Optimal Linear Commodity Taxation under Optimal Non-Linear Income Taxation

Bas Jacobs*and Robin Boadway†

May 14, 2012

Preliminary draft. Comments welcome.

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: H2
Key words: Atkinson-Stiglitz Theorem, optimal non-linear income taxation, optimal linear and non-linear indirect taxation

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

---

*Erasmus University Rotterdam, Tinbergen Institute, Netspar and CESifo. Address: Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Phone: +31–10–4081481/1491. Fax: +31–10–4089166. E-mail: bjacobs@ese.eur.nl. Homepage: http://people.few.eur.nl/bjacobs.

†Department of Economics, Queen’s University, Kingston, Ontario, Canada, K7L 3N6. Phone: +1–613–5332266. Fax: +1–613–5336668. E-mail: boadwayr@econ.queensu.ca. Homepage: http://econ.queensu.ca/faculty/boadway/.

1When 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 et al. (2001)
When preferences are not separable, differential commodity taxes should be deployed, but the analysis of the optimal tax structure remains somewhat opaque. There is a presumption, following the classical textbook treatment in Atkinson and Stiglitz (1980) and subsequently reported in Myles (1995) and Salanié (2011), that higher taxes should be imposed on goods that are more complementary with leisure. However, as we shall argue, their analyses implicitly assume that the government can impose non-linear commodity taxes since the control variables they use include quantities of taxed goods. These authors aim to decentralize second-best optimal allocations using non-linear income and linear commodity taxes. However, as we demonstrate in this paper, the commodity tax rates that implement the second-best allocation are necessarily different among households, and can therefore only be supported by non-linear commodity taxes.\footnote{In 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.}

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, as emphasized by Guesnerie (1995). Accordingly, it is of practical importance to clarify exactly how optimal linear commodity taxes should be set alongside the optimal non-linear income tax. Moreover, it is also relevant to understand how the optimal non-linear income tax should be modified when the government also optimizes optimal linear commodity taxes.

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 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, in the ambitious case where 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 formulation for demonstrate that the Atkinson-Stiglitz theorem also breaks down when individuals differ in unobserved endowments, besides skills.
optimal commodity taxes 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 et al. (1994) and Nava et al. (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 et al. (1996). We note, however, that this case can be quite special in settings with more than two commodities as analyzed by Edwards et al. (1994), since the presumed separability structure depends on the way in which we normalize the tax system.

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 et al. (1994), Guesnerie (1995), Nava et al. (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 et al. (1994) and Nava et al. (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. Guesnerie (1995) finds the same result. 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 and 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 et al. (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.

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 \( n \in \mathcal{N} = [\underline{n}, \overline{n}] \), where \( \underline{n} > 0 \) and \( \overline{n} \leq \infty \). 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 \( u(c, x^1, x^2, \ldots, x^I, \ell) \), where \( c \) is a numéraire commodity, \( x^1, x^2, \ldots, x^I \) are other commodities \( i = 1, \cdots, I \), and \( \ell \) is labor supply. For a type-\( n \) individual, we rewrite utility as:

\[
    u_n \equiv u(c_n, x_n, \ell_n), \quad \forall n \in \mathcal{N},
\]

where \( x_n = (x^1_n, x^2_n, \ldots, x^I_n) \) 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 \( \ell_n \) 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 amount of resources \( R \). Then, the resource constraint of the
economy is given by
\[ \int_{\mathcal{N}} \left( z_n - c_n - \sum_{i=1,\ldots,I} x^i_n - R \right) dF(n) = 0. \] (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, \( \Psi(u_n) \):
\[ \int_{\mathcal{N}} \Psi(u_n) dF(n), \quad \Psi'(u_n) > 0, \quad \Psi''(u_n) \leq 0. \] (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 \( n \). 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. However, 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):
\[ u_n = u(c_n, x_n, z_n/n) \geq u(c_{n'}, x_{n'}, z_{n'}/n), \quad \forall n, n' \neq n \in \mathcal{N}. \] (4)

Equivalently, \( u_n = \max_{n'} u(c_{n'}, x_{n'}, z_{n'}/n) \). Adopting the first-order approach, we can apply the envelope theorem with respect to \( n \) to obtain the first-order incentive compatibility constraint:
\[ u_n = - \ell_n u_\ell(c_n, x_n, \ell_n)/n, \] (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 $U(X_n,n) \equiv u(c_n,x_n,z_n/n)$, where $X \equiv (z,x_1,x_2,\ldots,x_I)$, then the following Spence-Mirrlees and monotonicity conditions must hold at the optimal allocation:

$$\frac{d(U_X/U_n)}{dn} > 0, \quad \frac{dX_n}{dn} \geq 0.$$ (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 applies 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 $\ell_n$ and $x^i_n$ 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 $c_n(x_n,\ell_n,u_n)$. The Lagrangian can then be written as:

$$\mathcal{L} \equiv \int_N \left( \Psi(u_n) + \eta\left(n\ell_n - c_n(x_n,\ell_n,u_n) - \sum_{i=1,\ldots,I} x^i_n - R \right) \right) f(n)dn$$ (7)

+ $\int_N \theta_n \left( \dot{u}_n + \frac{\ell_n u_{\ell}()}{n} \right)dn$,

where $\eta$ is the Lagrange multiplier on the economy’s resource constraint (2), and $\theta_n$ is the co-state variable on the incentive constraint (5). Applying partial integration to the term $\int_N \theta_n \dot{u}_n dn$, the Lagrangian can be rewritten as:

$$\mathcal{L} \equiv \int_N \left( \Psi(u_n) + \eta\left(n\ell_n - c_n(x_n,\ell_n,u_n) - \sum_{i=1,\ldots,I} x^i_n - R \right) \right) f(n)dn$$ (8)

+ $\int_N \left( \theta_n \frac{\ell_n u_{\ell}()n}{c_n(x_n,\ell_n,u_n),x_n,\ell_n} - u_n \dot{\theta}_n \right)dn + \theta_n u_{\tau} - \theta_{\pi} u_{\pi}$.

The first-order conditions with respect to $\ell_n$, $x^i_n$, and $u_n$ are given by

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

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

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

$$\frac{\partial \mathcal{L}}{\partial u_{\pi}} = -\theta_{\pi} = 0, \quad \frac{\partial \mathcal{L}}{\partial u_{\pi}} = \theta_{\tau} = 0.$$ (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 $t_i(x^i_n)$, where derivatives are assumed to be continuous and are denoted $T'(z_n)$ and $t_i'(x^i_n)$. 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. Assuming separate tax schedules, the household budget constraint is given by

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

Maximizing utility \( u(c_n, x_n, z_n/n) \) subject to the household budget constraint yields the following necessary first-order conditions:

\[ -\frac{u_{\ell}}{u_c} = (1 - T'(z_n)) n, \quad \frac{u_{x_i}}{u_c} = 1 + t_i'(x_n^i), \quad \forall n, i. \] (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 \( c \) 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.

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

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

where

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

and

\[ \frac{\theta_n}{\eta} = \int_{\pi_n} -\left( \frac{1}{u_c} - \frac{\Psi'(\cdot)}{\eta} \right) \exp \left[ -\int_{\pi_n} \left( \frac{\partial \ln u_c}{\partial \ln \ell} \right) ds \right] f(m)dm > 0, \quad \forall n \neq \pi, \pi. \] (17)

**Proof.** Recall that \( c_n(x_n, \ell_n, u_n) \) 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 \( \ell u_{\ell}/u_{\ell} = \partial \ln u_{\ell}/\partial \ln \ell \) and \( \ell u_{\ell}/u_c = \partial \ln u_c/\partial \ln \ell \). 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 \( \ell u_{\ell}/u_c = \partial \ln u_c/\partial \ln \ell \). Substituting these into (11) yields a linear differential equation in \( \theta \), which can be integrated, using a transversality condition from (12), to find (17). □

\[ ^3 \text{A formal proof that these tax schedules indeed implement the optimal second-best allocation is generally missing in the literature. However, recently Renes and Zoutman (2012) have demonstrated that, in economies with a one-dimensional type-space and no externalities, the Spence-Mirrlees and monotonicity conditions of Lemma 1 are sufficient conditions to implement the second-best optimal allocation using separate tax schedules on labor income and commodity demands. Consequently, our separate tax schedules are indeed implementable.} \]
The interpretation of the optimal income tax structure in Proposition 1 is familiar. The variable $\varepsilon_n^t$ 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 $\theta_n$ to the shadow price of government revenue $\eta$ 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), or Saez (2001). A detailed and intuitive interpretation is given in 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.\(^4\)

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

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

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

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 $\partial \ln (u_x/u_c)/\partial \ln \ell_n$. If the willingness to pay for $x^i$ in terms of the numéraire decreases (increases) with $\ell$, 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 $\partial \ln (u_x/u_c)/\partial \ln \ell_n \neq 0$. (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 heterogeneous 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 $\ell$, and let $\rho_{c\ell} \equiv -u_{x^c\ell}u/(u_{\ell}u_{x^c})$ denote the Hicksian partial

---

\(^4\)It might be wondered why the marginal tax rate is written as $t'_i(x^i_n)/(1 + t'_i(x^i_n))$ rather than simply $t'_i(x^i_n)$. The reason is that $t'_i(x^i_n)$ is the ad valorem marginal tax rate in terms of pre-tax (producer) prices, while $t'_i(x^i_n)/(1 + t'_i(x^i_n))$ 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.
elasticity of complementarity between \(c\) and \(\ell\), then the optimal non-linear marginal commodity tax rates are given by:

\[
\frac{t'_i(x^i_n)}{1 + t'_i(x^i_n)} = - \frac{u_i \theta_n}{\eta f(n)n} \omega_i (\rho_i - \rho_{i\ell}), \quad \forall n, i,
\]

(19)

where \(\omega_i \equiv -u_i \ell_n / u\) is the share of labor in utility.

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

The relationship with the Corlett-Hague rule follows by noting that if \(\rho_i > \rho_{i\ell}\), \(x^i\) 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 et al. (1994), Guesnerie (1995), and Nava et al. (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-skill 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 \(\theta\) on the incentive compatibility constraint.

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

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

\[
\frac{t'_i(x^i_n)}{1 + t'_i(x^i_n)} = - \frac{T'(z_n)}{1 - T'(z_n)} \left(1 + \frac{1}{\varepsilon^*_n}\right)^{-1} \frac{\partial \ln (u_{x^i}/u_c)}{\partial \ln \ell_n}. 
\]

(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).

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:

\[
u_n \equiv u(h(c_n, x_n), \ell_n), \quad \forall n.
\]

(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 $\ell_n$ 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

$$\dot{u}_n = -\ell_n u_{\ell} (h_n, \ell_n),$$

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

3.2.2 Non-optimality of linear commodity taxes

The above analysis uses $z_n$ and $x_n$ 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.

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). 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

$$\frac{u_{x^i}}{u_c} = \left(1 + \frac{u_c/\theta_{n}/\eta}{nf(n)} \left. \frac{\partial \ln (u_{x^i}/u_c)}{\partial \ln \ell_n} \right) \right)^{-1}, \forall n, i.$$  (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 $\ell_n$. 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. A number of papers have studied this problem, although the full characterization of the optimal commodity and income tax structure have yet to be worked out. As mentioned, Edwards et al. (1994), Guesnerie (1995), and Nava et al. (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 et al. (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. We continue to use the model of the previous section rather than the more general formulation of Mirrlees (1976), which allows for preference heterogeneity. Our analysis can then be related directly to Atkinson and Stiglitz (1976, 1980) and the subsequent literature.

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 $\ell_n$, which determines income $z_n$ and disposable income $y_n = z_n - T(z_n)$, given the non-linear income tax function $T(z_n)$. Second, disposable income is allocated among the $I + 1$ commodities, $c_n$ and $x_n$. 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 $\ell_n$, the household maximizes $u(c_n, x_n, \ell_n)$ with respect to $c_n$ and $x_n^i$, $\forall i$, subject to

$$c_n + \sum_{i=1}^{I} q_i x_n^i = y_n,$$

(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

\[
\frac{u_{x^i}}{u_c} = q_i = 1 + t_i, \quad \forall n, i. \tag{25}
\]

The solution yields the conditional commodity demands \( c_n(q, y_n, \ell_n) \) and \( x_n(q, y_n, \ell_n) \), where \( q \equiv (q_1, q_2, \ldots, q_I) \) is the vector of consumer prices.

Substitution in the utility function yields a conditional indirect utility function \( v_n \):

\[
v_n \equiv v(q, y_n, \ell_n) \equiv u(c_n(\cdot), x_n(\cdot), \ell_n), \quad \forall n. \tag{26}
\]

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

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

where \( v_y(q, y_n, \ell_n) \) 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 \( \ell_n \) to find:

\[
v_{q_i, \ell}(q, y_n, \ell_n) = -v_y(q, y_n, \ell_n)\frac{\partial x_n^i(q, y_n, \ell_n)}{\partial \ell_n} - v_{y\ell}(q, y_n, \ell_n)x_n^i(q, y_n, \ell_n). \tag{28}
\]

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 \( u(c_n, x_n, \ell_n) = v_n \). The solution yields compensated demands \( c_n^*(q, \ell_n, v_n) \) and \( x_n^{*i}(q, \ell_n, v_n) \), and the expenditure function \( e(q, \ell_n, v_n) \), where by the envelope theorem

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

In the consumer’s optimum, \( y_n \equiv e(q, \ell_n, v_n) \), \( c_n^* = c_n \) 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 \( c_n^*(q, \ell_n, v_n) \) and \( x_n^{*i}(q, \ell_n, v_n) \). The type-\( n \) household’s conditional budget constraint can be expressed as

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

Differentiating (30) with respect to prices \( q_j \), labor \( \ell_n \), and utility \( v_n \) and using the envelope properties of the expenditure function (29) yields:

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

\[
\frac{\partial c_n^*}{\partial \ell_n} + \sum_{i=1, \ldots, I} (1 + t_i) \frac{\partial x_n^{*i}}{\partial \ell_n} = -\frac{u_{\ell}}{u_c}, \quad \forall n. \tag{32}
\]
\[
\frac{\partial c^*_n}{\partial v_n} + \sum_{i=1, \ldots, t} (1 + t_i) \frac{\partial x^*_{ni}}{\partial v_n} = e_v = \frac{1}{u_c}, \quad \forall n. \tag{33}
\]

Moreover, we have \( x^*_{ni}(q, \ell_n, v_n) = x^*_{ni}(q, e(\cdot), \ell_n) \), since compensated and uncompensated conditional demands are the same, and \( y_n(\cdot) = e(\cdot) \). Differentiating with respect to \( v_n \) and \( \ell_n \), and using the envelope properties (29) yields
\[
\frac{\partial x^*_{ni}}{\partial v_n} = \frac{1}{u_c} \frac{\partial x^*_{ni}}{\partial y_n}, \tag{34}
\]
\[
\frac{\partial x^*_{ni}}{\partial \ell_n} = \frac{\partial x^*_{ni}}{\partial \ell_n} + \frac{u_i}{u_c} \frac{\partial x^*_{ni}}{\partial y_n}, \quad \forall n, i. \tag{35}
\]

4.1.2 Stage 1

In this stage, a type-\( n \) household chooses labor supply \( \ell_n \) to maximize the stage-2 partial utility function \( v_n = v(q, y_n, \ell_n) \), subject to the budget constraint \( y_n = n\ell_n - T(n\ell_n) \). This yields the first-order condition for labor supply:
\[
\frac{-v_i(q, y_n, \ell_n)}{v_y(q, y_n, \ell_n)} = -\frac{u_i}{u_c} = (1 - T'(n\ell_n)) n, \quad \forall n. \tag{36}
\]

4.2 The government’s problem

The government takes as given individual behavior summarized in the conditional indirect utility functions \( v_n \equiv v(q, y_n, \ell_n) = v_n(q, y_n, z_n/n) \) for all skill-types \( n \). The incentive compatibility constraint can readily be formulated as earlier:
\[
\dot{v}_n = -\frac{\ell_n v_i(q, e(q, \ell_n, v_n), \ell_n)}{n}. \tag{37}
\]

Again, we introduce a co-state variable \( \theta_n \) associated with the incentive constraint (37). After multiplying (37) by \( \theta_n \), and integrating by parts, we can write the Lagrangian for this optimal control problem as:
\[
\mathcal{L} = \int_{\mathcal{N}} \left( \Psi(v_n) + \eta \left( n\ell_n - c_n(q, \ell_n, v_n) - \sum_i x^*_{ni}(q, \ell_n, v_n) - R \right) \right) f(n)\,dn \tag{38}
\]
\[
+ \int_{\mathcal{N}} \left( \theta_n \ell_n v_i(q, e(q, \ell_n, v_n), \ell_n) - v_n \dot{\theta}_n \right) \,dn + \theta_{\pi} v_{\pi} - \theta_2 v_w.
\]

The control variables are now \( \ell_n \) and \( q \) (or \( t \)), and the state variable is \( v_n \). The first-order conditions with respect to the control and state variables are:
\[
\frac{\partial \mathcal{L}}{\partial \ell_n} = \eta \left( n - \frac{\partial c_n}{\partial \ell_n} - \sum_i \frac{\partial x^*_{ni}}{\partial \ell_n} \right) f(n) + \frac{\theta_n}{n} (v_{\ell} + \ell_n v_{\ell} + \ell_n v_{\ell e} e_{\ell}) = 0, \quad \forall n, \tag{39}
\]
\[
\frac{\partial \mathcal{L}}{\partial q_j} = \int_{\mathcal{N}} \left( -\eta \left( \frac{\partial c_n}{\partial q_j} + \sum_i \frac{\partial x^*_{ni}}{\partial q_j} \right) f(n) + \frac{\theta_n \ell_n}{n} (v_{\ell q_j} + v_{\ell e} e_{\ell j}) \right) \,dn = 0, \quad \forall j, \tag{40}
\]
\[
\frac{\partial \mathcal{L}}{\partial v_n} = \left( \Psi' - \eta \frac{\partial c_n}{\partial v_n} - \eta \sum_i \frac{\partial x^*_{ni}}{\partial v_n} \right) f(n) + \frac{\theta_n \ell_n}{n} v_{\ell e} e_n - \dot{\theta}_n = 0, \quad \forall n \neq n, \pi, \tag{41}
\]
\[
\frac{\partial \mathcal{L}}{\partial v_{\pi}} = -\theta_{\pi} = 0, \quad \frac{\partial \mathcal{L}}{\partial v_{\ell_2}} = \theta_{\ell_2} = 0. \tag{42}
\]

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 $s^{ij}_n \equiv \partial x^*_n / \partial q^j$ the substitution effect of taxing commodity $i$ with commodity tax $t_j$. Let $E[x^j] \equiv \int_N x^j f(n) dn$ be the aggregate demand for $x^j$, and define $\varepsilon_{j\ell}$ as the following uncompensated demand elasticity for $x^j$:

$$\varepsilon_{j\ell} \equiv \frac{\partial x^j_n}{\partial \ell_n} \cdot \frac{\ell_n}{x^j_n}. \quad (43)$$

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

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

**Proof.** Use (28) and note that $v_{y\ell} = v_{\ell e}$, $x^*_j = e_{qj}$, and $v_y = u_c$. Substitute these results in (40), using (31) and (43), and impose symmetry of the substitution effects, i.e., $s^{ji}_n = s^{ji}_n$.

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 are imposed. Alternatively, the left-hand side captures the marginal deadweight loss of distorting the demand of commodity $j$ using 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 $u_c \theta_n / (\eta n)$ represents the distortion created by the income tax at skill level $n$ (see below). 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 magnitude of the right-hand side depends on the size of $\theta_n$, 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 all taxes $t^i$ on commodity demand $x^j$ is positive (negative) if the commodity-weighted cross-elasticity of labor supply with respect

---

6The elasticity $\varepsilon_{j\ell}$ is comparable to a conditional elasticity in the quantity-rationing literature. See Neary and Roberts (1980).
to commodity demand is negative (positive), i.e. average labor supply decreases (increases) with a larger level of aggregate average commodity demand.

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 \( x^j \), 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 \( x^j \):

\[
\varepsilon_{ji}^* \equiv -s_{ni}^j \frac{(1 + t_i)}{x^j_{ni}}, \quad \forall i, j. \tag{45}
\]

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

\[
\bar{\varepsilon}^*_{ji} \equiv \left[ \int x^j_{ni} dF(n) \right] \left[ \int x^i_{ni} dF(n) \right]^{-1}, \quad \forall i, j. \tag{46}
\]

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

\[
\sum_i t_i \frac{\varepsilon_{ji}^*}{1 + t_i} = \frac{t_j}{1 + t_j} \bar{\varepsilon}^*_{jj} + \sum_{i \neq j} \frac{t_i}{1 + t_i} \varepsilon_{ji}^* = -\frac{\theta_n u_c}{\eta n f(n)} \varepsilon_{jt}, \quad \forall j. \tag{47}
\]

**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 (\( \varepsilon_{ji}^* \)) and the uncompensated cross-elasticity of commodity demands with respect to labor supply (\( \varepsilon_{jt} \)). 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:

\[
\omega_c \equiv \frac{u_{c} c_n}{u}, \quad \omega_j \equiv \frac{u_{x^j} x^j_n}{u}, \quad \omega_{\ell} \equiv \frac{-u_{\ell} \ell_n}{u}, \quad \forall j. \tag{48}
\]

Recall the Hicksian elasticities of complementarity:

\[
\rho_{cc} \equiv -\frac{u_{cc}}{u_c^2}, \quad \rho_{jj} \equiv -\frac{u_{x^j x^{j^2}}}{u_{x^j}^2}, \quad \rho_{\ell\ell} \equiv -\frac{u_{\ell\ell}}{u_{\ell}^2}, \tag{49}
\]

\[
\rho_{cj} = \rho_{jc} \equiv -\frac{u_{x^j c}}{u_{x^j} u_c}, \quad \rho_{\ell j} = \rho_{j\ell} \equiv \frac{-u_{x^j \ell}}{u_{x^j} u_{\ell}}, \quad \forall j.
\]

The cross-elasticities \( \varepsilon_{jt} \) and the price elasticities \( \varepsilon_{ji}^* \) follow from the following Lemma.
Lemma 2 The cross-elasticities \( \varepsilon_{jt} \) of individual \( n \) follow from

\[
\begin{bmatrix}
\frac{\partial u}{\partial c_i} \varepsilon_{1t} \\
\frac{\partial u}{\partial c_i} \varepsilon_{2t} \\
\vdots \\
\frac{\partial u}{\partial c_i} \varepsilon_{nt}
\end{bmatrix} = S^{-1}
\begin{bmatrix}
\rho_{t1} - \rho_{ct} \\
\rho_{t2} - \rho_{ct} \\
\vdots \\
\rho_{tI} - \rho_{ct}
\end{bmatrix}, \quad \forall n
\] (50)

and the compensated price elasticities \( \omega_j \varepsilon_{jt}^* \) of individual \( n \) follow from

\[
\omega_j \varepsilon_{jt}^* = \frac{C_{ij}}{|S|}, \quad \forall i, j, n.
\] (51)

where

\[
S \equiv \begin{bmatrix}
\rho_{cc} + \rho_{j1} - 2\rho_{ct} & \rho_{cc} + \rho_{j2} - \rho_{ct} & \cdots & \rho_{cc} + \rho_{jI} - \rho_{ct} \\
\rho_{cc} + \rho_{t1} - \rho_{ct} & \rho_{cc} + \rho_{t2} - 2\rho_{ct} & \cdots & \rho_{cc} + \rho_{tI} - \rho_{ct} \\
\vdots & \vdots & \ddots & \vdots \\
\rho_{cc} + \rho_{j1} - \rho_{ct} & \rho_{cc} + \rho_{j2} - \rho_{ct} & \cdots & \rho_{cc} + \rho_{jI} - 2\rho_{ct}
\end{bmatrix}, \quad \forall n.
\] (52)

is a positive semi-definite matrix from the second-order conditions and \( C_{ij} \) is the co-factor from matrix \( S \) where column \( j \) has been replaced by the relative changes in the first-order conditions.

Proof. See Appendix. ■

From Lemma 2 follows that the cross-elasticities \( \varepsilon_{jt} \) are determined by two main factors. The first are the cross-substitution effects captured by \( S \). 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 \( \rho_{jt} - \rho_{ct} \). We again find that what is crucial for the determination of commodity taxes is whether commodities are more or less complementary to labor than the numéraire commodity is, i.e., whether \( \rho_{jt} \geq \rho_{ct} \). Lemma 2 proves that \( \varepsilon_{jt} \) can be signed if the Hicksian elasticities of complementarity \( \rho_{jt} \) and \( \rho_{ct} \) 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 \( \varepsilon_{jt} \) cannot be signed unless the entire pattern of Hicksian elasticities of complementarity \( \rho_{jt} \) for each commodity \( j \) 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, \( \rho_{jt} > \rho_{ct} \) for all commodities \( j \), and we can be sure that for all commodities \( \varepsilon_{jt} > 0 \) 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 \( \varepsilon_{jt}^* \) 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 (43)), 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 \( u(c_n, h(x_n), \ell_n) \) and \( h(.) \) is a homothetic sub-utility function over all commodities \( x_j \). This result is given in the next proposition.
Proposition 9 If utility is given by \( u(x, h(x), \ell, n) \), and \( h(\cdot) \) is a homothetic sub-utility function over all commodities \( x \), then the optimal linear commodity taxes are uniform \( (t = t_j) \) and positive (negative), i.e., \( t > 0 \) \((t < 0)\), if \( \varepsilon_{it} = \varepsilon_{jt} < 0 \) \((> 0)\), \( \forall j \). The optimal, uniform commodity taxes are given by

\[
\frac{t}{1 + t} \left( \varepsilon_{jj}^* + (I - 1) \varepsilon_{-jj}^* \right) = - \frac{\theta_n u_c}{\eta m f(n)} \varepsilon_{jt}, \quad \forall j,
\]

where \( \varepsilon_{-jj}^* = \varepsilon_{ij}^* \), for \( -j = i, \forall i \neq j \).

Proof. First, derive that \( u_{x_j} = u_h h_{x_j} \) and \( u_{x_i x_i} = u_{h} h_{x_i} h_{x_j} + u_{h} h_{x_i} h_{x_j} \), so that the elasticities of complementarity are given by

\[
\rho_{ji} = \frac{u_{x_j} u_{x_i}}{u_{x_i}^2} = \frac{(u_h h_{x_j} h_{x_i} + u_h h_{x_i} h_{x_j}) u}{(u_h)^2 h_{x_j} h_{x_i}} = \sigma_{ji} - \rho_{hh},
\]

where \( \omega_h \equiv \frac{u_h}{u} \), \( \rho_{hh} \equiv -\frac{u_{hh}}{u_h} \), and \( \sigma_{ji} \equiv \frac{h_{x_j} h_{x_i}}{h_{x_j} h_{x_i}} \) \( \forall i, j \). Second, if \( h \) is homothetic the elasticity of substitution \( \sigma_{ji} \) between \( x_j \) and \( x_i \) is constant, hence \( \sigma_{ji} = \sigma \) \( \forall i, j \). As a result \( \rho_{jj} = \rho_{jj} = \sigma - \rho_{hh} \) is constant \( \forall j \), \( i \). This implies that matrix \( S \) is given by

\[
S = (\rho_{cc} + \rho_{jj} - 2\rho_{ij}) I
\]

where \( 1 \) is an \( I \times I \) unit matrix. All the elements of \( S \) are therefore identical. From (51) then follows that all own-price and cross-price elasticities pre-multiplied with their shares are identical: \( \omega_i \varepsilon_{ij}^* = \omega_j \varepsilon_{ji}^* \), \( \forall i, j \). Third, all goods \( x \) are equally complementary to work, since with weakly separable and homothetic preferences we have \( \rho_{ij} = \frac{u_{x_i} u_{x_j}}{u_{x_i}^2} = \frac{u_{hh}}{u_{x_i} u_{x_j}} = \rho_{li} \). From (50) it then follows that \( \omega_i \varepsilon_{ij} = \omega_j \varepsilon_{ji} \), \( \forall i, j \). Homotheticity of \( h \) also implies that the shares \( \omega_j \) are independent from ability type \( n \). Hence, we can multiply both sides of (47) with \( \omega_j/\omega_i \) for all \( n \) to get:

\[
\sum_i \frac{t_i}{1 + t_i} \frac{\omega_j}{\omega_i} \varepsilon_{ij}^* = - \frac{\theta_n u_c}{\eta m f(n)} \frac{\omega_j}{\omega_i} \varepsilon_{jt}, \quad \forall j.
\]

Since the barred terms are equal for all \( j \), it must follow that all commodity tax rates \( t_i \) are identical \((t = t_i = t_j)\). Introducing the notation \( \varepsilon_{-jj}^* = \varepsilon_{ij}^* \), for \( -j = i, \forall i \neq j \), the last expression can be written as in the proposition. From the second-order conditions follows that \( \varepsilon_{jj}^* + (I - 1) \varepsilon_{-jj}^* > 0 \). Hence, if \( \varepsilon_{jt} < 0 \), we find that \( t > 0 \), and vice versa.

This case provides some additional intuition as to the desirability of commodity taxes. Indeed, in this special case commodity taxes are set at positive rates \((t > 0)\) if doing so boosts labor supply \((\varepsilon_{jt} < 0)\), and this indirectly alleviates distortions on labor supply caused by income taxation \((\text{and vice versa})\). This special case was not identified by Edwards et al. (1994) and adds to our understanding as to what determines commodity taxes. We note also that the multi-commodity case is quite special, since the presumed separability structure depends on the way we normalize the tax system.

The special case of Proposition 9 also applies to the case analyzed by Nava et al. (1996) where there is only one commodity \( x \) besides \( c \). In that case we have \( I = 1 \), and the term \((I - 1) \varepsilon_{-jj}^* \) vanishes. Indeed, like Nava et al. (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.

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  \textit{Linear commodity taxes are superfluous if utility is weakly separable between commodities and labor:}

\[ u_n \equiv u(h(c_n, x_n), \ell_n), \quad \forall n. \]  

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

Clearly, when all commodities are equally complementary to labor as the numéraire commodity, we have that \( \rho_{j\ell} = \rho_{c\ell} \) for all commodities \( j \), and the elasticities \( \varepsilon_{j\ell} = 0 \) \( \forall j \), by Lemma 2. This finding casts a different light on 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 as\textsuperscript{7}

\[ \varepsilon^*_n = \left( \frac{\ell_n v_{\ell\ell}}{v_{\ell}} + \frac{\ell_n v_{\ell\ell}}{v_{\ell}} \right)^{-1}, \quad \forall n. \]  

\text{(56)}

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

\[ \gamma_n^i = \frac{(1 + t_i)x_n^i}{(1 - T_n)n\ell_n}, \quad \forall n, \]  

\text{(57)}

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

\[ \sigma_n = \frac{1 - T'}{1 - T_n}, \quad \forall n, \]  

\text{(58)}

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

Proposition 11  \textit{Optimal non-linear income taxes at optimal commodity taxes are given by}

\[ \frac{T'(z_n)}{1 - T'(z_n)} + \frac{1}{\sigma_n} \sum_i \left( \frac{t_i}{1 + t_i} \gamma_n^i \varepsilon_{i\ell} \right) = \frac{u_{\ell\ell} \theta_n}{\eta f(n) n} \left( 1 + \frac{1}{\varepsilon^*_n} \right), \quad \forall n, \]  

\text{(59)}

\textbf{Proof.}  Substitute (32) into (39), use (36), noting that \( v_y = u_c \). \hfill \blacksquare

Equation (59) 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

\textsuperscript{7} \partial \ln u_c / \partial \ln \ell = \ell u_{\ell\ell} / u_\ell \text{ and } \partial \ln u_c / \partial \ln \ell = \ell u_{\ell c} / u_c.
if utility is weakly separable so $\varepsilon_{it} = 0$ for all $i$. As (59) 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 $t_i\gamma^i_n/(1 + t_i)$.

Intuitively, if an individual spends $\gamma^i_n$ of his net earnings on consumption of $x^i$, then the tax on $x^i$ creates an additional marginal tax burden on labor of $t_i\gamma^i_n/(1 + t_i)$. In addition, there is a correction for the rate of tax progression through $\sigma_n$. The larger is the marginal tax rate above the average tax rate, the lower is the coefficient of residual income progression $\sigma_n$. 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 $\ell$ 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 (59). The reason is that $\varepsilon_{it}$ 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 (59) 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, $\varepsilon_{it}$ 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 et al. (1996).

The term $\frac{1}{\sigma_n} \sum_i \left( \frac{t_i}{1 + t_i} \gamma^i_n \varepsilon_{it} \right) < 0$ 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 (59), the government will therefore set different marginal tax rates on labor income, cf. the left-hand side of (59), if linear commodity taxes are used.

We can more closely relate the optimal income tax in (59) 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

$$\frac{\theta_n}{\eta} = \int_n^\infty \left( \frac{\alpha}{u_c} - \frac{\Psi'}{\eta} \right) \exp \left[ - \int_n^m \left( \frac{\partial \ln u_c}{\partial \ln \ell_n} \right) ds' \right] f(m)dm, \quad \forall n, \tag{60}$$

where

$$\alpha \equiv 1 - \sum_i t_i \frac{\partial x^*_{ni}}{\partial v_n}. \tag{61}$$

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

$$\frac{\partial c^*_n}{\partial v_n} + \sum_i \frac{\partial x^*_{ni}}{\partial v_n} = \frac{1}{u_c} - \sum_i t_i \frac{\partial x^*_{ni}}{\partial y_n} \frac{1}{u_c} \equiv \frac{\alpha}{u_c}. \tag{62}$$
Using this, and noting that $e_u = 1/v_y$ and $y_n = e(\cdot)$, (41) can be rewritten as:

$$
\dot{\theta}_n = (\Psi' - \eta \alpha) f(n) + \frac{\theta_n l_n v_y}{n v_y}.
$$

(63)

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

The expression (60) for $\theta_n / \eta$ is basically identical to the standard Mirrlees (1971) expression, except for the presence of $\alpha$ 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, $\theta_n = \theta_\pi = 0$ by (42), so the total effective tax wedge in (59) becomes zero at $n = n$ and $n = \pi$ as also noted by Edwards et al. (1994) and Nava et al. (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:

$$
\frac{T'(z_n)}{1 - T'(z_n)} = -\frac{1}{\sigma_n} \sum_i \left( \frac{t_i}{1 + t_i} \gamma_i^j \xi_i \right), \quad n = n, \pi.
$$

(64)

**Proof.** Substitute (42) into (59). From (47), it follows that $\xi_i t_i$ 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 et al. (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:

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

(65)
where

\[ A_n \equiv 1 + \frac{1}{\varepsilon_n^*}, \]  
\[ B_n \equiv \frac{u_c}{1 - F(n)} \int_0^n \left( \alpha \frac{\partial \ln u_c}{\partial \ln \ell_n} \right) ds \exp \left[ - \int_n^m \left( \frac{\partial \ln u_c}{\partial \ln \ell_n} \right) ds \right] f(m) dm, \]  
\[ C_n \equiv 1 - F(n) \frac{nf(n)}{n}, \]  
\[ D_n \equiv \frac{1}{\sigma_n} \sum_i \left( \frac{t_i}{1 + t_i} \gamma_i^* \varepsilon_{it} \right). \]

**Proof.** Substitute (60) in (59) and rearrange to find the result. \(\blacksquare\)

Eq. (65) 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 \(\alpha \equiv 1 - \sum_i t_i \partial x_n^i / \partial y_n\), 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 \(\alpha\). 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 \(\Psi'/\eta\) 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 \(\varepsilon_{it}\) are of opposite sign, and provided that \(\varepsilon_{it}\) 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:

\[ \tau'(z_n) = \frac{T'(z_n)}{1 - \tau'(z_n)} + \frac{1}{\sigma_n} \sum_i \left( \frac{t_i}{1 + t_i} \gamma_i^* \varepsilon_{it} \right), \forall n, \]  

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

\[ \frac{t_j}{1 + t_j \varepsilon_{jj}^*} + \sum_{i \neq j} \frac{t_i}{1 + t_i \varepsilon_{ij}^*} = - \frac{\tau'(z_n)}{1 - \tau'(z_n)} \cdot \frac{\varepsilon_{jt}}{1 + 1/\varepsilon_n^*} \quad \forall j. \]
Proof. Using (70) in (59) gives:
\[
\frac{\tau'(z_n)}{1 - \tau'(z_n)} \left( 1 + \frac{1}{\varepsilon_n} \right)^{-1} = \frac{u_c \theta_n}{\eta f(n)n}, \quad \forall n.
\] (72)
Substitution in (47) proves the result. ■

The left-hand side of (71) again represents the marginal excess burden of distorting demand for commodity \( x^j \). The right-hand side of (71) 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 \( \varepsilon_n^* \)). 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 \( \varepsilon_{jt} \). 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 \( \varepsilon^n_{ji} \).

As earlier, the optimal structure of commodity and income taxes is not directly dependent on the particular social welfare function used, or 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 joint 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.
In future research one may investigate the implications of preference heterogeneity for the setting of optimal linear commodity taxes under non-linear income taxation. One may then conjecture that linear commodity taxes will then also be used for redistributive purposes. Another interesting route for further research is to analyze differentiated commodity taxes for setting the optimal non-linear income tax schedules using simulations.

Appendix - Proof lemma 2

Totally differentiating the first-order conditions for commodity demands $u_x(c_n, x_n, \ell_n) = q_j u_c(c_n, x_n, \ell_n) \forall j$ gives

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

We can rewrite this expression by dividing by $u_{x_j}$ and multiplying with some terms:

$$\frac{c_n u_c}{u} u_{x_j} dc_n + \sum_i \frac{x_n^i u_{x_j} u_{x_j x_i}}{u} dx_n^i + \frac{\ell_n u_{\ell_j}}{u} u_{x_j \ell} d\ell_n = \frac{c_n u_c}{u} u_{cc} dc_n + \sum_i \frac{x_n^i u_{x_j x_i}}{u} dx_n^i + \frac{\ell_n u_{\ell_j}}{u} u_{c\ell} d\ell_n + \frac{dq_j}{q_j}, \quad \forall j.$$  

Use a tilde (‘˜’) to denote a relative change:

$$\tilde{c_n} \equiv \frac{dc_n}{c_n}, \quad \tilde{x_n^i} \equiv \frac{dx_n^i}{x_n^i}, \quad \tilde{\ell_n} \equiv \frac{d\ell_n}{\ell_n}, \quad \tilde{q_j} \equiv \frac{dq_j}{q_j}, \quad \forall j.$$  

Then, using the definitions in (48) and (49), we can rewrite (74) as

$$\sum_i (\rho_{ji} - \rho_{ci}) \omega_i \tilde{x_n^i} = (\rho_{cc} - \rho_{c\ell}) \omega_c \tilde{c_n} + (\rho_{\ell j} - \rho_{c\ell}) \omega_{\ell} \tilde{\ell_n} - \tilde{q_j}, \quad \forall j.$$  

Derivation $\varepsilon_{j\ell}$

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

$$\omega_c \tilde{c_n} + \sum_i \omega_i \tilde{x_n^i} = 0.$$  

Next substitute (77) in (76) to find the following system of equations:

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

We can find the cross-elasticity $\varepsilon_{j\ell}$ by noting that $\frac{\omega_j \tilde{x_n^j}}{\omega c} |_{dx_{i\neq j}=0, dq=0} = \frac{\omega_j \omega_{\ell}}{\omega c} \varepsilon_{j\ell}$ since

$$\varepsilon_{j\ell} \equiv \frac{\partial \tilde{x_n^j}}{\partial \tilde{\ell_n}} \frac{\ell_n}{x_n^j} \left|_{dx_{i\neq j}=0, dq=0} \right. = \frac{\tilde{x_n^j}}{\ell_n}, \quad \forall j.$$  

23
The relative changes in $\omega$ determinant obtained from $d\omega$ for

\[ \sum_i (\rho_{cc} + \rho_{ji} - \rho_{ci} - \rho_{cj}) \frac{\omega_i}{\omega} \xi_{it} = \rho v_j - \rho c_{it}, \quad \forall j. \]  

which can be rewritten in matrix form as

\[ \begin{bmatrix} \frac{\omega_1}{\omega} \xi_{1t} \\ \frac{\omega_2}{\omega} \xi_{2t} \\ \vdots \\ \frac{\omega_m}{\omega} \xi_{mt} \end{bmatrix} = S^{-1} \begin{bmatrix} \rho t_1 - \rho c_t \\ \rho t_2 - \rho c_t \\ \vdots \\ \rho t_m - \rho c_t \end{bmatrix}, \]  

where

\[ S \equiv \begin{bmatrix} \rho_{cc} + \rho_{ji} - 2\rho_{ci} & \rho_{ce} + \rho_{i2} - \rho_{ie} & \cdots & \rho_{ce} + \rho_{i1} - \rho_{ie} \\ \rho_{ce} + \rho_{i1} - \rho_{ce} - \rho_{ci} & \rho_{ce} + \rho_{i2} - \rho_{ce} - \rho_{cj} & \cdots & \rho_{ce} + \rho_{i1} - \rho_{ie} - \rho_{cj} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{ce} + \rho_{j1} - \rho_{ce} - \rho_{cj} & \rho_{ce} + \rho_{j2} - \rho_{ce} - \rho_{cj} & \cdots & \rho_{ce} + \rho_{j1} - \rho_{ie} - \rho_{cj} \end{bmatrix}. \]  

**Derivation $\varepsilon_{ji}^*$**

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

\[ \omega_c x_{ni}^* + \sum_i \omega_i x_{ni}^* = 0. \]  

Substitution in (76) at $d\ell_n = 0$ yields

\[ \sum_i (\rho_{cc} + \rho_{ji} - \rho_{ci} - \rho_{cj}) \omega_i x_{ni}^* = -\tilde{q}_j, \quad \forall j. \]  

This last equations can be written in matrix form as

\[ \begin{bmatrix} \omega_1 x_{ni}^* \\ \omega_2 x_{ni}^* \\ \vdots \\ \omega_m x_{ni}^* \end{bmatrix} = -S^{-1} \begin{bmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_m \end{bmatrix}. \]  

Applying Cramer's rule to (85) yields the elasticities. In particular, let $D_j$ denote the determinant obtained from $|S|$ by replacing the $j$-th column with vector $[\tilde{q}_1, \tilde{q}_2, \cdots, \tilde{q}_m]^T$:

\[ D_j \equiv \begin{vmatrix} \rho_{cc} + \rho_{ji} - 2\rho_{ci} & \cdots & \rho_{ce} + \rho_{j1} - \rho_{cj} - \rho_{cj1} & \cdots & \rho_{ce} + \rho_{j1} - \rho_{cj} - \rho_{cj1} \\ \rho_{ce} + \rho_{i2} - \rho_{ci} - \rho_{cj} & \cdots & \rho_{ce} + \rho_{j2} - \rho_{cj} - \rho_{cj2} & \cdots & \rho_{ce} + \rho_{j2} - \rho_{cj} - \rho_{cj2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{ce} + \rho_{i1} - \rho_{ci} - \rho_{cj} & \cdots & \rho_{ce} + \rho_{j1} - \rho_{cj} - \rho_{cj1} & \cdots & \rho_{ce} + \rho_{j1} - \rho_{cj} - \rho_{cj1} \end{vmatrix}. \]  

The relative changes in $\omega_j x_{ni}^*$ are then given by

\[ \omega_j x_{ni}^* = -\frac{D_j}{|S|}, \quad \forall j. \]  

Let $C_{nm}$ denote the co-factors of row $n$ and column $m$ in matrix $D_j$, then we can write for $\omega_j x_{ni}^*$

\[ \omega_j x_{ni}^* = -\left( \frac{C_{1j}}{|S|} \tilde{q}_1 + \frac{C_{2j}}{|S|} \tilde{q}_2 + \cdots + \frac{C_{1j}}{|S|} \tilde{q}_m \right), \quad \forall j. \]
Next, note that the elasticities can be written as:

\[
- \omega_j \tilde{x}_j = \lim_{dx_j \neq 0 \to 0} = \omega_j \varepsilon_{ji}, \quad \forall i, j, \tag{89}
\]

since

\[
\varepsilon_{ij} = -\frac{\partial x^{*i}}{\partial q_j} \frac{q_j}{x^{*i}} = -\left. \frac{dx^{*i}/x^{*i}}{dq_j/q_j} \right|_{dx_j \neq 0 \to 0} = -\tilde{x}_{ni} \frac{q_j}{q_j}, \quad \forall i, j. \tag{90}
\]

Therefore, we find that the elasticities are given by:

\[
\omega_j \varepsilon_{ji} = C_{ij} \left| \frac{S}{S} \right|, \quad \forall i, j. \tag{91}
\]

References


