On optimal commodity taxes when consumption is time consuming

Firouz Gahvari*

Department of Economics
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

July 2004
Revised, January 2006

*I thank an anonymous referee, Helmuth Cremer and the Editor, John Conley, for helpful comments.

Abstract

This paper studies the problem of optimal taxation of commodities when consumption is a time consuming activity. This is done under two distinct preference separability assumptions: between goods and labor supply, and between goods and leisure. It argues that with labor separability, the traditional uniform taxation results of optimal tax theory continue to hold. With leisure separability, on the other hand, consumption time is a major ingredient of optimal tax rates. However, the relationship between consumption time and optimal tax rates depends crucially on the representation of the economy. In representative consumer economies, time differences determine the pattern of optimal tax rates so that goods whose consumption take more time are subjected to higher tax rates. When individuals have different earning abilities, redistributive, incentive and efficiency considerations also come into play resulting in a complex relationship. The paper derives formulas for optimal commodity taxes in this case on the basis of three different tax structures: linear commodity taxes in combination with linear and nonlinear income taxes, and nonlinear commodity taxes in combination with nonlinear income taxes.

JEL classification: H21: D13; J22.

Keywords: Household production; optimal taxation; separability in leisure; separability in labor; heterogeneous agents.

1 Introduction

Labor supply plays a critical role in the theory of optimal commodity taxation. It is thus rather surprising that the optimal tax literature has paid little attention to the model of labor supply that underlies it. An obvious alternative to the standard model of labor supply is Becker's (1965) household production function. This is particularly relevant from the perspective of optimal tax theory in that Becker's accounting of consumption time, breaks the artificial dichotomous division of time into "labor supply" and "leisure". This approach teaches us that leisure is *not identically equal to* time away from work.

Given that the time taken in consumption varies across goods, one might reasonably expect that with Becker's formulation, the nature of second-best taxes would be different from the traditional literature. Gahvari and Yang (1993) were the first to study this question. Specifically, they examined the nature of optimal commodity taxes (on market-purchased goods), using Becker's formulation, for three special cases studied in the traditional literature: a perfectly inelastic labor supply; directly additive preferences; and the inverse elasticity rule. More recently, Kleven (2004) uses Gahvari and Yang's (1993) framework[1] to study the implication of "weak separability" between leisure and goods —yet another special case of interest studied in the traditional literature.[2]

Weak separability, however, can take two distinctly different meanings, and formulations, when one moves away from the standard model of labor supply and into the Becker's approach. Whereas in the traditional model, the separability between labor supply and goods is the same as separability between leisure and goods, this is no longer the case with Becker's formulation. The two assumptions yield radically different results for what the optimal tax rates should be. The aim of the current paper is to examine the

[1] Although Kleven (2004) asserts that his paper is close to Gahvari and Yang (1993) only in "spirit," he uses an identical setup in which time and market-purchased goods are used in fixed-proportions.

[2] By goods I mean non-leisure goods. This terminology is adopted for brevity and used throughout the paper.

nature of these differences. It studies the implications of the two separability assumptions in models with homogeneous *and* heterogeneous agents. Gahvari and Yang (1993) and Kleven (2004) have been concerned solely with one-consumer economies. Moreover, in studying the heterogeneous agents case, I differentiate between models when the government has access only to a linear income tax (the "Ramsey approach" to optimal taxation) and when it can also levy a nonlinear income tax (the "Mirrlees approach" to optimal taxation.)

In the traditional optimal tax theory, the role of the separability assumption is to block differential commodity taxes from having any efficiency properties, and, if individuals have different earning abilities, redistributive concerns as well. When consumption is time consuming, this blockage is no longer as absolute. Moreover, its degree of effectiveness depends on whether the agents are homogeneous or heterogeneous. When individuals are all alike, the only effect of differential commodity taxation that separability cannot isolate (when consumption is time consuming), is the differential time spent on consumption of different goods. On the other hand, when individuals are heterogeneous, an additional factor comes into play. This is due to the fact that time spent by different individuals in consumption (even for the same good) have differing opportunity costs. This additional factor results in a different "rule" for optimal taxation in many-consumer economies despite the separability assumption.

The traditional literature contains three types of results on ineffectual differential commodity taxation (commonly called tax uniformity results). The first is due to Sandmo (1974). He proved that if preferences are separable in leisure (or labor supply) and goods, with the subutility for goods being homothetic, optimal commodity taxes are proportionately uniform. Sandmo's result is derived within the context of the traditional one-consumer Ramsey problem and embodies only efficiency considerations. Kleven (2004) studies how Becker's approach, as modeled by Gahvari and Yang (1993), changes Sandmo's one-consumer-economy-tax-uniformity result. He shows that separability between leisure and goods implies that optimal commodity tax on every good is inversely related to the ratio of its market price to its "full cost" (the *market price* plus the opportunity cost of time used in consumption). He calls this result "the inverse factor share rule." This is in fact another facet of Gahvari and Yang's(1993) original result (derived under the assumption that labor, but not leisure, is perfectly inelastic in supply). As they noted "Market goods that are more time intensive in consumption ... should be taxed at higher rates than market goods that are less time intensive in consumption" (p. 484).

The case presented by Kleven, however, is not the only counterpart to Sandmo's traditional result under separability when consumption is time consuming. Sandmo's result can be stated *equivalently* in terms of separability between leisure and goods, *and* between labor supply and goods. This equivalence is an integral feature of the traditional model of labor supply. With Becker's specification, on the other hand, separability between labor supply and goods is no longer identical to separability between leisure and goods. Interestingly, Becker's formulation will not change the uniformity result of the traditional literature, if the presumed separability is between labor supply and goods.

The second type of uniformity results has its origin in Atkinson and Stiglitz (1976). They showed that, for a particular example of preferences, a *linear* income tax calls for proportionately uniform commodity taxes. Deaton (1979) generalized this result and proved that it holds for all preferences that are weakly separable between labor supply and goods, provided that all consumers have linear Engel curves for goods in terms of income.[3] Within this framework, i.e. assuming individuals differ in earning

[3] Atkinson and Stiglitz (1976) had claimed that their result holds as long as preferences over labor supply and goods correspond to the linear expenditure system. This was later proved by Atkinson (1977). Deaton (1981) re-examined the uniformity issue on the basis of quasi or implicit separability. Deaton and Stern (1986) generalized Deaton's (1979) result, who had assumed all individuals have identical tastes, by allowing for a bit of taste differentiation. Engel curves can now have different intercepts while the government is enabled to make differential lump-sum grants conditioned on observable household characteristics. Finally, Besley and Jewitt (1995) have discussed the uniformity issue with particular

ability, and ruling out a general income tax, I show that the traditional optimal tax results remain unchanged if preferences are separable between labor supply and goods and the subutility of goods is homothetic.[4] Next, I assume that the separability is between leisure (as opposed to labor supply) and goods and derive a formula for optimal commodity taxes when consumption is time consuming. I show that, ceteris paribus, goods whose consumption take more time should be taxed at a higher rate. Similarly, goods whose consumption take the same time, if there are any, should be subjected to the same tax rate. Nevertheless, unlike the one-consumer case studied by Kleven (2004), the optimal tax rates cannot be characterized by a simple inverse relationship between tax rates on goods and the ratios of goods' market prices to their full costs. Distributive considerations enter into the picture as well.

That distributive concerns matter is quite intuitive. The additional efficiency considerations arise because when the opportunity cost of time differs across individuals, the "full opportunity cost of consumption" (opportunity cost of time taken for consumption plus the *producer* price of the good) will be different for different individuals. This implies that the compensated price elasticity of leisure with respect to goods will also differ across individual types (despite the separability and homotheticity assumptions). Consequently, the excess burden of commodity taxes will depend on *who* consumes any given taxed good. Such considerations are plainly absent in one-consumer economies. They are also absent in traditional optimal tax models where consumption is not time consuming (and the full opportunity cost of consumption is simply the producer price of the good in question).

The third and most influential uniformity result is due to Atkinson and Stiglitz (1976). This classic paper on the design of tax structures was particularly concerned with the usefulness of commodity taxes in the presence of a general income tax in many-consumer

emphasis on the circumstances under which first-order conditions are sufficient for the problem.

[4] This assumption implies that Engel curves are not only linear but that they go through the origin.

economies.[5] It proved that if the government can levy a *general* income tax, and if preferences are weakly separable in labor supply (or leisure) and goods (homotheticity is no longer required), then commodity taxes are not needed as instruments of optimal tax policy.[6] Using Stiglitz's (1987) reformulation of the Mirrlees (1971) optimal tax problem, and Becker's household production approach, I derive a general formula for optimal commodity taxes in an economy with many types of agents and many goods. The formula is quite general. In particular, it is not based on the "single-crossing" property and I make no assumptions regarding which self-selection constraints bind. I show that if separability is between labor supply and goods, Atkinson and Stiglitz's uniformity result continues to hold. On the other hand, assuming preferences are separable in leisure and goods, I show that time taken in consumption does matter for optimal taxation. In particular, I prove that any two goods whose consumption take the same time, should have identical tax rates. Nevertheless, the optimal tax formula is more complex than what the inverse factor share rule of one-consumer economies implies. Redistributive *and* incentive terms also enter into the optimal tax formula. Additionally, the relative commodity tax rates faced by different individuals differ across types, with the highest ability type facing no differential taxes.[7]

Finally, I apply Becker's household approach to a version of the Mirrlees (1971) op-

[5]The ineffectiveness of commodity taxes and their proportionately uniform tax treatment boil down to the same thing. In the absence of exogenous incomes, the government will have an extra degree of freedom in setting its income and commodity tax instruments. This reflects the fact that demand function will then be homogeneous of degree zero in consumer prices and lump-sum grants, if any. In consequence, the government can, without any loss of generality, set one of the commodity taxes at zero (i.e. set one of the commodity prices at one). Under this normalization, uniform rates imply absence of commodity taxes.

[6]Atkinson and Stiglitz assumed that individuals have identical preferences and differ only in their earning abilities. Mirrlees (1976) generalized the result by examining the type of preferences which make commodity taxes redundant while allowing for taste differentiation.

[7]Thus Kleven (2004, p. 554) is incorrect when he claims that with heterogeneous agents, "we may apply the Atkinson and Stiglitz (1976) proposition stating that the optimum involves uniform taxation of consumption goods, provided that leisure is weakly separable in utility. In the Becker context, however, this result translates into a uniform tax on household activities which gives us the inverse factor share rule."

timal tax problem in which commodity taxes are constrained to be linear. This is the tax structure one can justify best on informational grounds. Whereas Mirrlees's original framework requires the tax authorities to observe individual consumption levels, this reformulation requires information only on anonymous transactions. I derive the formulas for the optimal commodity taxes in this case and show that, as previously, time taken in consumption, incentive and redistributive motives all play a role in determining the optimal tax rates. If consumption time is the same for all goods, optimal commodity taxes are uniform. However, when consumption times differ across goods, higher consumption time does not necessarily imply a higher tax rate. The inverse factor share rule again fails to hold.

2 The one-consumer economy

Consider an economy in which the representative consumer derives utility from leisure, l, and n produced goods $\underline{x} \equiv (x_1, x_2, \ldots, x_n)$. The individual is endowed with one unit of time. The economy's production technology is linear with labor, L, the only factor used in production of $\underline{x}$. Relative producer prices of $\underline{x}$ and L are thus constant. By choice of units, prices of x_i's are all set equal to one. The wage is denoted by w. Consumption of one unit of x_i takes a_i units of time. The individual then faces a time constraint of the form,

$$L + l + \sum_{i=1}^{n} a_i x_i = 1, \tag{1}$$

in addition to the budget constraint

$$\sum_{i=1}^{n} q_i x_i = wL + G, \tag{2}$$

where $q_i = 1 + t_i$ is the consumer price of x_i, t_i is the unit excise tax on x_i and G is lump-sum income (if any).[8]

[8]Observe that this formulation is formally identical to Gahvari and Yang's (1993), as well as Kleven's (2004). They stipulate that the consumer derives utility from n "consumption activities":

In describing the representative individual's behavior, one may proceed in two ways. First, represent the individual's preferences by a utility function written in terms of labor supply and other goods

$$U = U(L, \underline{x}), \tag{3}$$

where $U(\cdot)$ is strictly monotonic, quasi-concave and twice differentiable. Alternatively, one can represent the preferences by a utility function written in terms of leisure and other goods,

$$\Omega = \Omega(l, \underline{x}), \tag{4}$$

with $\Omega(\cdot)$ being strictly monotonic, quasi-concave and twice differentiable. Of course, the two formulations are equivalent and related to one another according to[9]

$$\Omega(l, \underline{x}) = \Omega(1 - L - \sum_{i=1}^{n} a_i x_i, \underline{x}) \equiv U(L, \underline{x}).$$

In the traditional model of labor supply, it makes no difference if one assumes that preferences are separable in leisure and produced goods, or in labor supply and goods. The two assumptions are identical. This is not the case with Becker's formulation of labor supply. Specifically, if the assumption is made in terms of labor supply then we will have the traditional uniform taxation result (assuming that homotheticity assumption is also satisfied). On the other hand, if the separability assumption is made with respect to leisure, tax rates will be directly related to the ratio of their "full cost" to their market price. The following two subsections prove these assertions. It must also be clear that

---

$z_1, z_2, \ldots, z_n$, and leisure, where z_j ($j = 1, 2, \ldots, n$) is "produced" using a market-purchased good, x_j, and a time input, t_j, according to the Leontieff production technology

$$z_j = \min\left(\frac{x_j}{b_j}, \frac{t_j}{a_j}\right), \qquad j = 1, 2, \ldots, n.$$

If b_i is normalized at one (for all $i = 1, 2, \ldots, n$), the two formulations coincide.

[9] Or

$$U(L, \underline{x}) = U(1 - l - \sum_{i=1}^{n} a_i x_i, \underline{x}) \equiv \Omega(l, \underline{x}).$$

which assumption one should adopt is not a question that can be answered on a priori reasoning.

2.1 Labor separability

Consider the behavior of an individual who maximizes (3) subject to (2). This yields the first-order conditions

$$-\frac{U_L}{U_i} = \frac{w}{q_i}, \qquad i = 1, 2, \ldots, n.$$

These conditions plus equation (2) lead to the labor supply function $L = \hat{L}(\underline{q}, w, G)$, demand functions $x_i = \hat{x}_i(\underline{q}, w, G)$, $i = 1, 2, \ldots, n$, and the indirect utility function $v = \hat{v}(\underline{q}, w, G)$. Assume now that $U(L, \underline{x})$ is weakly separable in L and $\underline{x}$ and that the subutility in $\underline{x}$ is homothetic. One can then easily establish, following Sandmo (1974), that optimal tax rates on x_i's are uniform.

2.2 Leisure separability

Now consider the behavior of an individual who maximizes (4) subject to (2). Substitute for L from (1) into (2) to rewrite it as

$$\sum_{i=1}^{n} (q_i + a_i w) x_i = w(1 - l) + G. \tag{5}$$

Then choose l and $\underline{x}$ to maximize $\Omega(\cdot)$ subject to (5). This yields the first-order conditions

$$\frac{\Omega_l}{\Omega_i} = \frac{w}{q_i + a_i w}, \qquad i = 1, 2, \ldots, n.$$

Denote $\tilde{q}_i \equiv q_i + a_i w$ $(i = 1, 2, \ldots, n)$ and $\underline{\tilde{q}} = (\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_n)$. The above conditions plus equation (5) then lead to the demand functions $l = l(\underline{\tilde{q}}, w, G)$, $x_i = x_i(\underline{\tilde{q}}, w, G)$, and the indirect utility function $v = v(\underline{\tilde{q}}, w, G)$.

Next assume that $\Omega(\cdot)$ is separable in l and $\underline{x}$, with the latter the subutility being homothetic. Writing the Ramsey tax problem for this case, using a result from

Sandmo (1974), I prove in the Appendix that

$$\frac{t_i}{q_i} = \frac{\mu - \gamma}{\mu\,\epsilon}\,\frac{q_i + a_i w}{q_i}, \qquad i = 1, 2, \ldots, n, \tag{6}$$

where μ is the shadow cost of public funds (the Lagrange multiplier associated with the government's revenue constraint), $\gamma \equiv \alpha + \sum_i t_i(\partial x_i/\partial G)$ is the "social marginal utility of income" (α is the marginal utility of income), and ϵ is the cross-price elasticity of compensated demand for l with respect to any one of the goods (same for all goods when preferences are separable in leisure and goods). Equation (6) shows that there is a direct relationship between the optimal tax rates and the ratios of the goods' "full cost" to their market price. This can also be stated in terms of an inverse relationship between tax rates and the ratios of the goods' market price to their full cost—what Kleven calls the "inverse factor rule."[10]

One may also rewrite (6) in another informative way:

$$\frac{t_s}{t_i} = \frac{1 + a_s w}{1 + a_i w}, \qquad i, s = 1, 2, \ldots, n. \tag{7}$$

That is, the tax rates are such that they keep the relative "full opportunity costs" of goods (their *producer* price plus time cost of consumption) unchanged. This is of course a variant of the uniform taxation principle.

These results are summarized as:

Proposition 1 *Assume consumption of one unit of* x_i *takes* a_i *units of time* ($i = 1, 2, \ldots, n$)*, so that, in a representative consumer economy, each person faces a time constraint given by equation (1).*

(i) If preferences are weakly separable in L *and* $\underline{x}$ *with the subutility in* $\underline{x}$ *being homothetic, optimal commodity taxes are uniform.*

[10]As observed earlier, this is very much related to Gahvari and Yang's (1993) original result (in spirit?) that relatively more time intensive goods should be taxed at higher rates. Observe also that Kleven (2004) proves this result by minimizing the excess burden of commodity taxes. However, I prove this by the standard method of maximizing the indirect utility function with respect to the tax rates.

(ii) If preferences are weakly separable in l and $\underline{x}$ with the subutility in $\underline{x}$ being homothetic, optimal commodity taxes are characterized by equation (6).

3 The many-consumer economy à la Ramsey

Consider the same economy as in the previous section, but now assume that it is inhabited by individuals who differ in earning ability (they continue to have identical taste). Specifically, assume there are H different earning abilities, with a person of type h earning w^h $(h = 1, 2, \ldots, H)$. Continue to normalize the population size at one, and denote the proportion of persons of type h in the population by π^h.

Observe that the optimization problem of individuals in this setting remains the same as that in the representative consumer case. Denote a h-type person's indirect utility function by $v^h(\cdot)$ and let

$$\sum_{h=1}^{H} \pi^h W\left(v^h(\cdot)\right), \tag{8}$$

represent the social welfare function for the economy (where $W(.)$ is concave and twice differentiable). Let G denote the lump-sum rebate that the government pays to each person regardless of his type. The optimal tax problem is to determine the values of G and the commodity tax vector $(t_1, t_2, \ldots, t_n)$ which maximize social welfare subject to the government's external revenue requirement, $\bar{R}$,

$$\sum_{i=1}^{n} t_i \left(\sum_{h=1}^{H} \pi^h x_i^h\right) - G \geq \bar{R}. \tag{9}$$

3.1 Labor separability

Assume that the direct utility function behind $v^h(.)$ is $U(L, \underline{x})$ so that $v^h = v(\underline{q}, w^h, G)$. If $U(L, \underline{x})$ is weakly separable in L and $\underline{x}$, one can reformulate the individual's problem using a two-stage optimization procedure (optimizing first over consumption goods conditional on labor supply); see Deaton and Muellbauer (1980). The "conditional" indirect utility function will then take the form of $v^h = v(\underline{q}, w^h L^h + G)$.[11] The structure of the optimal tax problem will then be identical to the standard model. Thus, if the subutility in $\underline{x}$ is homothetic, the traditional uniform tax result, as originally proved by Deaton (1979), holds.[12]

3.2 Leisure separability

Assume now that the direct utility function behind $v^h(.)$ is $\Omega(l, \underline{x})$ so that $v^h = v(\underline{\tilde{q}}, w^h, G)$. If $\Omega(l, \underline{x})$ is weakly separable in l and $\underline{x}$, again using a two-stage optimization procedure on the part of the individual, the conditional indirect utility function will take the form of $v^h = v\left(\underline{q}, w^h(1 - l^h) + G\right)$. I show in the Appendix that with these preferences, assuming that the subutility in $\underline{x}$ is homothetic, the optimal commodity taxes are characterized by

$$\frac{t_s}{q_s} = \sum_h \pi^h \frac{\overline{\gamma} - \gamma^h}{\overline{\gamma}\, \epsilon^h} \frac{q_s + a_s w^h}{q_s}, \tag{10}$$

where

$$\beta^h \equiv W'(v^h) \frac{\partial v^h}{\partial G},$$

is the h-type individual's social marginal utility of income,

$$\gamma^h \equiv \beta^h + \mu \sum_i t_i \frac{\partial x_i^h}{\partial G},$$

is the h-type individual's *net* social marginal utility of income, μ is the shadow price of government revenue (the Lagrangian multiplier on the revenue constraint), $\overline{\gamma} = \sum_h \pi^h \gamma^h$ is the mean of γ^h's, and ϵ^h is the cross-price elasticity of the h-type's demand for leisure with respect to any one of the produced goods. Observe that the separability and homotheticity assumptions imply that ϵ^h is the same for all goods.

Equation (10) is the counterpart of (6) that characterized optimal tax rates in a one-consumer economy. If $a_s = 0$ for all $s = 1, 2, \ldots, n$, $t_s/q_s = \sum_h \pi^h (\overline{\gamma} - \gamma^h)/\overline{\gamma}\, \epsilon^h$, and one

[11] This entails an abuse of notation in that $v(\underline{q}, w^h L^h + G)$ has a different functional form from $v(\underline{q}, w^h, G)$.

[12] This will be the case as long as one has interior solutions for l and $\underline{x}$.

has the traditional uniform tax result. When $a_s \neq 0$, we have some version of the inverse factor share rule; albeit *not* the simple version of the one-consumer economy case. It is clear from (10) that all goods whose consumption take the same time should have the same tax rate. It is also clear that when a_s differs across goods, t_s moves positively with a_s; that is, the more time intensive goods will have a higher tax rate. Nevertheless, the optimal tax rates are not governed only by the ratio of their "full cost" to their market price. Distributive considerations enter into the picture through γ^h. Moreover, efficiency terms also come into play as reflected by ϵ^h—the h-type's compensated price elasticity of leisure with respect to goods. Intuitively, because the opportunity cost of time differs across individuals, the full opportunity cost of consumption will be different for different individuals. One would then expect that ϵ^h's also to differ across individuals, despite the separability and homotheticity assumptions.

The impact of redistributive and efficiency considerations on the relative tax rates is seen most clearly from the following expression derived directly from (10) by substituting $1 + t_s$ for q_s.

$$\frac{t_s}{t_i} = \frac{1 + \Theta\, a_s}{1 + \Theta\, a_i}, \qquad i, s = 1, 2, \ldots, n, \tag{11a}$$

where

$$\Theta \equiv \frac{\sum_h \pi^h(\overline{\gamma} - \gamma^h) w^h/\epsilon^h}{\sum_h \pi^h(\overline{\gamma} - \gamma^h)/\epsilon^h}. \tag{11b}$$

Equation (11a) corresponds to equation (7), its counterpart for the one-consumer case. Observe that any different configuration of γ^h's *or* ϵ^h's imply a different value for Θ and, with it, a different ratio of tax rates for the *same* a_s and a_i (as long as $a_s \neq a_i$).[13]

The discussion in this subsection is summarized as

Proposition 2 *Assume the economy is inhabited by individuals who differ in earning ability but have identical preferences. It takes each person a_i units of time to consume*

[13] Time differences in consumption become irrelevant if the social planner has no equity objectives. Setting $\gamma^h = \overline{\gamma}$ (for all $h = 1, 2, \ldots, H$) in (10) implies that $t_s = 0$ for all $s = 1, 2, \ldots, n$. All revenues are then raised from a head tax.

one unit of good i ($i = 1, 2, \ldots, n$). Tax instruments are constrained to be linear.

(i) If preferences are weakly separable in labor supply and goods, with the goods subutility being homothetic, optimal commodity taxes are uniform.

(ii) If preferences are weakly separable in leisure and goods, with goods subutility in $\underline{x}$ being homothetic, the optimal commodity taxes are characterized by equation (10).

4 The many-consumer economy à la Mirrlees

The presence of a general income tax changes the landscape for optimal taxation. In the traditional model, weak separability between goods and leisure (or labor supply) becomes sufficient for the redundancy of commodity taxes; homotheticity of the goods subutility is no longer required. To characterize the optimal tax rates, I consider the standard equivalent problem of the government first choosing optimal allocations subject to resource balance and self-selection constraints. Having derived the optimal allocation, I then describe the tax structure that can implement it.[14]

Let w^k denote the wage of an individual of "type" k, with $w^k > w^h$ whenever $k > h$. Introduce a type-specific utility function describing preferences over x_i's, and $I = wL$,

$$\boldsymbol{u^h}(I, \underline{x}) \equiv \Omega \left(1 - \frac{I}{w^h} - \sum_i a_i x_i, \underline{x} \right) \equiv U \left(\frac{I}{w^h}, \underline{x} \right). \tag{12}$$

Denote the utility level of a h-type individual by u^h when he chooses the allocation intended for him, and by u^{hk} when he chooses a k-type person's bundle; namely,

$$u^h = \boldsymbol{u^h}(I^h, \underline{x}^h), \tag{13a}$$

$$u^{hk} = \boldsymbol{u^h}(I^k, \underline{x}^k). \tag{13b}$$

[14] Observe that as long as a_i's do not differ across individuals, this problem remains very much within Mirrlees's original formulation with a single source of heterogeneity (ability). On optimal tax models with multi-dimensional heterogeneity see, among others, Cremer and Gahvari (1998, 2002); Cremer, Gahvari and Ladoux (1998, 2001); and Cremer, Pestieau and Rochet (2001).

One can describe the set of Pareto-efficient allocations as follows. Let δ^h $(h = 1, 2, \ldots, H)$ denote a positive constant with the normalization $\sum_{h=1}^{H} \delta^h = 1$. Maximize

$$\sum_{h=1}^{H} \delta^h u^h, \tag{14}$$

with respect to $\underline{x}^h, y^h$ and I^h; subject to the resource constraint

$$\sum_{h=1}^{H} \pi^h \big(I^h - \sum_i x_i^h - y^h \big) \geq \bar{R}, \tag{15}$$

and the self-selection constraints

$$u^h \geq u^{hk}, \quad h \neq k;\ h, k = 1, 2, \ldots, H. \tag{16}$$

4.1 Labor separability

Ignore $\Omega(.)$ in (12). The problem (12)–(16) has precisely the same structure as the traditional optimal tax problem with different individual types. One can then derive the first-order conditions for the government's optimization problem. These conditions, in conjunction with the economy's resource constraint, characterize the economy's Pareto-efficient allocations constrained by self-selection. Using the standard arguments, one can show that if $U(L, \underline{x}) = U(L, \phi(\underline{x}))$, commodity taxes are not needed for the implementation of these allocations.[15]

[15] A referee asks "where the time constraint enters the analysis here". The answer is that it depends on whether preferences are originally defined over labor supply and goods or leisure and goods. In the former case, the time constraint comes into the picture only with respect to the determination of leisure (assuming an interior solution). This parallels the traditional model with $L + l = T$ where, with such a preference specification, T affects l but plays no role in the choice of L and $\underline{x}$. In the latter case, the time constraint is embedded in $\phi(\underline{x})$. To see this most clearly, consider preferences that are quasi-linear in l so that $\Omega = l + b_1 \ln x_1 + b_2 \ln x_2$. Using the time constraint, one can rewrite this as $\Omega = -L + 1 + (b_1 \ln x_1 - a_1 x_1) + (b_2 \ln x_2 - a_2 x_2) = U(L, \phi(\underline{x}))$.

Observe also that the uniform commodity tax result holds as long as one has a interior solution for l and $\underline{x}$. In the traditional optimal tax problem, this implies that at the equilibrium $0 < L^h < 1$, or alternatively $0 < I^h < w^h$. In the present setting, with the time constraint being different, the implication is that at the equilibrium, $0 < I^h < w^h \left(1 - \sum_i a_i x_i\right)$.

4.2 Leisure separability

Now ignore $U(.)$ in (12) and assume that preferences are weakly separable in leisure and goods: $\Omega(l, \underline{x}) = \Omega\left(1 - I/w^h - \sum_i a_i x_i, f(\underline{x})\right)$. In the Appendix, I derive the first-order conditions for the government's optimization problem in this case. Again, these equations, in conjunction with the economy's resource constraint, characterize the economy's constrained Pareto-efficient allocations. The characterizations that I derive are quite general. In particular, they are not based on the "single-crossing" property. The self-selection constraints associated with a h-type mimicking a k-type and a k-type mimicking a h-type can simultaneously bind. It is also possible to have "bunching". On the basis of the characterizations, denoting the (non-negative) Lagrangian multipliers associated with the resource constraint (15) by μ, and with the self-selection constraints (16) by λ^{hk}, I am able to show that the relative optimal tax rates on goods s and i facing a h-type individual (t_s^h, t_i^h) has the following specification $(s, i = 1, 2, \ldots, n; h = 1, 2, \ldots, H)$:

$$\frac{1 + t_s^h}{1 + t_i^h} = 1 + \frac{a_s \Omega_i^h - a_i \Omega_s^h}{-a_i \Omega_l^h + \Omega_i^h} \sum_{k \neq h} \frac{\lambda^{kh} \Omega_s^{kh}}{\mu \pi^h} \left(\frac{\Omega_l^h}{\Omega_s^h} - \frac{\Omega_l^{kh}}{\Omega_s^{kh}} \right), \tag{17}$$

with the notation

$$\Omega_i^h = \Omega_i \left(1 - \frac{I^h}{w^h} - \sum_i a_i x_i^h,\ f(x_1^h, x_2^h, \ldots, x_n^h) \right), \tag{18a}$$

$$\Omega_i^{kh} = \Omega_i \left(1 - \frac{I^h}{w^k} - \sum_i a_i x_i^h,\ f(x_1^h, x_2^h, \ldots, x_n^h) \right), \tag{18b}$$

where $\Omega_i = (\partial \Omega / \partial x_i^h)|_l$, with similar notation for $\Omega_s^h, \Omega_s^{kh}, \Omega_l^h$ and Ω_l^{kh}.[16]

[16] A referee is concerned about implementability. He writes "that the commodity tax rates differ between types ... immediately raises questions about implementability. Income is taxed because wage rates are not observable. How then can commodity taxes be conditioned on wage rates?" The answer is that having commodity taxes differ across types does not mean that their implementation is conditioned on wages (if by "conditioned" one means that the government should know a person's wage in order to levy a particular tax on him). Nor does the appearance of wage rates in the optimal tax formulas for different types [as in (17)] mean that implementation is conditioned on wages. The problem here is akin to Mirrlees's original optimal tax problem in which income tax rates depend on wages, but their implementation does not. The point is that the government knows the distribution of wages

It is easy to show that if $a_s = a_i$, goods s and i should have the same tax rate (see the Appendix). This holds for any two goods regardless of how long it takes to consume *other* goods. However, this characterization is quite a bit more complicated than the simple inverse factor share rule.[17] Observe also that the effects of a_s and a_i *differ* for individuals of different types depending on their wage rates (and, of course, the pattern of binding incentive compatibility constraints). Also, as is typical in these models, if there is a group of individuals whom no one wants to mimic in equilibrium (i.e. if there exists a group H for whom $\lambda^{kH} = 0$ for all $k = 1, 2, \ldots, H-1$), they should face no differential commodity tax rates even when the goods they consume entail different time consumption (a_i's are different for different goods).[18] If redistribution is from higher to lower ability persons, these individuals will be those with the highest w.

The following proposition summarizes the results of this subsection.

Proposition 3 *Assume the economy is inhabited by individuals who differ in earning ability but have identical preferences. It takes each person a_i units of time to consume one unit of good i ($i = 1, 2, \ldots, n$). Assume further that personal purchases are publicly observable so that nonlinear commodity taxes are feasible.*

---

and thus can calculate the optimal tax rates according to (17). What the government does not know is who earns what wages. But that information is not needed for implementation. The solution to the optimal tax problem posed here is incentive compatible and must thus be implementable by a tax schedule conditioned on observables (particularly given the discrete type setup of the problem). The implementation requires that the government imposes a general transfer function on I and $\underline{x}$, the publicly observable variables here. This implies that a consumer's marginal price for x_i depends on his income as well as his consumption of $\underline{x}$. (It also implies that his marginal income tax rate will depend on his consumption of $\underline{x}$ as well as his income.) See, Cremer and Gahvari (2002) for a discussion of implementation through separate implementing functions: a pricing function that depends on $\underline{x}$ and a tax function that depends on I.

[17] This refutes Kleven's (2004) claim that as long as $\Omega(\cdot)$ is weakly separable in l and $\underline{x}$, the inverse factor share rule applies via the application of the Atkinson and Stiglitz (1976) theorem. Indeed, with individuals of different types earning different wages, the concept of an inverse factor cost share is not well defined. Even ignoring the fact that commodity taxes differ across types, the ratio of market price of x_i to its full opportunity cost will be $q_i/(q_i + a_i w^h)$ for an $h-$type person and $q_i/(q_i + a_i w^j) \neq q_i/(q_i + a_i w^j)$ for a $j-$type person.

[18] In the absence of the single-crossing (agent-monotonicity) property, the existence of such a group is not guaranteed.

(i) If preferences are weakly separable in labor supply and goods, optimal commodity taxes are uniform.

(ii) Assume preferences are weakly separable in leisure and goods. Then:

- *If there exists a group H for whom $\lambda^{kH} = 0$ for all $k = 1, 2, \ldots, H-1$, this group should face no differential commodity taxes.*
- *All other groups, $h = 1, 2, \ldots, H-1$, face a relative tax on good s and i given by (17). This rate differs across types.*
- *If there are two goods whose consumption take the same time ($a_i = a_s$), they should be subjected to the same tax rate.*

5 Anonymous transactions

The results of the previous section rests on the assumption that personal consumption levels are publicly observable. This assumption, which enables the government to levy non-linear commodity taxes, is rather hard to justify on informational grounds. It is more realistic to assume that the tax administration has information on anonymous transactions (i.e. aggregate sales of a commodity rather than who purchases how much). This is the standard assumption in the literature, so much so that it has been used as part of very definition of indirect taxes. [See, e.g., Atkinson and Stiglitz (1980, p. 427)]. Under this circumstance, non-linear commodity taxes are not feasible. If , for instance, the tax rate is linked to the quantity purchases, the buyer can avoid higher taxes by splitting the transactions. As a rule, only linear commodity taxes are available.

Given this informational structure, one may proceed to characterize Pareto-efficient allocations that are constrained not only by the standard self-selection constraints and the resource balance, but also by the linearity of commodity taxes. To do this, I derive an optimal revelation mechanism. The mechanism consists of a set of type-specific before-tax incomes, I^h's, aggregate expenditures on private goods, c^h's, and a vector of commodity tax rates (same for everyone): $\underline{t} = (t_1, t_2, \ldots, t_n)$. This procedure determines the commodity tax rates right from the outset. A complete solution to the optimal tax problem per-se then requires only the design of a general income tax function. Note that instead of commodity taxes, the mechanism may equivalently specify the consumer prices: $\underline{q} = (q_1, q_2, \ldots, q_n)$, where $q_i = 1 + t_i$ $(i = 1, 2, \ldots, n)$. The mechanism assigns $(\underline{q}, c^h, I^h)$ to an individual who reports type h; the consumer then allocates c^h between the produced goods, $\underline{x}$.[19]

Formally, given any vector $(\underline{q}, c, I)$, an individual of type h solves

$$\max_{\underline{x}} \quad \boldsymbol{u^h}(\underline{x}, I) \tag{19a}$$

$$\text{subject to} \quad \sum_{i=1}^{n} q_i x_i = c. \tag{19b}$$

Denote, with some abuse of notation, the resulting "conditional" demand functions by $\boldsymbol{x_i^h}(\underline{q}, c, I)$,[20] and the indirect utility function by

$$\boldsymbol{v^h}(\underline{q}, c, I) \equiv \boldsymbol{u^h}\left(\underline{x}^h(\underline{q}, c, I), I\right).$$

Next, define

$$x_i^h = \boldsymbol{x_i^h}(\underline{q}, c^h, I^h), \;\; x_i^{hk} = \boldsymbol{x_i^h}(\underline{q}, c^k, I^k), \;\; v^h = \boldsymbol{v^h}(\underline{q}, c^h, I^h), \;\; v^{hk} = \boldsymbol{v^h}(\underline{q}, c^k, I^k). \tag{20}$$

Pareto-efficient "allocations" (constrained by incentive compatibility and linearity of commodity taxes) can then be described as follows. Maximize[21]

$$\sum_h \delta^h v^h, \tag{21}$$

[19]Strictly speaking, this procedure does not characterize "allocations" as such; the optimization is over a mix of quantities *and* prices. However, given the commodity prices, utility maximizing individuals would choose the quantities themselves. We can thus think of the procedure as indirectly determining the final allocations.

[20]These functions are *conditional* on c and I; they differ from the customary Marshallian demand functions. Specifically, $\boldsymbol{x_i^h}(.)$ as defined here has a different functional form from its counterpart in Section 3.

[21]With one extra degree of freedom in setting commodity tax rates, t_1 is set equal to zero so that $q_1 = 1$.

with respect to $q_2, q_3, \ldots, q_n, c^h$ and I^h; subject to the resource constraint

$$\sum_h \pi^h \left[(I^h - c^h) + \sum_{i=2}^{n} (q_i - 1) x_i^h \right] \geq \bar{R}, \tag{22}$$

and the self-selection constraints

$$v^h \geq v^{hk}, \quad h, k = 1, 2, \ldots, H. \tag{23}$$

Denote the Lagrange multipliers associated with the resource balance (22) by μ and the self-selection constraints (23) by λ^{hk} $(h, k = 1, 2, \ldots, H)$. Let $\hat{x}_i^h$ denote the *compensated* version of individual h's conditional demand for x_i as determined by problem (19a)–(19b). Finally, define

$$\boldsymbol{A} \equiv \begin{pmatrix} \sum_h \pi^h \frac{\partial \hat{x}_2^h}{\partial q_2} & \sum_h \pi^h \frac{\partial \hat{x}_2^h}{\partial q_3} & \cdots & \sum_h \pi^h \frac{\partial \hat{x}_2^h}{\partial q_n} \\ \sum_h \pi^h \frac{\partial \hat{x}_3^h}{\partial q_2} & \sum_h \pi^h \frac{\partial \hat{x}_3^h}{\partial q_3} & \cdots & \sum_h \pi^h \frac{\partial \hat{x}_3^h}{\partial q_n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_h \pi^h \frac{\partial \hat{x}_n^h}{\partial q_2} & \sum_h \pi^h \frac{\partial \hat{x}_n^h}{\partial q_3} & \cdots & \sum_h \pi^h \frac{\partial \hat{x}_n^h}{\partial q_n} \end{pmatrix}. \tag{24}$$

I prove in the Appendix that an interior solution to the problem (21)–(23) satisfies the following equations.

$$\begin{pmatrix} t_2 \\ t_3 \\ \vdots \\ t_n \end{pmatrix} = -\frac{1}{\mu} \boldsymbol{A^{-1}} \begin{pmatrix} \sum_h \sum_{k \neq h} \lambda^{kh} (x_2^{kh} - x_2^h) \frac{\partial v^{kh}}{\partial c^h} \\ \sum_h \sum_{k \neq h} \lambda^{kh} (x_3^{kh} - x_3^h) \frac{\partial v^{kh}}{\partial c^h} \\ \vdots \\ \sum_h \sum_{k \neq h} \lambda^{kh} (x_n^{kh} - x_n^h) \frac{\partial v^{kh}}{\partial c^h} \end{pmatrix}. \tag{25}$$

5.1 Labor separability

Assume U is weakly separable in L and $\underline{x}$, so that $\boldsymbol{u^h}(\underline{x}, I) = U\left(\frac{I}{w^h}, f(\underline{x})\right)$. It then follows from the first-order conditions of problem (19a)–(19b) that x_i^h $(i = 1, 2, \ldots, n;$ $k, h = 1, 2, \ldots, H)$ is the solution to

$$\frac{f_i(\underline{x}^h)}{f_1(\underline{x}^h)} = q_i, \tag{26}$$

$$\sum_{i=1}^{n} q_i x_i^h = c^h. \tag{27}$$

Similarly, x_i^{kh} is the solution to

$$\frac{f_i(\underline{x}^{kh})}{f_1(\underline{x}^{kh})} = q_i, \tag{28}$$

$$\sum_{i=1}^{n} q_i x_i^{kh} = c^h. \tag{29}$$

Consequently, $x_i^h = x_i^{kh}$. It then follows from (25) that $t_2 = t_3 = \ldots = t_n = 0$: Optimal commodity taxes are uniform.

5.2 Leisure separability

Assume Ω is weakly separable in l and $\underline{x}$, so that $\boldsymbol{u^h}(\underline{x}, I) = \Omega\left(1 - \frac{I}{w^h} - \sum_{s=1}^{n} a_s x_s, f(\underline{x})\right).$ It now follows from the first-order conditions of problem (19a)–(19b) that x_i^h $(i = 1, 2, \ldots, n;\ k, h = 1, 2, \ldots, H)$ is the solution to

$$\frac{-a_i \Omega_l^h\big|_{\underline{x}^h} + \Omega_i^h\big|_l}{-a_1 \Omega_l^h\big|_{\underline{x}^h} + \Omega_1^h\big|_l} = \frac{q_i}{q_1} = q_i, \tag{30}$$

$$\sum_{i=1}^{n} q_i x_i^h = c^h, \tag{31}$$

where $\Omega_l^h|_{\underline{x}^h}$ and $\Omega_i^h|_l$ $(i = 1, 2, \ldots, n)$ are defined as previously through (18a). Similarly, x_i^{kh} is the solution to

$$\frac{-a_i \Omega_l^{kh}\big|_{\underline{x}^{kh}} + \Omega_i^{kh}\big|_l}{-a_1 \Omega_l^{kh}\big|_{\underline{x}^{kh}} + \Omega_1^{kh}\big|_l} = \frac{q_i}{q_n} = q_i, \tag{32}$$

$$\sum_{i=1}^{n} q_i x_i^{kh} = c^h. \tag{33}$$

Note that, in contrast to (18a), x_i^{kh} and not x_i^h, appears as an argument for $\Omega_l^{kh}\big|_{\underline{x}^{kh}}$ and $\Omega_i^{kh}\big|_l$ $(i = 1, 2, \ldots, n)$. Optimality of uniform taxation with leisure separability thus depends on whether the systems of equations (30)–(31) and (32)–(33) have identical solutions ($\underline{x}^h = \underline{x}^{kh}$) or not. The equality is not obvious because where I^h/w^h appears in the arguments of (30)–(31), it is replaced by I^h/w^k in the arguments of (32)–(33).

It is plain from equations (30)–(31) and (32)–(33) that if $a_i = 0$ for all $i = 1, 2, \ldots, n$ (which transforms this model to the traditional setup where consumption does not take time), separability in leisure calls for uniform taxation. More interestingly, one can show that if a_i is the same for all $i = 1, 2, \ldots, n$ (and nonzero), leisure separability implies uniform taxation (see the Appendix). As with the case with nonlinear commodity taxes, however, separability in leisure does *not* necessarily imply that the tax rates on the goods move positively with their time consumption coefficients, a_i's. There is even no presumption now that tax rates must be equal for two goods i and s with the property $a_i = a_s$—a result which held when commodity taxes were nonlinear.

The following proposition summarizes the results of this subsection.

Proposition 4 *Assume the economy is inhabited by individuals who differ in earning ability but have identical preferences. It takes each person a_i units of time to consume one unit of good i ($i = 1, 2, \ldots, n$). Assume further that only aggregate purchases are publicly observable so that feasible commodity taxes are linear.*

(i) If preferences are weakly separable in labor supply and goods, optimal commodity taxes are uniform.

(ii) Assume preferences are weakly separable in leisure and goods. Then:

- *Optimal commodity taxes are characterized by equations (25) when t_1 is normalized at zero.*
- *If time taken in consumption is the same for all goods, optimal commodity taxes are uniform.*
- *Suppose it takes more time to consume one unit of good i than one unit of good s ($a_i > a_s$). It does not follow that good i should necessarily have a higher tax rate than good s.*
- *Two goods whose consumption take the same time ($a_i = a_s$) should not necessarily have identical tax rates.*

6 Conclusion

This paper has studied the problem of optimal taxation of commodities when consumption is a time consuming activity. The first important lesson that has emerged is that assuming preferences are separable in goods and labor supply, or separable in goods and leisure, have markedly different implications for the structure of optimal commodity taxes. The two assumptions are not the same with Becker's (1965) formulation of labor supply. With labor separability, the traditional uniform taxation results of optimal tax theory continue to hold. With leisure separability, on the other hand, consumption time is a major ingredient of optimal tax rates. Which separability assumption is more "realistic" is of course an empirical question. And to the extent that the separability assumption per se has been put to test (within the traditional model of labor supply in which the two are equivalent), it does not seem to have fared that well.[22]

The second lesson that has emerged concerns one- versus many-consumer economies when separability holds between leisure and goods. In one-consumer economies, time taken in consumption is the sole determinant of optimal tax rates, and the optimal commodity tax on every good is related directly to the ratio of its full cost to its market price. The rule does *not* generalize in its simple form to economies with heterogeneous agents; however. When consumers are all alike, the only effect of differential commodity taxation is the differential time spent on consumption of different goods. On the other hand, when individuals differ in earning ability, the time spent by different individuals in consumption (even for the same good) also become relevant (because their opportunity costs of time differ). This enables commodity taxes to have redistributive and incentive

[22] See, e.g., Browning and Meghir (1991). Kleven (2004, p. 553) claims that the traditional results of optimal tax theory are not useful because they either (i) "rely on unreasonable simplifications" or/and (ii) "the information needed for optimal policy, essentially global knowledge of preferences and demand, is simply not obtainable." On the other hand "the theory of optimal taxation presented here is less vulnerable to the above critique." However, this is bogus. The "inverse factor share rule" is based on precisely the same "unreasonable simplifications" that Kleven criticizes. At the same time, without those assumptions, we will need precisely the same kind of information that he says is "simply not obtainable."

effects that separability can no longer isolate. These considerations too become relevant in determining the optimal commodity taxes. In particular, commodity tax rates do not necessarily move positively with time taken in consumption.

Appendix

Proof of (6): The optimal tax problem is summarized by the Lagrangian

$$\mathcal{L} = v(\underline{\tilde{q}}, w, G) + \mu \left[\sum_{i=1}^{n} t_i x_i - G - \bar{R} \right].$$

The first-order condition for this problem is

$$\frac{\partial \mathcal{L}}{\partial t_s} = \frac{\partial v}{\partial \tilde{q}_s} + \mu \left[\sum_{i=1}^{n} t_i \frac{\partial x_i}{\partial \tilde{q}_s} + x_s \right] = 0.$$

Simplifying these equations, using Roy's identity, results in

$$\sum_{i=1}^{n} t_i \frac{\partial x_i}{\partial \tilde{q}_s} = -\frac{\mu - \alpha}{\mu} x_s,$$

where α is the representative individual's marginal utility of income. This can be rewritten, using the Slutsky equation,

$$\sum_{i=1}^{n} t_i \frac{\partial x_i^c}{\partial \tilde{q}_s} = -\frac{\mu - \gamma}{\mu} x_s, \tag{A1}$$

where x_i^c denotes the compensated demand for good i, and $\gamma \equiv \alpha + \sum_i t_i(\partial x_i / \partial G)$ is the "social marginal utility of income".

Next, from the properties of the Slutsky matrix, one has

$$\sum_{i=1}^{n} \tilde{q}_i \frac{\partial x_i^c}{\partial \tilde{q}_s} + w \frac{\partial l^c}{\partial \tilde{q}_s} = 0, \tag{A2}$$

where $l^c(.)$ is the compensated demand for leisure. Now, Sandmo (1974) has shown that when preferences are separable in l and $\underline{x}$, with the latter subutility being homothetic, it will be the case that

$$\frac{\partial l^c}{\partial \tilde{q}_s} = \eta x_s, \tag{A3}$$

where η is independent of s. Substituting in (A2)

$$\sum_{i=1}^{n} \tilde{q}_i \frac{\partial x_i^c}{\partial \tilde{q}_s} = -\eta w x_s. \tag{A4}$$

Eliminating x_s between (A1) and (A4) yields,

$$\sum_{i=1}^{n}\left[\frac{\mu t_i}{\mu-\gamma}-\frac{\tilde{q}_i}{\eta w}\right]\frac{\partial x_i^c}{\partial \tilde{q}_s}=0, \qquad s=1,2,\ldots,n. \tag{A5}$$

Assuming that the matrix $[\partial x_i^c/\partial \tilde{q}_s]$ is non-singular, the solution to the system of equations (A5) is characterized by

$$\frac{t_i}{\tilde{q}_i}=\frac{1}{\eta w}\frac{\mu-\gamma}{\mu}. \tag{A6}$$

Finally, observe that from (A3) and symmetry of the Slutsky matrix,

$$\eta=\frac{1}{x_s}\frac{\partial l^c}{\partial \tilde{q}_s}=\frac{1}{x_s}\frac{\partial x_s^c}{\partial w}=\frac{\epsilon}{w},$$

where ϵ is the cross-price elasticity of compensated demand for l with respect to any one of the produced goods (same for all goods). Substituting this value for η in (A6) and rearranging, one arrives at equation (6) in the text.

Proof of (10): The optimal tax problem is summarized by the Lagrangian

$$\mathcal{L}=\sum_{h=1}^{H}\pi^h W\left(v^h(\cdot)\right)+\mu\left[\sum_{i=1}^{n}t_i\left(\sum_{h=1}^{H}\pi^h x_i^h\right)-G-\bar{R}\right],$$

where $v^h(\cdot)=v(\underline{\tilde{q}}^h, w^h, G)$. The first-order conditions are, for all $s=1,2,\ldots,n$,

$$\frac{\partial\mathcal{L}}{\partial t_s} = \sum_{h=1}^{H}\pi^h W'\left(v^h(\cdot)\right)\frac{\partial v^h}{\partial \tilde{q}_s^h}+\mu\left[\sum_{h=1}^{H}\pi^h x_s^h+\sum_{i=1}^{n}t_i\left(\sum_{h=1}^{H}\pi^h\frac{\partial x_i^h}{\partial \tilde{q}_s^h}\right)\right]=0,$$

$$\frac{\partial\mathcal{L}}{\partial G} = \sum_{h=1}^{H}\pi^h W'\left(v^h(\cdot)\right)\frac{\partial v^h}{\partial G}+\mu\left[\sum_{i=1}^{n}t_i\left(\sum_{h=1}^{H}\pi^h\frac{\partial x_i^h}{\partial G}\right)-1\right]=0.$$

Manipulation of these equations, using Roy's identity, yields

$$\sum_h(\mu-\beta^h)\pi^h x_s^h+\mu\sum_i t_i\left(\sum_h \pi^h\frac{\partial x_i^h}{\partial \tilde{q}_s^h}\right) = 0, \tag{A7}$$

$$\sum_h \pi^h\left[\beta^h+\mu\sum_i t_i\frac{\partial x_i^h}{\partial G}\right] = \mu. \tag{A8}$$

Slutsky decomposition of $\partial x_i^h / \partial \tilde{q}_s^h$ terms in (A7), denoting the elements of the associated Slutsky matrix by S_{is}^h, allows one to simplify (A7) further and rewrite equations (A7)–(A8) as

$$\begin{aligned}\sum_h (\mu - \gamma^h)\pi^h x_s^h + \mu \sum_h \sum_i t_i \pi^h S_{is}^h &= 0, & \text{(A9)}\\ \sum_h \pi^h \gamma^h &= \mu. & \text{(A10)}\end{aligned}$$

Next, from the properties of Slutsky matrix,

$$\sum_i (q_i + a_i w^h) S_{is}^h + w^h S_{ls}^h = 0. \tag{A11}$$

As with (A3), following Sandmo (1974), one can write

$$S_{ls}^h = \eta^h x_s^h, \tag{A12}$$

where η^h is independent of s. Substituting from (A12) into (A11) then results in

$$\sum_i (q_i + a_i w^h) S_{is}^h = -\eta^h w^h x_s^h. \tag{A13}$$

Finally, substituting from (A13) into (A9) and simplifying, one arrives at the following system of n equations (for $s = 1, 2, \ldots, n$),

$$\sum_h \sum_i \left[t_i - \frac{\mu - \gamma^h}{\mu\, \eta^h w^h} (q_i + a_i w^h) \right] \pi^h S_{is}^h = 0.$$

Observe that the assumptions of weak-separability and homotheticity in goods imply that $S_{is}^h = (w^h L^h) S_{is}$. Then, assuming that the matrix S_{is} is non-singular, the solution to above is given by

$$t_i = \frac{\mu - \gamma^h}{\mu\, \eta^h w^h} (q_i + a_i w^h). \tag{A14}$$

Now from (A12),

$$\eta^h w^h = \frac{S_{ls}^h}{x_s^h}\, w^h = \epsilon^h,$$

where ϵ^h is type h's cross-price elasticity of compensated demand for l with respect to any one of the produced gods (same for all goods). Substituting for η^h from this relationship, and for μ from (A10), into (A14) and rearranging, one arrives at equation (10) in the text.

Proof of (17): Summarize the government's optimization problem by

$$\mathcal{L} = \sum_h \delta^h u^h + \mu\big[\sum_h \pi^h (I^h - \sum_i x_i^h) - \bar{R}\big] + \sum_h \sum_{k \neq h} \lambda^{hk}(u^h - u^{hk}).$$

Rearrange the terms in the above expression to rewrite it as

$$\mathcal{L} = \sum_h \big(\delta^h + \sum_{k \neq h} \lambda^{hk}\big) u^h + \mu\big[\sum_h \pi^h (I^h - \sum_i x_i^h) - \bar{R}\big] - \sum_h \sum_{k \neq h} \lambda^{hk} u^{hk}.$$

This yields the following first-order conditions for $h = 1, 2, \ldots, H$, and $i = 1, 2, \ldots, n$.

$$\frac{\partial \mathcal{L}}{\partial x_i^h} = (\delta^h + \sum_{k \neq h} \lambda^{hk}) u_i^h - \mu \pi^h - \sum_{k \neq h} \lambda^{kh} u_i^{kh} = 0, \tag{A15}$$

$$\frac{\partial \mathcal{L}}{\partial I^h} = (\delta^h + \sum_{k \neq h} \lambda^{hk}) u_I^h + \mu \pi^h - \sum_{k \neq h} \lambda^{kh} u_I^{kh} = 0. \tag{A16}$$

Note that the calculation of the derivatives of $\sum_h \sum_{k \neq h} \lambda^{hk} u^{hk}$ results in the transposition of their h and k indices.[23] The system of equations (A15)–(A16), in conjunction with the economy's resource constraint, characterize the Pareto-efficient allocations of the economy constrained by self-selection.

To characterize the optimal commodity taxes that implement the (constrained) Pareto efficient allocations, manipulate (A15) to arrive at, for $h = 1, 2, \ldots, H$, and $i, s = 1, 2, \ldots, n$,

$$\frac{u_s^h}{u_i^h}\Big|_I = \frac{1 + \sum_{k \neq h} \lambda^{kh} u_s^{kh} / \mu \pi^h}{1 + \sum_{k \neq h} \lambda^{kh} u_i^{kh} / \mu \pi^h}. \tag{A17}$$

[23] To simplify notation, I use $\sum_h$ for $\sum_{h=1}^H$ and $\sum_{k \neq h}$ for $\sum_{\substack{k=1 \\ k \neq h}}^H$.

where a subscript on u denotes a partial derivative. This equation can be rewritten in the following useful form,

$$\left.\frac{u_s^h}{u_i^h}\right|_I = 1 + \sum_{k\neq h} \frac{\lambda^{kh} u_i^{kh}}{\mu\pi^h} \left(\frac{u_s^{kh}}{u_i^{kh}} - \frac{u_s^h}{u_i^h} \right). \tag{A18}$$

This is achieved by multiplying equation (A17) by $1+\sum_{k\neq h}\lambda^{kh}u_i^{kh}/\mu\pi^h$ and collecting terms. Observe that u_s^h/u_i^h in above denotes the h-type individual's marginal rate of substitution between goods s and i *assuming* income, or labor supply, is kept constant. It is this (conditional) marginal rate of substitution that consumers set equal to the relative consumer prices of goods s and i, when they make their decisions on labor supply and consumption in the market and face a tax schedule conditioned on income. Consequently, at the optimum, u_s^h/u_i^h, as characterized by (A18), is equal to $(1+t_s^h)/(1+t_i^h)$ where t_s^h and t_i^h are the commodity taxes h faces on s and h.

Finally, using (12), one may rewrite the right-hand side of (A18) in terms of the derivatives of $\Omega(.)$ to utilize the separability property between leisure and goods. It follows from (12), and the notation used in the text, that

$$\begin{aligned}
\left.u_s^h\right|_I &= -a_s \left.\Omega_l^h\right|_{\underline{x}^h} + \left.\Omega_s^h\right|_l, \\
\left.u_i^h\right|_I &= -a_i \left.\Omega_l^h\right|_{\underline{x}^h} + \left.\Omega_i^h\right|_l, \\
\left.u_s^{kh}\right|_I &= -a_s \left.\Omega_l^{kh}\right|_{\underline{x}^h} + \left.\Omega_s^{kh}\right|_l, \\
\left.u_i^{kh}\right|_I &= -a_i \left.\Omega_l^{kh}\right|_{\underline{x}^h} + \left.\Omega_i^{kh}\right|_l.
\end{aligned}$$

Substitute these expressions in the right-hand side of (A18) and simplify to arrive at

$$\begin{aligned}
(-a_s\Omega_l^h + \Omega_s^h) \; - \; (-a_i\Omega_l^h + \Omega_i^h) &= \sum_{k\neq h} \frac{\lambda^{kh}}{\mu\pi^h} \Big\{ \Omega_l^h \Omega_s^{kh} \left(a_s \frac{\Omega_i^{kh}}{\Omega_s^{kh}} - a_i \right) \\
- \; \Omega_s^h \Omega_l^{kh} \left(a_s \frac{\Omega_i^h}{\Omega_s^h} - a_i \right) &+ \Omega_s^h \Omega_s^{kh} \left(\frac{\Omega_i^h}{\Omega_s^h} - \frac{\Omega_i^{kh}}{\Omega_s^{kh}} \right) \Big\}.
\end{aligned} \tag{A19}$$

Next, note that from weak separability of $\Omega(.)$,

$$\frac{\Omega_i^{kh}}{\Omega_s^{kh}} = \frac{\Omega_i^h}{\Omega_s^h}.$$

Substituting the above in (A19) and simplifying, one obtains

$$\frac{-a_s\Omega_l^h + \Omega_s^h}{-a_i\Omega_l^h + \Omega_i^h} - 1 = \frac{a_s\Omega_i^h - a_i\Omega_s^h}{-a_i\Omega_l^h + \Omega_i^h} \sum_{k\neq h} \frac{\lambda^{kh}\Omega_s^{kh}}{\mu\pi^h} \left(\frac{\Omega_l^h}{\Omega_s^h} - \frac{\Omega_l^{kh}}{\Omega_s^{kh}} \right), \tag{A20}$$

which is equivalent to equation (17) in the text.

Proof of $a_s = a_i \Rightarrow t_s^h = t_i^h$, $h = 1, 2, \ldots, H$, when commodity taxes are nonlinear: Set $a_s = a_i = a$ in (A20) and simplify to arrive at

$$(\Omega_s^h - \Omega_i^h) \left[1 + \frac{a}{-a\Omega_l^h + \Omega_i^h} \sum_{k\neq h} \frac{\lambda^{kh}\Omega_s^{kh}}{\mu\pi^h} \left(\frac{\Omega_l^h}{\Omega_s^h} - \frac{\Omega_l^{kh}}{\Omega_s^{kh}} \right) \right] = 0.$$

A solution to this is $\Omega_s^h = \Omega_i^h$. It then follows that

$$\frac{1+t_s^h}{1+t_i^h} = \frac{u_s^h}{u_i^h}\bigg|_I = \frac{-a_s\Omega_l^h\big|_{\underline{x}^h} + \Omega_s^h\big|_l}{-a_i\Omega_l^h\big|_{\underline{x}^h} + \Omega_i^h\big|_l} = 1.$$

Derivation of (25): Summarize the problem by the Lagrangian

$$\mathcal{L} = \sum_h \delta^h v^h(\underline{q}, c^h, I^h) + \mu\{ \sum_h \pi^h \big[(I^h - c^h) + \sum_{i=2}^{n} (q_i - 1) x_i^h \big] - \bar{R} \} + \sum_h \sum_{k\neq h} \lambda^{hk} (v^h - v^{hk}). \tag{A21}$$

Rearranging the terms, one may usefully rewrite the above Lagrangian expression as

$$\begin{aligned} \mathcal{L} &= \sum_h \big(\delta^h + \sum_{k\neq h} \lambda^{hk} \big) v^h - +\mu\{ \sum_h \pi^h \big[(I^h - c^h) \\ &+ \sum_{i=2}^{n} (q_i - 1) x_i^h \big] - \bar{R} \} - \sum_h \sum_{k\neq h} \lambda^{hk} v^{hk}. \end{aligned} \tag{A22}$$

The first-order conditions are, for all $h, k = 1, 2, \ldots, H$, and $t = 2, 3, \ldots, n$,

$$\frac{\partial \mathcal{L}}{\partial I^h} = \big(\delta^h + \sum_{k \neq h} \lambda^{hk}\big) v_I^h + \mu \pi^h \big[1 + \sum_{i=2}^{n} (q_i - 1) \frac{\partial x_i^h}{\partial I^h}\big] - \sum_{k \neq h} \lambda^{kh} v_I^{kh} = 0, \quad \text{(A23)}$$

$$\frac{\partial \mathcal{L}}{\partial c^h} = \big(\delta^h + \sum_{k \neq h} \lambda^{hk}\big) v_c^h + \mu \pi^h \big[-1 + \sum_{i=2}^{n} (q_i - 1) \frac{\partial x_i^h}{\partial c^h}\big] - \sum_{k \neq h} \lambda^{kh} v_c^{kh} = 0, \quad \text{(A24)}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial q_t} &= \sum_h \big(\delta^h + \sum_{k \neq h} \lambda^{hk}\big) v_t^h + \mu \sum_h \pi^h \big[\sum_{1=2}^{n} (q_i - 1) \frac{\partial x_i^h}{\partial q_t} + x_t^h\big] \\ &- \sum_h \sum_{k \neq h} \lambda^{hk} v_t^{hk} = 0. \end{aligned} \quad \text{(A25)}$$

Multiply equation (A24) by x_t^h, sum over h, add the resulting equation to (A25), and simplify. We will have the following system of equations for $i = 2, 3, \ldots, n$,

$$\begin{aligned} \sum_h \big(\delta^h + \sum_{k \neq h} \lambda^{hk}\big)\big(v_t^h + x_t^h v_c^h\big) &+ \mu \sum_h \pi^h \big[\sum_{i=1}^{n-1} (q_i - 1)\big(\frac{\partial x_i^h}{\partial q_t} + x_t^h \frac{\partial x_i^h}{\partial c^h}\big)\big] \\ &- \sum_h \sum_{k \neq h} \lambda^{kh} \big(v_t^{kh} + x_t^h v_c^{kh}\big) = 0. \end{aligned} \quad \text{(A26)}$$

Next, make use of Roy's identity to set, for all $t = 1, 2, \ldots, n$, and $h, k = 1, 2, \ldots, H$,

$$v_t^h + x_t^h v_c^h = 0, \quad \text{(A27)}$$

$$v_t^{kh} + x_t^{kh} v_c^{kh} = 0. \quad \text{(A28)}$$

Then use the Slutsky equation to write, for all $i, t = 1, 2, \ldots, n$, and $h = 1, 2, \ldots, H$,

$$\frac{\partial x_i^h}{\partial q_t} = \frac{\partial \hat{x}_i^h}{\partial q_t} - x_t^h \frac{\partial x_i^h}{\partial c^h}. \quad \text{(A29)}$$

Substituting from equations (A27)–(A28) and (A29) in (A26), making use of the symmetry of the Slutsky matrix, setting $q_i - 1 = t_i$, upon further simplification and rearrangement, one arrives at

$$\sum_{i=2}^{n} \big(\sum_h \pi^h \frac{\partial \hat{x}_t^h}{\partial q_i}\big) t_i = -\sum_h \sum_{k \neq h} \lambda^{kh} \big(x_t^{kh} - x_t^h\big) \frac{v_c^{kh}}{\mu}, \; t = 2, 3, \ldots, n. \quad \text{(A30)}$$

Equation (A30) is one way of characterizing the optimal commodity tax rates: t_i's.

To arrive at (25), use the definition of $\boldsymbol{A}$ in (24) to write out equations (A30) in matrix notation:

$$\boldsymbol{A}\begin{pmatrix} t_2 \\ t_3 \\ \vdots \\ t_n \end{pmatrix} = -\frac{1}{\mu}\begin{pmatrix} \sum_h \sum_{k\neq h} \lambda^{kh}\big(x_2^{kh} - x_2^h\big)\frac{\partial v^{kh}}{\partial c^h} \\ \sum_h \sum_{k\neq h} \lambda^{kh}\big(x_3^{kh} - x_3^h\big)\frac{\partial v^{kh}}{\partial c^h} \\ \vdots \\ \sum_h \sum_{k\neq h} \lambda^{kh}\big(x_n^{kh} - x_n^h\big)\frac{\partial v^{kh}}{\partial c^h} \end{pmatrix}. \tag{A31}$$

Premultiplying (A31) by $\boldsymbol{A^{-1}}$ then yields the system of equations (25) in the text.

Proof of $a_i = a$ $(i = 1, 2, \ldots, n) \Rightarrow$ uniform taxation when commodity taxes are linear: It follows from equations (30) and (33) that

$$\frac{-a_s\Omega_l^h + \Omega_s^h}{-a_i\Omega_i^h + \Omega_i^h} = \frac{-a_s\Omega_l^{kh} + \Omega_s^{kh}}{-a_i\Omega_i^{kh} + \Omega_i^{kh}}.$$

Multiplying through, simplifying and collecting terms yields

$$\frac{-a_s\Omega_l^h + \Omega_s^h}{-a_i\Omega_i^h + \Omega_i^h} = \frac{\Omega_s^{kh}}{\Omega_i^{kh}} + \frac{\Omega_l^{kh}}{\Omega_i^{kh}}\,\frac{a_i\Omega_s^h - a_s\Omega_i^h}{-a_i\Omega_l^h + \Omega_i^h}. \tag{A32}$$

I now show that if $a_i = a$ $(i = 1, 2, \ldots, n)$, $\underline{x}^h = \underline{x}^{kh}$ satisfies (A32) so that optimal taxes will be uniform.

Observe that if $\underline{x}^h = \underline{x}^{kh}$ optimal taxes are uniform and the left-hand side of (A32), which is equal to u_s^h/u_i^h, will be reduced to one. This implies that $\Omega_s^h = \Omega_i^h$. Substituting this into the right-hand side of (A32), its value will be reduced to $\Omega_s^{kh}/\Omega_i^{kh}$. But, with $\underline{x}^h = \underline{x}^{kh}$, weak separability in leisure implies that $\Omega_s^{kh}/\Omega_i^{kh} = \Omega_s^h/\Omega_i^h = 1$. Consequently, the right-hand side of (A32) will also be equal to one.

References

Atkinson, A.B. (1977) "Optimal taxation and the direct versus indirect tax controversy," *Canadian Journal of Economics*, **10**, 590–606.

Atkinson, A.B. and J.E. Stiglitz (1976) "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics*, **6**, 55–75.

Atkinson, A.B. and J.E. Stiglitz (1980) *Lectures on Public Economics*. New-York: Mc-Graw-Hill.

Becker, G.S. (1965) "A theory of the allocation of time," *Economic Journal*, **75**, 493–517.

Besley, T. and I. Jewitt (1995) "Uniform taxation and consumer preferences," *Journal of Public Economics*, **58**, 73–84.

Browning, M. and C. Meghir (1991) "The effects of male and female labor supply of commodity demands," *Econometrica*, **59**, 925–951.

Cremer, H. and F. Gahvari (1998) "On optimal taxation of housing," *Journal of Urban Economics*, **43**, 315–335.

Cremer, H., Gahvari, F. and N. Ladoux (1998) "Externalities and optimal taxation," *Journal of Public Economics*, **70**, 343–364.

Cremer, H., Gahvari, F. and N. Ladoux (2001) "Second-best pollution taxes and the structure of preferences," *Southern Economic Journal*, **68**, 258–280.

Cremer, H., and F. Gahvari (2002) "Nonlinear pricing, redistribution and optimal tax policy," *Journal of Public Economic Theory*, **4**, 139–161.

Cremer, H., P. Pestieau and J-C. Rochet (2001) "Direct versus indirect taxation: The design of the tax structure revisited," *International Economic Review*, **42**, 781–799.

Deaton, A. (1979) "Optimally uniform commodity taxes," *Economics Letters*, **2**, 357–361.

Deaton, A. and J. Muellbauer (1980) *Economics and consumer behavior*. Cambridge: Cambridge University Press.

Deaton, A. (1981) "Optimal taxes and the structure of preferences," *Econometrica*, **49**, 1245–1260.

Deaton, A. and N. Stern (1986) "Optimally uniform commodity taxes, taste differences and lump-sum grants," *Economics Letters*, **20**, 263–266.

Gahvari, F. and C.C. Yang (1993) "Optimal commodity taxation and household consumption activities," *Public Finance Quarterly*, **21**, 479–487.

Kleven, H.J. (2004) "Optimum taxation and the allocation of time," *Journal of Public Economics*, **88**, 545-557.

Mirrlees, J.A. (1971) “An exploration in the theory of optimal income taxation,” *Review of Economic Studies*, **38**, 175–208.

Mirrlees, J.A. (1976) “Optimal tax theory: a synthesis,” *Journal of Public Economics*, **6**, 327–358.

Sandmo, A. (1974) “A note on the structure of optimal taxation,” *American Economic Review*, **64**, 701–706.

Stiglitz, J.E. (1987) “Pareto efficient and optimal taxation and the new new welfare economics,” in *Handbook of Public Economics, Vol 2*, ed. Auerbach, A. and M. Feldstein. Amsterdam: North-Holland, 991–1042.