes are superfluous if utility is weakly separable between commodities and labor:*

$$u_n \equiv u(h(c_n, \mathbf{x}_n), \ell_n), \quad \forall n. \quad (21)$$

Proof. The result follows from substituting $\partial(\ln u_{x^i}/u_c)/\partial \ln \ell_n = 0$ into (18). ■

The marginal willingness to pay for x^i , or u_{x^i}/u_c , is independent from ℓ_n for all i since $u_{x^i}/u_c = h_{x^i}/h_c$. Therefore, introducing a commodity tax is impotent to stimulate labor supply since all goods are equally complementary to work (leisure), and it only distorts the optimal composition of consumption. Consequently, the optimal commodity tax should be zero.

In terms of incentive constraints, with weakly separable utility of the form in equation (21), the incentive compatibility constraint is given by

$$\dot{u}_n = -\frac{\ell_n u_\ell(h_n, \ell_n)}{n}, \quad \forall n, \quad (22)$$

where $h_n \equiv h(c_n, \mathbf{x}_n)$. Consequently, the incentive compatibility constraint only depends on aggregate sub-utility h_n , but not on its composition over c_n and \mathbf{x}_n . Introducing a distortion

⁷This also follows from Christiansen (1984), who demonstrates that introducing commodity-tax differentiation, starting from zero commodity taxation, while adjusting the income-tax schedule to off-set the distributional effects of commodity taxes, generally yields a Pareto-improvement under non-separable preferences.

in the composition of spending will only produce a lower level of h_n , which is symmetric across all agents. Hence, the incentive constraints cannot be relaxed by distorting the consumption of c_n and \mathbf{x}_n .

3.2.2 Non-optimality of linear commodity taxes

The above analysis uses z_n and \mathbf{x}_n as controls, so implicitly assumes the government can observe them and can apply non-linear taxes to both income and commodity demands. As (20) indicates, optimal marginal commodity tax rates generally differ across households. Some authors suggest that the above formulation also applies when commodity taxes are linear and have the same rate for all households, including Atkinson (1977, p. 597, 601), Atkinson and Stiglitz (1980, p. 435–7), Myles (1995, pp. 163–6), Salanié (2011, pp. 125–7). In particular, they assume that the marginal tax rates in (18) can be interpreted as linear tax rates.⁸ If this were so, it would simplify the analysis and interpretation of optimal linear commodity taxes. Unfortunately, this is not the case if preferences are not weakly separable as the following proposition shows.

Proposition 6 *The optimal second-best allocation cannot be implemented using a non-linear income tax and linear commodity taxes if preferences are not weakly separable in commodities and leisure.*

Proof. The optimal second-best allocation is described by the first-order and transversality conditions (9)–(12) and the resource constraint (2). To implement the optimal second-best allocation using a non-linear income tax and a linear commodity tax, it must be the case that, at the optimal allocation, the marginal rates of substitution between commodity x^i and c are the same for all skill types, i.e., u_{x^i}/u_c is constant across all skill types. However, from the first-order conditions (10), it follows that

$$\frac{u_{x^i}}{u_c} = \left(1 + \frac{u_c \theta_n / \eta}{nf(n)} \frac{\partial \ln(u_{x^i}/u_c)}{\partial \ln \ell_n} \right)^{-1}, \quad \forall n, i. \quad (23)$$

The right-hand side of (23) is obviously not constant across skill types, except with weakly separable preferences, so that u_{x^i}/u_c is independent of ℓ_n . Therefore, in the absence of weak separability, the optimal second-best allocation cannot be implemented with constant tax rates on consumption of goods x^i . ■

Atkinson and Stiglitz (1976, 1980) have derived the condition under which *non-linear* commodity taxes are redundant. In this case, linear commodity taxes are redundant as well. However, suppose the government can only levy linear commodity taxes with a common rate for all individuals, because it is unable to observe individuals' commodity consumption levels. The assumption that consumption taxes should be linear therefore constrains the set of admissible second-best allocations to those allocations where the marginal rates of substitution between c and x^i are identical for all agents. This must be imposed as an additional constraint. In linear-tax models this constraint is the implementability constraint if the primal approach to the optimal-tax problem is employed. See also the original contribution by Ramsey (1927).

⁸See eq. (14-39) in Atkinson and Stiglitz (1980, p. 435–7), eq. (5.96) in Myles (1995, pp. 163–6), and eq. (1) in Salanié (2011, pp. 125–7).

A number of papers have also studied this problem, although the full characterization of the optimal commodity and income tax structure have yet to be worked out. As mentioned, Edwards, Keen, and Tuomala (1994), Guesnerie (1995), and Nava, Schroyen, and Marchand (1996) derive the optimal commodity tax rate in a model with two ability types, and provide an interpretation in terms of complementarity with leisure. Cremer, Pestieau, and Rochet (2001) have generalized the analysis to many goods and many discrete skill types, and have considered some properties of the optimal commodity tax structure, though their main focus is on the case where households differ in more than one characteristic. The following section fully characterizes the structure of linear commodity taxes and the non-linear income tax when there is a continuum of skill types and households differ only in their skill level.

4 Optimal linear commodity taxes under optimal non-linear income taxation

Suppose now that the government is restricted to linear commodity taxes. Instead of adopting the primal approach to find the optimal commodity taxes, we follow Mirrlees (1976), who employs a ‘mixed’ primal-dual approach to determine the optimal income and commodity tax schedules. This approach is ‘mixed’ in the sense that the non-linear income tax is found by determining the optimal quantities of labor and utility, as in Mirrlees (1971), whereas the optimal commodity tax rates are found by determining the optimal prices (rather than quantities) for each commodity, as in Diamond and Mirrlees (1971).

4.1 Individual behavior

All individuals face the same commodity tax rates on goods x^i denoted by t_i . Following the approach initially suggested by Mirrlees (1976) and Christiansen (1984), it is useful to disaggregate individual optimization into two stages. First, a type- n person chooses labor supply ℓ_n , which determines income z_n and disposable income $y_n \equiv z_n - T(z_n)$, given the non-linear income tax function $T(z_n)$. Second, disposable income is allocated among the $I + 1$ commodities, c_n and \mathbf{x}_n . The individual anticipates the outcome of the second stage when choosing labor supply, and it is convenient to start with the second-stage problem.

4.1.1 Stage 2

Given ℓ_n , the household maximizes $u(c_n, \mathbf{x}_n, \ell_n)$ with respect to c_n and x_n^i , subject to

$$c_n + \sum_{i=1, \dots, I} q_i x_n^i = y_n, \quad \forall n, \quad (24)$$

where $q_i \equiv 1 + t_i$ is the consumer price of commodity x^i . The first-order conditions for this partial maximization problem are

$$\frac{u_{x^i}}{u_c} = q_i = 1 + t_i, \quad \forall n, i. \quad (25)$$

The solution yields the conditional commodity demands $c_n(\mathbf{q}, y_n, \ell_n)$ and $\mathbf{x}_n(\mathbf{q}, y_n, \ell_n)$, where $\mathbf{q} \equiv (q_1, q_2, \dots, q_I)$ is the vector of consumer prices.

Substitution in the utility function yields a conditional indirect utility function v_n :

$$v_n \equiv v(\mathbf{q}, y_n, \ell_n) \equiv u(c_n(\cdot), \mathbf{x}_n(\cdot), \ell_n), \quad \forall n. \quad (26)$$

Applying the envelope theorem, we obtain Roy's identity for the conditional indirect utility function:

$$v_{q_i}(\mathbf{q}, y_n, \ell_n) = -v_y(\mathbf{q}, y_n, \ell_n)x_n^i(\mathbf{q}, y_n, \ell_n), \quad \forall n, i, \quad (27)$$

where $v_y(\mathbf{q}, y_n, \ell_n)$ is the private marginal utility of income and is equal to $u_c(\cdot)$. As above, we use v_n as the state variable in the government's optimization problem. For later use, we differentiate Roy's identity with respect to ℓ_n to find:

$$v_{q_i \ell}(\mathbf{q}, y_n, \ell_n) = -v_y \frac{\partial x_n^i(\mathbf{q}, y_n, \ell_n)}{\partial \ell_n} - v_{y \ell}(\mathbf{q}, y_n, \ell_n)x_n^i(\mathbf{q}, y_n, \ell_n). \quad (28)$$

It is also useful for what follows to consider the dual to the above problem: the conditional expenditure-minimizing problem. Consumer n chooses c_n and x_n^i to minimize expenditures $c_n + \sum_i q_i x_n^i$, subject to $u(c_n, \mathbf{x}_n, \ell_n) = v_n$. The solution yields compensated demands $c_n^*(\mathbf{q}, \ell_n, v_n)$ and $x_n^{*i}(\mathbf{q}, \ell_n, v_n)$, and the expenditure function $e(\mathbf{q}, \ell_n, v_n)$, where by the envelope theorem

$$e_{q_i}(\mathbf{q}, \ell_n, v_n) = x_n^{*i}(\mathbf{q}, \ell_n, v_n), \quad e_v(\mathbf{q}, \ell_n, v_n) = \frac{1}{u_c}, \quad e_{\ell}(\mathbf{q}, \ell_n, v_n) = -\frac{u_{\ell}}{u_c}, \quad \forall n, i. \quad (29)$$

In the consumer's optimum, $y_n \equiv e(\mathbf{q}, \ell_n, v_n)$, $c_n^* = c_n$ and $x_n^{*i} = x_n^i$. Note that these compensated demands are conditional in the sense that labor supply is being held constant. In what follows, unless otherwise mentioned, all compensated demands and the expenditure function are conditional.

For later reference, we can derive the following properties of the compensated demand functions $c_n^*(\mathbf{q}, \ell_n, v_n)$ and $x_n^{*i}(\mathbf{q}, \ell_n, v_n)$. The type- n household's conditional budget constraint can be expressed as

$$c_n^*(\mathbf{q}, \ell_n, v_n) + \sum_i (1 + t_i)x_n^{*i}(\mathbf{q}, \ell_n, v_n) = e(\mathbf{q}, \ell_n, v_n), \quad \forall n. \quad (30)$$

Differentiating (30) with respect to prices q_j (recall $q_j \equiv 1 + t_j$), labor ℓ_n , and utility v_n and using the envelope properties of the expenditure function (29) yields:

$$\frac{\partial c_n^*}{\partial q_j} + \sum_{i=1, \dots, I} (1 + t_i) \frac{\partial x_n^{*i}}{\partial q_j} = 0, \quad \forall n, j, \quad (31)$$

$$\frac{\partial c_n^*}{\partial \ell_n} + \sum_{i=1, \dots, I} (1 + t_i) \frac{\partial x_n^{*i}}{\partial \ell_n} = -\frac{u_{\ell}}{u_c}, \quad \forall n, \quad (32)$$

$$\frac{\partial c_n^*}{\partial v_n} + \sum_{i=1, \dots, I} (1 + t_i) \frac{\partial x_n^{*i}}{\partial v_n} = e_v = \frac{1}{u_c}, \quad \forall n. \quad (33)$$

Moreover, we have $x_n^{*i}(\mathbf{q}, \ell_n, v_n) = x_n^i(\mathbf{q}, e(\cdot), \ell_n)$, since compensated and uncompensated conditional demands are the same, and $y_n(\cdot) = e(\cdot)$. Differentiating with respect to v_n and ℓ_n , and using the envelope properties (29) yields

$$\frac{\partial x_n^{*i}}{\partial v_n} = \frac{1}{u_c} \frac{\partial x_n^i}{\partial y_n}, \quad \forall n, i, \quad (34)$$

$$\frac{\partial x_n^{*i}}{\partial \ell_n} = \frac{\partial x_n^i}{\partial \ell_n} + \frac{u_\ell}{u_c} \frac{\partial x_n^i}{\partial y_n}, \quad \forall n, i. \quad (35)$$

4.1.2 Stage 1

In this stage, a type- n household chooses labor supply ℓ_n to maximize the stage-2 partial utility function $v_n \equiv v(\mathbf{q}, y_n, \ell_n)$, subject to the budget constraint $y_n = n\ell_n - T(n\ell_n)$. This yields the first-order condition for labor supply:

$$-\frac{v_\ell(\mathbf{q}, y_n, \ell_n)}{v_y(\mathbf{q}, y_n, \ell_n)} = -\frac{u_\ell}{u_c} = (1 - T'(n\ell_n))n, \quad \forall n. \quad (36)$$

4.2 The government's problem

The government takes as given individual behavior summarized in the conditional indirect utility functions $v_n \equiv v(\mathbf{q}, y_n, \ell_n) = v_n(\mathbf{q}, y_n, z_n/n)$ for all skill-types n . The incentive compatibility constraint can readily be formulated as earlier:

$$\dot{v}_n = -\frac{\ell_n v_\ell(\mathbf{q}, e(\mathbf{q}, \ell_n, v_n), \ell_n)}{n}, \quad \forall n. \quad (37)$$

Again, we introduce a co-state variable θ_n associated with the incentive constraint (37). After multiplying (37) by θ_n , and integrating by parts, we can write the Lagrangian for this optimal control problem as:

$$\begin{aligned} \mathcal{L} \equiv & \int_{\mathcal{N}} \left(\Psi(v_n) + \eta \left(n\ell_n - c_n^*(\mathbf{q}, \ell_n, v_n) - \sum_i x_n^{*i}(\mathbf{q}, \ell_n, v_n) - R \right) \right) f(n) dn \\ & + \int_{\mathcal{N}} \left(\theta_n \frac{\ell_n v_\ell(\mathbf{q}, e(\mathbf{q}, \ell_n, v_n), \ell_n)}{n} - v_n \dot{\theta}_n \right) dn + \theta_{\bar{n}} v_{\bar{n}} - \theta_{\underline{n}} v_{\underline{n}}. \end{aligned} \quad (38)$$

The control variables are now ℓ_n and \mathbf{q} (or \mathbf{t}), and the state variable is v_n . The first-order conditions with respect to the control and state variables are:

$$\frac{\partial \mathcal{L}}{\partial \ell_n} = \eta \left(n - \frac{\partial c_n^*}{\partial \ell_n} - \sum_i \frac{\partial x_n^{*i}}{\partial \ell_n} \right) f(n) + \frac{\theta_n}{n} (v_\ell + \ell_n v_{\ell\ell} + \ell_n v_{\ell e} e_\ell) = 0, \quad \forall n, \quad (39)$$

$$\frac{\partial \mathcal{L}}{\partial q_j} = \int_{\mathcal{N}} \left(-\eta \left(\frac{\partial c_n^*}{\partial q_j} + \sum_i \frac{\partial x_n^{*i}}{\partial q_j} \right) f(n) + \frac{\theta_n \ell_n}{n} (v_{\ell q_j} + v_{\ell e} e_{q_j}) \right) dn = 0, \quad \forall j, \quad (40)$$

$$\frac{\partial \mathcal{L}}{\partial v_n} = \left(\Psi' - \eta \frac{\partial c_n^*}{\partial v_n} - \eta \sum_i \frac{\partial x_n^{*i}}{\partial v_n} \right) f(n) + \frac{\theta_n \ell_n v_{\ell e} e_u}{n} - \dot{\theta}_n = 0, \quad \forall n \neq \underline{n}, \bar{n}, \quad (41)$$

$$\frac{\partial \mathcal{L}}{\partial v_{\underline{n}}} = -\theta_{\underline{n}} = 0, \quad \frac{\partial \mathcal{L}}{\partial v_{\bar{n}}} = \theta_{\bar{n}} = 0. \quad (42)$$

We use these conditions first to characterize the optimal linear commodity tax system, and then the optimal non-linear income tax.

4.3 Optimal linear commodity taxation

To characterize optimal commodity taxes, denote by $s_n^{ij} \equiv \partial x_n^{*i} / \partial q_j$ the substitution effect of taxing commodity i with commodity tax t_j . Let $E[x^j] \equiv \int_{\mathcal{N}} x_n^j f(n) dn$ be the aggregate demand for x^j , and define $\varepsilon_{j\ell}$ as the following uncompensated demand elasticity for x_n^j :⁹

$$\varepsilon_{j\ell} \equiv \frac{\partial x_n^j}{\partial \ell_n} \frac{\ell_n}{x_n^j}, \quad \forall n, j. \quad (43)$$

Armed with these definitions, we can provide the optimal commodity tax structure in Proposition 7.

Proposition 7 *The optimal linear commodity tax structure at the optimal non-linear income tax satisfies*

$$\frac{1}{E[x^j]} \int_{\mathcal{N}} \sum_i t_i s_n^{ji} f(n) dn = \int_{\mathcal{N}} \frac{\theta_n u_c}{\eta n f(n)} \frac{x^j}{E[x^j]} \varepsilon_{j\ell} f(n) dn, \quad \forall j. \quad (44)$$

Proof. Use (28) and note that $v_{y\ell} = v_{\ell e}$, $x_n^{*j} = e_{q_j}$, and $v_y = u_c$. Substitute these results in (40), using (31) and (43), and impose symmetry of the substitution effects, i.e., $s_n^{ij} = s_n^{ji}$. ■

The left-hand side of (44) is similar to that found when all taxes are linear (Atkinson and Stiglitz, 1980, eq. (12-55)) and is analogous to the so-called index of discouragement of commodity j (Mirrlees, 1976, eq. (86)). Roughly speaking, it represents the proportional reduction in the compensated aggregate demand for commodity j when commodity taxes on all commodities are all marginally increased. Alternatively, the left-hand side captures the marginal deadweight loss of distorting the demand of commodity j by marginally increasing all commodity taxes t_i . The right-hand side of (44) measures the marginal reduction in distortions on labor supply by discouraging (encouraging) the demand of commodity x^j . If $\varepsilon_{j\ell} > 0$ then consumption of commodity x^j boosts labor supply. By encouraging the consumption of this commodity, the government indirectly stimulates labor supply, and thereby alleviates the distortions of the income tax on work effort. Demand for commodity x^j will be encouraged more, the larger is $\varepsilon_{j\ell}$. Similarly, when commodity demand for x^j reduces labor supply, $\varepsilon_{j\ell} < 0$, the consumption of this commodity should be discouraged so as to alleviate the distortions of the income tax on labor supply.

Again, our findings establish the close connection between the Corrlet-Hague rule for optimal linear commodity taxation in a revenue-raising setting with homogeneous agents and optimal linear commodity taxation in the Mirrlees framework with optimal non-linear income taxes. Our results illustrate that commodity taxes play largely an efficiency role. The term $u_c \theta_n / (\eta n f(n))$ represents the distortion created by the income tax at skill level n , see also below in equation (58). The larger is the income tax, the greater should be the size of commodity taxes (or subsidies) deployed to alleviate the labor-supply distortions created by the income tax. The

⁹The elasticity $\varepsilon_{j\ell}$ is comparable to a conditional elasticity in the quantity-rationing literature. See Neary and Roberts (1980).

magnitude of the right-hand side depends on the size of θ_n , which is the multiplier on the incentive constraint.

The intuition for the use of commodity taxes is summarized in the following corollary.

Corollary 1 *The total marginal distortion of commodity taxes t^i on commodity demand x^j is positive (negative) if the commodity-weighted elasticity of demand for commodity j with respect to labor supply ℓ is negative (positive), i.e. aggregate average commodity demand decreases (increases) with a larger level of average labor supply.*

Using the terminology of incentive constraints, it is apparent that commodity taxes are employed only to deter the high-skill types from mimicking low-skill types. If the mimickers value good x^j relatively more than the mimicked individuals, then there should be a downward distortion on demand for good x^j . Intuitively, the mimicker reduces his labor supply to mimic a lower skill type. If the mimicker has a larger willingness to pay for from x^j , this good should thus be taxed so as to make mimicking less attractive. As a result, the incentive constraint is relaxed.

4.4 Optimal linear commodity taxes in terms of elasticities

We are able to write the optimal commodity tax structure in terms of elasticities by using the definitions of the compensated elasticities of demand for x_n^j :

$$\varepsilon_{ji}^* \equiv -s_n^{ji} \frac{(1+t_i)}{x_n^j}, \quad \forall i, j. \quad (45)$$

And, let a 'bar' denote a commodity demand-weighted variable, e.g.,

$$\bar{\varepsilon}_{ji}^* \equiv \left[\int_{\mathcal{N}} \varepsilon_{ji}^* x_n^j dF(n) \right] \left[\int_{\mathcal{N}} x_n^j dF(n) \right]^{-1}, \quad \forall i, j. \quad (46)$$

Using these definitions, optimal commodity taxes in (44) can be readily seen to satisfy the following proposition.

Proposition 8 *The optimal linear commodity taxes at optimal non-linear income taxes are given by*

$$\sum_i \frac{t_i}{1+t_i} \bar{\varepsilon}_{ji}^* = \frac{t_j}{1+t_j} \bar{\varepsilon}_{jj}^* + \sum_{i \neq j} \frac{t_i}{1+t_i} \bar{\varepsilon}_{ji}^* = - \overline{\left(\frac{\theta_n u_c}{\eta n f(n)} \varepsilon_{j\ell} \right)}, \quad \forall j. \quad (47)$$

Proof. Substitute (45) into (44) and employ the bar notation as in (46). ■

The structure of commodity taxes relates in a complex way to the compensated own and cross-elasticities of commodity demands with respect to the commodity taxes (ε_{ji}^*) and the uncompensated cross-elasticity of commodity demands with respect to labor supply ($\varepsilon_{j\ell}$). We gain more intuition into the factors determining the optimal commodity-tax structure by formally deriving those elasticities.

To that end, define the shares of each commodity in the utility function as before:

$$\omega_c \equiv \frac{u_c c_n}{u}, \quad \omega_j \equiv \frac{u_{x^j} x_n^j}{u}, \quad \omega_\ell \equiv \frac{-u_\ell \ell_n}{u}, \quad \forall j. \quad (48)$$

Recall the Hicksian elasticities of complementarity:

$$\begin{aligned}\rho_{cc} &\equiv -\frac{u_{cc}u}{u_c^2}, & \rho_{jj} &\equiv -\frac{u_{x^j x^j} u}{u_{x^j}^2}, & \rho_{\ell\ell} &\equiv -\frac{u_{\ell\ell} u}{u_\ell^2}, \\ \rho_{cj} = \rho_{jc} &\equiv -\frac{u_{x^j c} u}{u_{x^j} u_c}, & \rho_{c\ell} = \rho_{\ell c} &\equiv -\frac{u_{c\ell} u}{u_c u_\ell}, \\ \rho_{ij} = \rho_{ji} &\equiv -\frac{u_{x^j x^i} u}{u_{x^j} u_{x^i}}, & \rho_{\ell j} = \rho_{j\ell} &\equiv -\frac{u_{x^j \ell} u}{u_{x^j} u_\ell}, & \forall j.\end{aligned}\quad (49)$$

The cross-elasticities $\varepsilon_{j\ell}$ and the price elasticities ε_{ji}^* follow from the following Lemma.

Lemma 2 *The cross-elasticities $\varepsilon_{j\ell}$ of individual n follow from*

$$\begin{bmatrix} \frac{\omega_1}{\omega_\ell} \varepsilon_{1\ell} \\ \frac{\omega_2}{\omega_\ell} \varepsilon_{2\ell} \\ \vdots \\ \frac{\omega_I}{\omega_\ell} \varepsilon_{I\ell} \end{bmatrix} = \mathbf{S}^{-1} \begin{bmatrix} \rho_{\ell 1} - \rho_{c\ell} \\ \rho_{\ell 2} - \rho_{c\ell} \\ \vdots \\ \rho_{\ell I} - \rho_{c\ell} \end{bmatrix}, \quad \forall n \quad (50)$$

and the compensated price elasticities ε_{ji}^* of individual n follow from

$$\omega_j \varepsilon_{ji}^* = \frac{C_{ji}}{|\mathbf{S}|}, \quad \forall i, j, n. \quad (51)$$

where

$$\mathbf{S} \equiv \begin{bmatrix} \rho_{cc} + \rho_{11} - 2\rho_{c1} & \rho_{cc} + \rho_{12} - \rho_{c1} - \rho_{c2} & \cdots & \rho_{cc} + \rho_{1I} - \rho_{c1} - \rho_{cI} \\ \rho_{cc} + \rho_{21} - \rho_{c2} - \rho_{c1} & \rho_{cc} + \rho_{22} - 2\rho_{c2} & \cdots & \rho_{cc} + \rho_{2I} - \rho_{c2} - \rho_{cI} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{cc} + \rho_{I1} - \rho_{cI} - \rho_{c1} & \rho_{cc} + \rho_{I2} - \rho_{cI} - \rho_{c2} & \cdots & \rho_{cc} + \rho_{II} - 2\rho_{cI} \end{bmatrix}, \quad \forall n. \quad (52)$$

is a positive semi-definite matrix from the second-order conditions and C_{ji} is the co-factor from matrix \mathbf{S} where column i has been replaced by the proportional changes in the I consumer prices, dq_i/q_i .

Proof. See Appendix. ■

From Lemma 2 follows that the cross-elasticities $\varepsilon_{j\ell}$ are determined by two main factors. The first are the cross-substitution effects captured by \mathbf{S} . The second is the pattern of complementarities of commodities with labor as reflected in the vector containing the differences in the elasticities of complementarity of commodity j and the numéraire commodity ($\rho_{\ell j} - \rho_{c\ell}$). We again find that what is crucial for the determination of commodity taxes is whether commodities are more or less complementary to work than the numéraire commodity is, i.e., whether $\rho_{\ell j} \gtrless \rho_{c\ell}$. Lemma 2 proves that $\varepsilon_{j\ell}$ can be signed if the Hicksian elasticities of complementarity $\rho_{\ell j}$ and $\rho_{c\ell}$ are known. This is strongly related to our findings with non-linear commodity taxes in Section 3.2, but not completely the same.

Lemma 2 proves that $\varepsilon_{j\ell}$ cannot be signed unless the entire pattern of Hicksian elasticities of complementarity $\rho_{\ell i}$ for all commodities i is known. This contrasts with the non-linear case, where only the difference in the Hicksian elasticity of complementarity of commodity j with labor and the numéraire commodity with labor determines whether commodity j needs to be taxed or subsidized. To illustrate, suppose that all commodities x^j are relatively more complementary to work effort than the numéraire commodity is. In that case, $\rho_{\ell j} > \rho_{c\ell}$ for

all commodities j , and we can be sure that for all commodities $\varepsilon_{j\ell} > 0$ for all j . In that case, the demand for x^j should be encouraged to alleviate the income-tax distortion on labor supply. The reverse is true when all commodities would be less complementary to work effort than the numéraire commodity.

The utility function determines the signs of the compensated own and cross price elasticities of commodity demands ε_{ji}^* for all i and j . The own compensated price elasticities of demand are always positive (note that the elasticities are defined with a minus sign, see (45), but the compensated cross-price elasticities cannot be signed unambiguously. Again, this depends on whether commodities are Hicksian complements or substitutes in utility.

No general statements can be made regarding the sign of commodity taxes without imposing structure on the utility function. However, it is possible to establish a useful result in the special case where utility is given by $u(c_n, h(\mathbf{x}_n), \ell_n)$ and $h(\cdot)$ is a homothetic sub-utility function over all commodities x^j , see also Sandmo (1974) and Deaton (1979). This result is given in the next proposition.

Proposition 9 *If utility is given by $u(c_n, h(\mathbf{x}_n), \ell_n)$, and $h(\cdot)$ is a homothetic sub-utility function over all commodities x^j , then the optimal linear commodity taxes are uniform ($t = t_j$) and positive (negative), i.e., $t > 0$ ($t < 0$), if $\varepsilon_{i\ell} = \varepsilon_{j\ell} < 0$ (> 0), $\forall j$. The optimal, uniform commodity taxes are given by*

$$\frac{t}{1+t} \left(\overline{\varepsilon_{jj}^*} + \sum_{i \neq j} \overline{\varepsilon_{ji}^*} \right) = - \left(\frac{\theta_n u_c}{\eta n f(n)} \varepsilon_{j\ell} \right), \quad \forall j. \quad (53)$$

Proof. See Appendix. ■

This case provides some additional intuition as to the desirability of commodity taxes on non-numéraire commodities. Indeed, in this special case commodity taxes are set at positive rates ($t > 0$) if doing so boosts labor supply ($\varepsilon_{j\ell} < 0$), and this indirectly alleviates distortions on labor supply caused by income taxation (and vice versa). This special case could not be identified by Edwards, Keen, and Tuomala (1994), since they only allow for two goods, one of which is the numéraire. We note also that the multi-commodity case is quite special, since since the taxed commodities need to be weakly separable from the untaxed numéraire commodity.

The special case of Proposition 9 also applies to the case analyzed by Nava, Schroyen, and Marchand (1996) where there is only one commodity x besides c . In that case the term $\sum_{i \neq j} \overline{\varepsilon_{ji}^*}$ vanishes. Indeed, like Nava, Schroyen, and Marchand (1996) we are able to demonstrate that commodity taxes can be signed unambiguously, depending on the relative complementarity of the commodity with labor compared to the numéraire commodity.

One could generalize this result to the case where the utility function consists of more than one sub-utility function, each of which is homothetic and weakly separable from the numéraire commodity. The goods within each sub-utility function could be aggregated into a composite commodity. There would be no commodity-tax differentiation within each commodity group and the commodity tax system would have a limited number of rates applying to the composite commodities, as is the case in the real world.

4.5 Atkinson-Stiglitz theorem

For completeness we can directly prove that the Atkinson and Stiglitz theorem applies under optimal linear commodity taxes with weakly separable preferences.

Proposition 10 *Linear commodity taxes are superfluous if utility is weakly separable between commodities and labor:*

$$u_n \equiv u(h(c_n, \mathbf{x}_n), \ell_n), \quad \forall n. \quad (54)$$

Proof. In stage 2, consumer n chooses c_n and x_n^i to maximize $u(h(c_n, \mathbf{x}_n), \ell_n)$, subject to $c_n + \sum_i q_i x_n^i = y_n$, given ℓ_n and y_n . First-order conditions for this conditional maximization problem are independent of ℓ_n : $h_{x^i}/h_c = q_i$, $\forall i$. Therefore, the uncompensated conditional commodity demands are functions only of prices q and income y_n , and not of labor supply ℓ_n , so we can write $c_n(\mathbf{q}, y_n)$ and $x_n^i(\mathbf{q}, y_n)$, $\forall i$. Hence, we establish that $\partial x_n^i/\partial \ell_n = 0$, so that $\varepsilon_{i\ell} = 0$, $\forall i$. Substitution of $\varepsilon_{j\ell} = 0$, $\forall j$, in (47) yields $t_j = 0$, $\forall j$. ■

Clearly, when all commodities are equally complementary to labor as the numéraire commodity, we have that $\rho_{\ell j} = \rho_{\ell c}$ for all commodities j , and the elasticities $\varepsilon_{j\ell} = 0$ $\forall j$, by Lemma 2. This finding casts further insight into the Atkinson-Stiglitz theorem. In particular, assuming weakly separable preferences is equivalent to assuming that all commodities are equally complementary to labor supply as the numéraire commodity is.

4.6 Optimal non-linear income taxation

The following variables are useful for expressing the optimal marginal income tax structure. The compensated labor supply elasticity with respect to the net marginal wage rate, defined in (16), can be rewritten as¹⁰

$$\varepsilon_n^* = \left(\frac{\ell_n v_{\ell\ell}}{v_\ell} + \frac{\ell_n v_{\ell e}}{v_\ell} e_\ell \right)^{-1}, \quad \forall n. \quad (55)$$

The net expenditure share of individual n on good x^i is

$$\gamma_n^i \equiv \frac{(1 + t_i)x_n^i}{(1 - T_a(z_n))n\ell_n}, \quad \forall n, \quad (56)$$

where $T_a(z_n) \equiv T(z_n)/z_n$ is the average income tax rate. And, the coefficient of residual income progression (Musgrave and Musgrave, 1979) for individual n is given by

$$\sigma_n \equiv \frac{1 - T'(z_n)}{1 - T_a(z_n)}, \quad \forall n, \quad (57)$$

Using these definitions, the optimal non-linear income tax is given in the following proposition.

Proposition 11 *Optimal non-linear income taxes at optimal commodity taxes are given by*

$$\frac{T'(z_n)}{1 - T'(z_n)} + \frac{1}{\sigma_n} \sum_i \left(\frac{t_i}{1 + t_i} \gamma_n^i \varepsilon_{i\ell} \right) = \frac{u_c \theta_n}{\eta f(n)n} \left(1 + \frac{1}{\varepsilon_n^*} \right), \quad \forall n. \quad (58)$$

¹⁰ $\partial \ln u_\ell / \partial \ln \ell = \ell u_{\ell\ell} / u_\ell$ and $\partial \ln u_c / \partial \ln \ell = \ell u_{\ell c} / u_c$.

Proof. Substitute (32) into (39), use (36), noting that $v_y = u_c$. ■

Equation (58) gives the total effective tax wedge on labor. It differs from the case in (15) with non-linear commodity taxes studied by Atkinson and Stiglitz (1976) and Atkinson and Stiglitz (1980) by the term in commodity taxes. This term will only disappear if utility is weakly separable so $\varepsilon_{i\ell} = 0$ for all i . As (58) shows, the total tax wedge on labor supply is determined not only by the direct marginal tax on labor earnings $T'(z_n)$, but also by the indirect taxes multiplied with their expenditure shares $\gamma_n^i t_i / (1 + t_i)$.

Intuitively, if an individual spends γ_n^i of his net earnings on consumption of x^i , then the tax on x^i creates an additional marginal tax burden on labor of $\gamma_n^i t_i / (1 + t_i)$. In addition, there is a correction for the rate of tax progression through σ_n . The larger is the marginal tax rate above the average tax rate, the lower is the coefficient of residual income progression σ_n . Consequently, the larger is the additional impact of the indirect tax on the total tax wedge on labor effort. Finally, the tax wedge on labor is determined by the elasticity of complementarity between labor ℓ the consumption of x^i . When consumption of x^i is associated with a larger labor supply, a higher tax on x^i imposes a larger distortion on labor supply, implying a larger total tax wedge on labor (and vice versa).

Without imposing structure on the utility function we cannot make unambiguous statements as to whether optimal non-linear income taxes are higher or lower under optimal linear commodity taxes for any given desire to redistribute income – represented by the right-hand side of (58). The reason is that $\varepsilon_{i\ell}$ and t_i can be of the same or opposite signs, as we have demonstrated previously. Hence, the second term on the left-hand side of (58) can both be positive or negative, precluding unambiguous statements on the effect of commodity taxation on level of marginal income taxation. However, in the special case of Proposition 9, $\varepsilon_{i\ell}$ and t_i are of opposite sign, and we can be certain that commodity taxation reduces the effective tax wedge on labor effort. Hence, for a given desire to redistribute income, marginal income taxes increase under optimal commodity taxes. These findings correspond to the two-commodity analyzed by Nava, Schroyen, and Marchand (1996).

The term $\frac{1}{\sigma_n} \sum_i \left(\frac{t_i}{1+t_i} \gamma_n^i \varepsilon_{i\ell} \right)$ in the expression for the optimal income tax is absent under optimal non-linear commodity taxes (see (15)). The reason is that, under linear instruments, there is an additional constraint on the allocation (besides resource and incentive constraints): all marginal rates of substitution between any pair of commodities should be identical for all households. One can associate the additional term with the efficiency loss of using linear commodity taxes relative to employing fully non-linear commodity taxes. For any desire to redistribute income, captured by the right-hand side of (58), the government will therefore set different marginal tax rates on labor income, cf. the left-hand side of (58), if linear commodity taxes are used.

We can more closely relate the optimal income tax in (58) to the now-standard results in Atkinson and Stiglitz (1980), Diamond (1998), and Saez (2001), as the next proposition shows.

Proposition 12 *The marginal social value of redistributing income to type- n persons is given by*

$$\frac{\theta_n}{\eta} = \int_n^{\bar{n}} \left(\frac{\alpha_n}{u_c} - \frac{\Psi'}{\eta} \right) \exp \left[- \int_n^m \left(\frac{\partial \ln u_c}{\partial \ln \ell_n} \right) \frac{ds}{s} \right] f(m) dm, \quad \forall n, \quad (59)$$

where

$$\alpha_n \equiv 1 - \sum_i t_i \frac{\partial x_n^{*i}}{\partial v_n}, \quad \forall n. \quad (60)$$

Proof. From (33), (34), and (35), it follows that

$$\frac{\partial c_n^*}{\partial v_n} + \sum_i \frac{\partial x_n^{*i}}{\partial v_n} = \frac{1}{u_c} - \sum_i t_i \frac{\partial x_n^{*i}}{\partial y_n} \frac{1}{u_c} \equiv \frac{\alpha_n}{u_c}. \quad (61)$$

Using this, and noting that $e_u = 1/v_y$ and $y_n = e(\cdot)$, (41) can be rewritten as:

$$\dot{\theta}_n = (\Psi' - \eta\alpha_n) f(n) + \frac{\theta_n \ell_n v_{\ell y}}{n v_y}. \quad (62)$$

Next substitute $\ell v_{\ell y}/v_y = \partial \ln u_c / \partial \ln \ell$, integrate by parts, and use (42) to find the result. ■

The expression (59) for θ_n/η is basically identical to the standard Mirrlees (1971) expression, except for the presence of α in the first term in brackets. This term just captures the multiple commodity setting that we analyze. It has an intuitive explanation, which is given below.

4.7 Endpoint results

Sadka (1976) and Seade (1977) showed that the marginal tax rates are zero at the endpoints of the skill distribution (in case of a finite upper bound in the skill distribution and no bunching at the bottom skill level). The above results imply that a zero marginal tax rate at the endpoints applies only to the total tax burden on labor. In particular, $\theta_{\underline{n}} = \theta_{\bar{n}} = 0$ by (42), so the total effective tax wedge in (58) becomes zero at $n = \underline{n}$ and $n = \bar{n}$, as also noted by Edwards, Keen, and Tuomala (1994) and Nava, Schroyen, and Marchand (1996). The following proposition replicates these latter results.

Proposition 13 *At the endpoints of the skill distribution, the optimal marginal tax rates on labor income are non-zero:*

$$\frac{T'(z_n)}{1 - T'(z_n)} = -\frac{1}{\sigma_n} \sum_i \left(\frac{t_i}{1 + t_i} \gamma_n^i \varepsilon_{i\ell} \right), \quad n = \underline{n}, \bar{n}. \quad (63)$$

Proof. Substitute (42) into (58). From (47), it follows that $\varepsilon_{i\ell}$ and $t_i/(1 + t_i)$ are of opposite sign if all the individual compensated cross elasticities have the same sign. ■

The reason that optimal rates at the endpoints are non-zero is that the indirect tax part of the tax wedge is always non-zero. Hence, a positive direct tax wedge is needed to keep the total effective tax wedge on labor effort at zero at the endpoints.

Again, no general statements on the level of marginal income taxation can be made at the endpoints, unless specific structure is imposed on the utility function. When the utility function employed in Proposition 9 is used, optimal marginal tax rates at the endpoints are unambiguously positive. This confirms the findings in the two-commodity setting analyzed by Nava, Schroyen, and Marchand (1996).

4.8 ABCD formula for the optimal non-linear income tax

Diamond (1998) has disaggregated the optimal marginal tax rate expression into three multiplicative components labeled A , B and C . The following proposition derives a modified ABC -type formula for the optimal non-linear income tax under optimal linear commodity taxation, which we will label as an $ABCD$ -formula.

Proposition 14 *The optimal income tax expression can be written as an ABCD-formula:*

$$\frac{T'(z_n)}{1 - T'(z_n)} = A_n B_n C_n - D_n, \quad \forall n, \quad (64)$$

where

$$A_n \equiv 1 + \frac{1}{\varepsilon_n^*}, \quad (65)$$

$$B_n \equiv \frac{u_c}{1 - F(n)} \int_n^{\bar{n}} \left(\frac{\alpha_n}{u_c} - \frac{\Psi'}{\eta} \right) \exp \left[- \int_n^m \left(\frac{\partial \ln u_c}{\partial \ln \ell_n} \right) \frac{ds}{s} \right] f(m) dm, \quad (66)$$

$$C_n \equiv \frac{1 - F(n)}{nf(n)}, \quad (67)$$

$$D_n \equiv \frac{1}{\sigma_n} \sum_i \left(\frac{t_i}{1 + t_i} \gamma_n^i \varepsilon_{il} \right). \quad (68)$$

Proof. Substitute (59) in (58) and rearrange to find the result. ■

Eq. (64) generalizes the Diamond ABC formula to include optimal linear commodity taxes in a straightforward way. Note that the A_n and C_n terms are the same as in Diamond's expression.

The B_n term is affected by commodity taxes through the term $\alpha_n \equiv 1 - \sum_i t_i \partial x_n^{*i} / \partial y_n$, which reflects the net revenue raised by extracting one unit of income tax from individual n . Following the tax-perturbation intuition of Saez (2001) and Kaplow (2008), if one increases the marginal income tax rate for type- n individuals, while holding all other marginal tax rates constant, the tax paid by all persons with skills above n must increase by an increment (with no change in their marginal tax rate). The value of the transfer of those revenues to the government is picked up by α . As individuals above n reduce their commodity demands when the government raises the tax burden on them, commodity tax revenue decreases (increases) if commodity taxes are positive (negative). This tax-base effect reduces (increases) the marginal value of raising tax revenue from individuals above type n . The term Ψ'/η is the loss in social welfare measured in terms of government revenue of these same individuals as a result of the unit increase in their income tax payments, that is, their reduction on disposable income y_n .

The additional term D_n captures the change in the effective marginal tax on labor due to the use of commodity taxes. This term can be positive or negative depending on particular cross-substitution patterns and the level of commodity taxation. In the special case of Proposition 9, the D -term is always positive. In that case, commodity taxes $t_i/(1 + t_i)$ and ε_{il} are of opposite sign, and provided that ε_{il} has the same sign for all skill levels, the optimal use of commodity taxes always allows for higher income taxes.

4.9 Reinterpretation of optimal linear commodity taxes

We can use the above results on the optimal effective tax wedge on labor income to further elucidate the nature of optimal commodity taxes. Define the total effective tax wedge on labor as:

$$\frac{\tau'(z_n)}{1 - \tau'(z_n)} \equiv \frac{T'(z_n)}{1 - T'(z_n)} + \frac{1}{\sigma_n} \sum_i \left(\frac{t_i}{1 + t_i} \gamma_n^i \varepsilon_{il} \right), \quad \forall n, \quad (69)$$

Then, the following characterization of optimal commodity tax rates applies.

Proposition 15 *The distortions of optimal commodity tax rates on good x^j can be written as*

$$\frac{t_j}{1 + t_j} \overline{\varepsilon_{jj}^*} + \sum_{i \neq j} \frac{t_i}{1 + t_i} \overline{\varepsilon_{ji}^*} = - \left(\frac{\tau'(z_n)}{1 - \tau'(z_n)} \cdot \frac{\varepsilon_{j\ell}}{1 + 1/\varepsilon_n^*} \right), \quad \forall j. \quad (70)$$

Proof. Using (69) in (58) gives:

$$\frac{\tau'(z_n)}{1 - \tau'(z_n)} \left(1 + \frac{1}{\varepsilon_n^*} \right)^{-1} = \frac{u_c \theta_n}{\eta f(n)n}, \quad \forall n. \quad (71)$$

Substitution in (47) proves the result. ■

The left-hand side of (70) again represents the marginal excess burden of distorting demand for commodity x^j . The right-hand side of (70) is a weighted average of the reduction in marginal labor supply distortions, and represents the labor-market benefits of distorting commodity demands. Labor-supply distortions increase in the total level of taxation on labor (larger $\tau'(z_n)$) and the compensated elasticity of labor supply (larger ε_n^*). The larger are tax wedges on labor or labor supply elasticities, the more useful are commodity taxes to reduce distortions in the labor market. Commodity taxes or subsidies reduce labor supply distortions more the more elastic is labor supply to a change in the commodity demand for x^j , that is, the larger is $\varepsilon_{j\ell}$. Commodity taxes are less attractive for alleviating labor supply distortions the more responsive is the demand for commodity x^j to the commodity tax rates t_i , that is, the larger are ε_{ji}^* .

As earlier, the optimal structure of commodity and income taxes is neither directly dependent on the particular social welfare function used, nor on the distribution of skills in the population. The use of commodity taxes is primarily associated with reducing labor-supply distortions, and not with direct income redistribution.

5 Conclusions

This paper analyzed optimal linear commodity taxes combined with non-linear income taxes. This was partly motivated by a misconception in the literature that the optimal commodity taxes derived by Atkinson and Stiglitz (1976, 1980) apply to linear tax commodity taxes. We have demonstrated that the optimal second-best allocation cannot be implemented with linear commodity taxes and non-linear income taxes. In addition to clarifying that, we have provided a full characterization of the optimal commodity and income tax structure in terms of measurable elasticities. To illustrate the results clearly, we have done so both for the case where non-linear commodity taxes are available and for the more realistic case where only linear commodity taxes

are feasible.

Our results demonstrate that there is a close link between the classical results of Corlett and Hague (1953) and those of Atkinson and Stiglitz (1976, 1980). Indeed, the intuition that goods should be taxed/subsidized if they are more/less complementary with leisure than the untaxed numéraire good fully carries over to the case with optimal non-linear income taxes. Intuitively, commodity taxes are only used for efficiency reasons to offset distortions of the income tax on labor supply by boosting labor supply. Hence, analyzing optimal income redistribution with heterogeneous agents does not change the nature of the Corlett-Hague conclusions.

Moreover, we have adjusted Diamond's *ABC*-formula for optimal non-linear income taxation to take into account the presence of optimal linear commodity taxes. We have shown that optimal commodity taxes are employed only to reduce the distortions of the income tax, but not to directly redistribute incomes. When linear commodity taxes are employed, optimal income tax rates are typically different from those under optimal non-linear commodity taxes, since for any desire to redistribute resources, a linear commodity-tax system creates larger inefficiencies compared to a non-linear commodity tax system. Furthermore, the optimal top rates at the endpoints of the distribution are non-zero.

Our theoretical results could have some important implications. First, as long as individuals have the same utility function and governments can employ non-linear income taxes, commodity-tax differentiation is desirable only if it boosts labor supply. This implication can readily be tested empirically. However, despite the central importance of the Atkinson-Stiglitz theorem in the optimal-tax literature, it is disappointing that not more evidence is available on its empirical validity. As we have noted earlier, Browning and Meghir (1991), Gordon and Kopczuk (2010), Pirttilä and Suoniemi (2010), Crawford, Keen, and Smith (2010) reject weak separability. However, no robust evidence has been presented that (groups of) commodities are strongly associated with labor supply, except, perhaps, for child-care facilities, housing expenditures and capital incomes.

Second, we can recast our model in an intertemporal, life-cycle setting where commodities are consumption levels at different dates. This generalizes the model of Ordober and Phelps (1979) to allow for linear commodity taxes. The results would imply that capital-income taxation is desirable only when doing so stimulates labor supply. The desirability of capital-income taxes thus critically hinges on their impact on labor supply. This is a condition that can also be tested empirically. Estimates by Pirttilä and Suoniemi (2010) indeed suggest that labor supply falls when capital incomes are larger, implying that capital income should be taxed for efficiency reasons. However, more research could be done to explore this issue further.

Appendix

Proof lemma 2

Totally differentiating the first-order conditions for commodity demands $u_{x^j}(c_n, \mathbf{x}_n, \ell_n) = q_j u_c(c_n, \mathbf{x}_n, \ell_n) \forall j$ gives

$$u_{x^j c} dc_n + \sum_i u_{x^j x^i} dx_n^i + u_{x^j \ell} d\ell_n = \frac{u_{x^j}}{u_c} \left(u_{cc} dc_n + \sum_i u_{cx^i} dx_n^i + u_{c\ell} d\ell_n \right) + u_c dq_j, \quad \forall j. \quad (72)$$

We can rewrite this expression by dividing by u_{x^j} and multiplying with some terms:

$$\begin{aligned} & \frac{c_n u_c}{u} \frac{u_{x^j c} u}{u_c u_{x^j}} \frac{dc_n}{c_n} + \sum_i \frac{x_n^i u_{x^i}}{u} \frac{u_{x^j x^i} u}{u_{x^j} u_{x^i}} \frac{dx_n^i}{x_n^i} + \frac{\ell_n u_\ell}{u} \frac{u_{x^j \ell} u}{u_{x^j} u_\ell} \frac{d\ell_n}{\ell_n} \\ &= \frac{c_n u_c}{u} \frac{u_{cc} u}{u_c^2} \frac{dc_n}{c_n} + \sum_i \frac{x_n^i u_{x^i}}{u} \frac{u_{cx^i} u}{u_c u_{x^i}} \frac{dx_n^i}{x_n^i} + \frac{\ell_n u_\ell}{u} \frac{u_{c\ell} u}{u_c u_\ell} \frac{d\ell_n}{\ell_n} + \frac{dq_j}{q_j}, \quad \forall j. \end{aligned} \quad (73)$$

Use a tilde ('~') to denote a relative change:

$$\tilde{c}_n \equiv \frac{dc_n}{c_n}, \quad \tilde{x}_n^j \equiv \frac{dx_n^j}{x_n^j}, \quad \tilde{\ell}_n \equiv \frac{d\ell_n}{\ell_n}, \quad \tilde{q}_j \equiv \frac{dq_j}{q_j} \quad \forall j. \quad (74)$$

Then, using the definitions in (48) and (49), we can rewrite (73) as

$$\sum_i (\rho_{ji} - \rho_{ci}) \omega_i \tilde{x}_n^i = (\rho_{cc} - \rho_{cj}) \omega_c \tilde{c}_n + (\rho_{\ell j} - \rho_{\ell c}) \omega_\ell \tilde{\ell}_n - \tilde{q}_j, \quad \forall j. \quad (75)$$

Derivation $\varepsilon_{j\ell}$

We can totally differentiate the household budget constraint at constant income ($dy_n = 0$), and constant prices ($d\mathbf{q} = 0$), using the first-order conditions $q_j \equiv \frac{u_{x^j}}{u_c} \forall j$ and the definitions (48) to find

$$\omega_c \tilde{c}_n + \sum_i \omega_i \tilde{x}_n^i = 0. \quad (76)$$

Next substitute (76) in (75) to find the following system of equations:

$$\sum_i (\rho_{cc} + \rho_{ji} - \rho_{ci} - \rho_{cj}) \omega_i \tilde{x}_n^i = (\rho_{\ell j} - \rho_{\ell c}) \omega_\ell \tilde{\ell}_n, \quad \forall j. \quad (77)$$

We can find the cross-elasticity $\varepsilon_{j\ell}$ by noting that $\left. \frac{\omega_j \tilde{x}_n^j}{\omega_\ell \tilde{\ell}_n} \right|_{dx_{i \neq j} = 0, d\mathbf{q} = 0} = \frac{\omega_j}{\omega_\ell} \varepsilon_{j\ell}$ since

$$\varepsilon_{j\ell} \equiv \frac{\partial x_n^j}{\partial \ell_n} \frac{\ell_n}{x_n^j} = \frac{dx_n^j / x_n^j}{d\ell_n / \ell_n} \bigg|_{dx_{i \neq j} = 0, d\mathbf{q} = 0} = \frac{\tilde{x}_n^j}{\tilde{\ell}_n}, \quad \forall j. \quad (78)$$

Hence,

$$\sum_i (\rho_{cc} + \rho_{ji} - \rho_{ci} - \rho_{cj}) \frac{\omega_i}{\omega_\ell} \varepsilon_{i\ell} = \rho_{\ell j} - \rho_{\ell c}, \quad \forall j. \quad (79)$$

which can be rewritten in matrix form as

$$\begin{bmatrix} \frac{\omega_1}{\omega_\ell} \varepsilon_{1\ell} \\ \frac{\omega_2}{\omega_\ell} \varepsilon_{2\ell} \\ \vdots \\ \frac{\omega_I}{\omega_\ell} \varepsilon_{I\ell} \end{bmatrix} = \mathbf{S}^{-1} \begin{bmatrix} \rho_{\ell 1} - \rho_{\ell c} \\ \rho_{\ell 2} - \rho_{\ell c} \\ \vdots \\ \rho_{\ell I} - \rho_{\ell c} \end{bmatrix}, \quad (80)$$

where

$$\mathbf{S} \equiv \begin{bmatrix} \rho_{cc} + \rho_{11} - 2\rho_{c1} & \rho_{cc} + \rho_{12} - \rho_{c1} - \rho_{c2} & \cdots & \rho_{cc} + \rho_{1I} - \rho_{c1} - \rho_{cI} \\ \rho_{cc} + \rho_{21} - \rho_{c2} - \rho_{c1} & \rho_{cc} + \rho_{22} - 2\rho_{c2} & \cdots & \rho_{cc} + \rho_{2I} - \rho_{c2} - \rho_{cI} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{cc} + \rho_{I1} - \rho_{cI} - \rho_{c1} & \rho_{cc} + \rho_{I2} - \rho_{cI} - \rho_{c2} & \cdots & \rho_{cc} + \rho_{II} - 2\rho_{cI} \end{bmatrix}. \quad (81)$$

Derivation ε_{ji}^*

We can find the compensated price and cross elasticities of demand by setting $d\ell_n = dv_n = 0$. Totally differentiating the utility constraint $v_n = u(c_n, \mathbf{x}_n, \ell_n)$ and using the definitions (48) gives

$$\omega_c \tilde{c}_n^* + \sum_i \omega_i \tilde{x}_n^{*i} = 0. \quad (82)$$

Substitution in (75) at $d\ell_n = 0$ yields

$$\sum_i (\rho_{cc} + \rho_{ji} - \rho_{ci} - \rho_{cj}) \omega_i \tilde{x}_n^{*i} = -\tilde{q}_j, \quad \forall j. \quad (83)$$

This last equations can be written in matrix form as

$$\begin{bmatrix} \omega_1 \tilde{x}_n^{*1} \\ \omega_2 \tilde{x}_n^{*2} \\ \vdots \\ \omega_I \tilde{x}_n^{*I} \end{bmatrix} = -\mathbf{S}^{-1} \begin{bmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_I \end{bmatrix}. \quad (84)$$

Applying Cramer's rule to (84) yields the elasticities. In particular, let D_j denote the determinant obtained from $|\mathbf{S}|$ by replacing the j -th column with vector $[\tilde{q}_1 \quad \tilde{q}_2 \quad \cdots \quad \tilde{q}_I]'$:

$$D_j \equiv \begin{vmatrix} \rho_{cc} + \rho_{11} - 2\rho_{c1} & \cdots & \rho_{cc} + \rho_{1j-1} - \rho_{c1} - \rho_{c_{j-1}} & \tilde{q}_1 & \rho_{cc} + \rho_{1j+1} - \rho_{c1} - \rho_{c_{j+1}} & \cdots & \rho_{cc} + \rho_{1I} - \rho_{c1} - \rho_{cI} \\ \rho_{cc} + \rho_{21} - \rho_{c2} - \rho_{c1} & \cdots & \rho_{cc} + \rho_{2j-1} - \rho_{c2} - \rho_{c_{j-1}} & \tilde{q}_2 & \rho_{cc} + \rho_{2j+1} - \rho_{c2} - \rho_{c_{j+1}} & \cdots & \rho_{cc} + \rho_{2I} - \rho_{c2} - \rho_{cI} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{cc} + \rho_{I1} - \rho_{cI} - \rho_{c1} & \cdots & \rho_{cc} + \rho_{Ij-1} - \rho_{cI} - \rho_{c_{j-1}} & \tilde{q}_I & \rho_{cc} + \rho_{Ij+1} - \rho_{cI} - \rho_{c_{j+1}} & \cdots & \rho_{cc} + \rho_{II} - 2\rho_{cI} \end{vmatrix}. \quad (85)$$

The relative changes in $\omega_j \tilde{x}_n^{*j}$ are then given by

$$\omega_j \tilde{x}_n^{*j} = -\frac{D_j}{|\mathbf{S}|}, \quad \forall j. \quad (86)$$

Let C_{nm} denote the co-factor of row n and column m in matrix D_j , then we can write for $\omega_j \tilde{x}_n^{*j}$

$$\omega_j \tilde{x}_n^{*j} = -\left(\frac{C_{1j}}{|\mathbf{S}|} \tilde{q}_1 + \frac{C_{2j}}{|\mathbf{S}|} \tilde{q}_2 + \cdots + \frac{C_{Ij}}{|\mathbf{S}|} \tilde{q}_I \right), \quad \forall j. \quad (87)$$

Next, note that the elasticities can be written as:

$$-\omega_j \frac{\tilde{x}_n^j}{\tilde{q}_i} \Big|_{dx_i \neq j = dl=0} = \omega_j \varepsilon_{ji}, \quad \forall i, j, \quad (88)$$

since

$$\varepsilon_{ij} \equiv -\frac{\partial x_n^{*i}}{\partial q_j} \frac{q_j}{x_n^{*i}} = -\frac{dx_n^{*i}/x_n^{*i}}{dq_j/q_j} \Big|_{dx_j \neq i = dl=0} = -\frac{\tilde{x}_n^{*i}}{\tilde{q}_j}, \quad \forall i, j. \quad (89)$$

Using the fact that the matrix D is completely symmetric, we can change rows and columns and have $C_{ij} = C_{ji}$, so that the elasticities are given by:

$$\omega_j \varepsilon_{ji} = \frac{C_{ji}}{|\mathbf{S}|}, \quad \forall i, j. \quad (90)$$

Proof proposition 9

The proof follows Deaton (1979). Homotheticity of h implies that the expenditure shares ω_j are independent from ability type n . Hence, we can multiply both sides of (47) with ω_j for all n . Moreover, suppose that all commodity taxes are constant, $t = t_i = t_j$ then we must be able to write (47) as

$$\frac{t}{1+t} \sum_i \omega_j \bar{\varepsilon}_{ji}^* = -\overline{\left(\frac{\theta_n u_c}{\eta n f(n)} \omega_j \varepsilon_{j\ell} \right)}, \quad \forall j. \quad (91)$$

If commodity tax rates are constant, then i) $\overline{\left(\frac{\theta_n u_c}{\eta n f(n)} \omega_j \varepsilon_{j\ell} \right)}$ should be constant for all j and ii) $\sum_i \omega_j \bar{\varepsilon}_{ji}^*$ should be constant for all j . First, the right-hand side of (91) is shown to be constant in four steps. i) Note that all goods x^j are equally complementary to work, since with weakly separable preferences we have $\rho_{\ell x} \equiv \rho_{\ell j} = \frac{u_{\ell x^j} u}{u_{\ell} u_{x^j}} = \frac{u_{\ell h} u}{u_{\ell} u_h} = \frac{u_{\ell x^i} u}{u_{\ell} u_{x^i}} = \rho_{\ell i}$. $u_{x^j} = u_h h_{x^j}$ and $u_{x^j x^i} = u_{hh} h_{x^i} h_{x^j} + u_h h_{x^j x^i}$ so that the elasticities of complementarity are given by $\rho_{ji} \equiv -\frac{u_{x^j x^i} u}{u_{x^j} u_{x^i}} = -\frac{(u_{hh} h_{x^j} h_{x^i} + u_h h_{x^j x^i}) u}{(u_h)^2 h_{x^j} h_{x^i}} = \rho_{hh} - \frac{\sigma_{ji}}{\omega_h}$, where $\rho_{hh} \equiv -\frac{u u_{hh}}{u_h^2}$, $\sigma_{ji} \equiv \frac{h_{x^j x^i} h}{h_{x^j} h_{x^i}}$, and $\omega_h \equiv \frac{u_h h}{u}$, $\forall i, j$. As a result $\rho_{ji} = \frac{\sigma_{ji}}{\omega_h} - \rho_{hh}$. ii) Note furthermore that $\rho_{ji} = \rho_{ij} \forall i, j$, since $u_{x^j x^i} = u_{x^i x^j}$ so that $\sigma_{ji} = \sigma_{ij}$. With uniform commodity taxes, the elasticities of substitution σ between commodities are identical for all commodities: $\sigma \equiv \sigma_{ji} = \sigma_{ij}$. In particular, for a homothetic sub-utility function $h(\mathbf{x}_n)$ we can find that the relative commodity demands are functions ϕ of their relative prices only $x_j/x_i = \phi(q_j/q_i)$. The function ϕ is the same across skill types. The elasticity of substitution σ is the relative change in the commodity demands with respect to a relative change in prices $\sigma_{ji} \equiv -\partial \ln(x_j/x_i) / \partial \ln(q_j/q_i)$. Substituting for the quantities yields $\sigma_{ji} \equiv -\partial \ln(\phi(q_j/q_i)) / \partial \ln(q_j/q_i)$. With uniform commodity taxes $q_j/q_i = 1$, hence σ_{ji} is constant and equal $\forall i, j$. Therefore, we have $\rho \equiv \rho_{ji} = \rho_{hh} - \frac{\sigma}{\omega_h}$ is constant $\forall i, j$. iii) $\rho_{ch} \equiv \rho_{cj} = \rho_{ci}$, since $\rho_{cj} \equiv \frac{u_{cx^j} u}{u_c u_{x^j}} = \frac{u_{ch} u}{u_c u_h} \equiv \rho_{ch} \forall j$. iv) Using the results from i) – iii) in (50) then results in the following set of equations

$$\sum_i (\rho_{cc} + \rho - 2\rho_{ch}) \frac{\omega_i}{\omega_{\ell}} \varepsilon_{i\ell} = \rho_{\ell x} - \rho_{\ell c}, \quad \forall j. \quad (92)$$

The right-hand side is constant $\forall j$. Hence, this set of equations can only be satisfied if $\frac{\omega_i}{\omega_\ell} \varepsilon_{i\ell}$ is constant $\forall i$, since $(\rho_{cc} + \rho - 2\rho_{ch})$ is constant $\forall j$. Therefore, we have that $\omega_j \varepsilon_{j\ell}$ is indeed constant $\forall j$ and the right-hand side is of (91) constant.

Next, we prove that the left-hand side of (91) is constant in four steps. i) The conditional demand functions for x^j are homogeneous of degree zero in prices, hence we can write

$$\sum_i q_i s_n^{ji} + q_c s_n^{jc} = 0, \quad \forall j, \quad (93)$$

where $s_n^{jc} \equiv \frac{\partial x_j^*}{\partial q_c}$ and $q_c (= 1)$ is the price of the numéraire commodity c . The last equation can be written in elasticity form as

$$\sum_i \varepsilon_{ji}^* = -\varepsilon_{jc}^*, \quad \forall j, \quad (94)$$

where $\varepsilon_{jc}^* \equiv -q_c \frac{s_n^{jc}}{x_n^*}$ is the compensated elasticity of commodity x^j with respect to a change in the price of the numéraire commodity. Consequently, we can write

$$\frac{t}{1+t} \omega_j \varepsilon_{jc}^* = \overline{\left(\frac{\theta_n u_c}{\eta n f(n)} \omega_j \varepsilon_{j\ell} \right)}, \quad \forall j, \quad (95)$$

ii) Matrix \mathbf{S} is given by $\mathbf{S} = (\rho_{cc} + \rho - 2\rho_{ch}) \mathbf{1}$ where $\mathbf{1}$ is an $I \times I$ unit matrix. All the elements of \mathbf{S} are therefore identical. From (51) then follows that all own-price and cross-price elasticities pre-multiplied with their shares are identical: $\omega_i \varepsilon_{ij}^* = \omega_j \varepsilon_{ji}^*, \forall i, j$. Therefore, we find that $\omega_j \varepsilon_{jc}^*$ is the same for all j from the Slutsky equation:

$$\sum_i \omega_j \varepsilon_{ji}^* = -\omega_j \varepsilon_{jc}^*, \quad \forall j, \quad (96)$$

Hence, the left-hand side of (91) is constant $\forall j$. Hence, if $\varepsilon_{j\ell} < 0$, we find that $t > 0$, and vice versa.

References

- Atkinson, Anthony B. 1977. "Optimal Taxation and the Direct versus Indirect Tax Controversy." *Canadian Journal of Economics* 10 (4):590–606.
- Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1-2):55–75.
- . 1980. *Lectures on Public Economics*. London and New York: McGraw-Hill.
- Boadway, Robin and Pierre Pestieau. 2002. "Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem." In *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz*, edited by Richard Arnott, Bruce Greenwald, Ravi Kanbur, and Barry Nalebuff. MIT Press, 387–403.
- Browning, Martin and Costas Meghir. 1991. "The Effects of Male and Female Labor Supply on Commodity Demands." *Econometrica* 59 (4):925–951.

- Christiansen, Vidar. 1984. "Which Commodity Taxes Should Supplement the Income Tax?" *Journal of Public Economics* 24 (2):195–220.
- Corlett, Wilfred J. and Douglas C. Hague. 1953. "Complementarity and the Excess Burden of Taxation." *Review of Economic Studies* 21 (1):21–30.
- Crawford, Ian, Michael Keen, and Stephen Smith. 2010. "Value Added Taxes and Excises." In *The Mirrlees Review – Dimensions of Tax Design*, edited by James A. Mirrlees, Stuart Adam, Timothy J. Besley, Richard Blundell, Steven Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James M. Poterba, chap. 4. Oxford: Oxford University Press, 275–422.
- Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet. 2001. "Direct versus Indirect Taxation: The Design of Tax Structure Revisited." *International Economic Review* 42 (3):781–799.
- Deaton, Angus. 1979. "Optimally Uniform Commodity Taxes." *Economics Letters* 2 (4):357–361.
- Diamond, Peter A. 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1):83–95.
- Diamond, Peter A. and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1):8–27.
- Edwards, Jeremy, Michael Keen, and Matti Tuomala. 1994. "Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide." *FinanzArchiv* 51 (4):472–487.
- Gordon, Roger H. and Wojciech Kopczuk. 2010. "The Choice of the Personal Income Tax Base." mimeo, Columbia University.
- Guesnerie, Roger. 1995. *A Contribution to the Pure Theory of Taxation*. Cambridge-UK: Cambridge University Press.
- Kaplow, Louis. 2006. "On the Undesirability of Commodity Taxation even when Income Taxation is not Optimal." *Journal of Public Economics* 90 (6-7):1235–1250.
- . 2008. *The Theory of Taxation and Public Economics*. Princeton and Oxford: Princeton University Press.
- Laroque, Guy R. 2005. "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof." *Economics Letters* 87 (1):141–144.
- Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2):175–208.
- . 1976. "Optimal Tax Theory: A Synthesis." *Journal of Public Economics* 6 (4):327–358.
- Musgrave, Richard A. and Peggy B. Musgrave. 1979. *Public Finance in Theory and in Practice*. New York: McGraw-Hill.

- Myles, Gareth D. 1995. *Public Economics*. Cambridge-UK: Cambridge University Press.
- Nava, Mario, Fred Schroyen, and Maurice Marchand. 1996. “Optimal Fiscal Policy and Expenditure in a Two Class Economy.” *Journal of Public Economics* 61 (1):119–137.
- Neary, J.Peter and Kevin W.S. Roberts. 1980. “The Theory of Household Behaviour under Rationing.” *European Economic Review* 13 (1):25–42.
- Ordovery, Janush A. and Edmund S. Phelps. 1979. “The Concept of Optimal Taxation in the Overlapping-Generations Model of Capital and Wealth.” *Journal of Public Economics* 12 (1):1–26.
- Pirttilä, Jukka and Ilpo Suoniemi. 2010. “Public Provision, Commodity Demand and Hours of Work: An Empirical Analysis.” Labour Institute for Economic Research Discussion Paper No. 261.
- Ramsey, Frank P. 1927. “A Contribution to the Theory of Taxation.” *Economic Journal* 37 (145):47–61.
- Renes, Sander S. and Floris Zoutman. 2012. “Extending the Atkinson-Stiglitz Theorem: Implementing Second-Best Allocations Through the Simplest Tax Schedule.” Mimeo, Erasmus University Rotterdam.
- Sadka, Efraim. 1976. “On Income Distribution, Incentive Effects and Optimal Income Taxation.” *Review of Economic Studies* 43 (2):261–267.
- Saez, Emmanuel. 2001. “Using Elasticities to Derive Optimal Income Tax Rates.” *Review of Economic Studies* 68 (1):205–229.
- . 2002. “The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes.” *Journal of Public Economics* 83 (2):217–230.
- Salanié, Bernard. 2011. *The Economics of Taxation, Second Edition*. Cambridge, MA: MIT Press.
- Sandmo, Agnar. 1974. “A Note on the Structure of Optimal Taxation.” *American Economic Review* 64 (4):701–706.
- Seade, Jesus K. 1977. “On the Shape of Optimal Tax Schedules.” *Journal of Public Economics* 7 (2):203–235.
- Stiglitz, Joseph E. 1982. “Self-Selection and Pareto Efficient Taxation.” *Journal of Public Economics* 17 (2):213–240.
- Tuomala, Matti. 1984. “On the Optimal Income Taxation: Some Further Numerical Results.” *Journal of Public Economics* 23 (3):351–366.