Queen's Economics Department Working Paper No. 1005

Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem

Robin Boadway
Department of Economics, Queen's University

Pierre Pestieau
Universitt'e de Liege

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

3-2002

Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem*

by

Robin Boadway
Queen's University, Canada

and

Pierre Pestieau
Université de Liege, Belgique

March 2002

Abstract

The Atkinson-Stiglitz Theorem states that if labor is weakly separable from goods in household utility functions, differential commodity taxation should not be not part of an optimal redistributive tax system. This Theorem, which is arguably the most policy-relevant result to come out of the optimal income tax literature, has come under considerable scrutiny in the literature. We consider how robust it is with respect to differences in needs or endowments of goods, more than one type of labor supply, differences in preference for leisure, and restrictions on policy instruments.

JEL Classification: H2

Key Words: Optimal taxation, indirect taxation

---

* Prepared for the Festschrift in Honor of Joe Stiglitz.

1. Introduction

One of the oldest controversies in tax theory involves the choice between direct and indirect taxation, in particular the issue of when differential commodity taxes are not a component of the optimal tax system. The early literature focussed on the efficiency role of commodity taxes: under what circumstances would the Ramsey tax system applied to a given household consist of a uniform tax on commodities, or equivalently a tax on income? The famous Corlett and Hague (1953-54) Theorem settled that. If all goods are 'equally substitutable' for leisure, differential commodity taxes should not be used. Otherwise, goods that are more complementary with leisure should bear higher commodity tax rates. As explained in Sandmo (1976), a utility function in which goods are separable from leisure, and which is homothetic in goods satisfies this property. This result, although an important methodological innovation, is of limited interest from a policy point of view since it abstracts from the redistributive role that the tax system plays.

The question of when differential commodity taxes should be used alongside a progressive income tax as part of a redistributive tax system was addressed in a well-known paper by Atkinson and Stiglitz (1976). Their result, the *Atkinson-Stiglitz (A-S) Separability Theorem*, has been seminal and has spawned a substantial literature.[1] Roughly speaking, the A-S Theorem states that if household utility functions are separable in goods and leisure, differential commodity taxes should not be used. This result is arguably the most relevant result for policy purposes to emerge from the optimal income tax literature initiated by Mirrlees (1971). It has been subject to considerable scrutiny in the literature, and special attention has been devoted to the circumstances in which it is violated and what it implies for the structure of commodity taxes. Interestingly, the analogue of the Corlett-Hague Theorem applies, albeit for different reasons. As shown by Edwards et al (1994) and Nava et al (1996), if weak separability is violated, higher tax rates should apply to goods that are relatively more complementary with leisure.

Our purpose in this paper is to revisit the A-S Theorem. We explore the robustness of the theorem to different specifications of household utility, of government information and of restrictions on policy instruments. We begin with a simple derivation of the A-S result, using a methodology that will be useful in synthesizing the various extensions. We then turn to those extensions, first focusing on the case where an optimal non-linear tax is in place, and then turning to the case where the government is restricted to a linear progressive tax.

The A-S Theorem

In this section, we adopt a simplified version of the model used by Atkinson and Stiglitz (1976), retaining their essential assumptions. For simplicity, we assume that there are two types of households who differ only in their wage rates $w_i \quad (i = 1, 2)$, where $w_2 > w_1$. There are n_i households of type i. We assume that their are only two goods, denoted x and z, along with labor ℓ, and that households have identical weakly separable utility functions of the form $u\left(g\left(x, z\right), \ell\right)$.[2] The utility function is strictly concave, and both goods as well as leisure are normal. The market (pre-tax) income of a type-i household is $y_i \equiv w_i \ell_i$. Following Guesnerie (1995), the government is assumed to be able to observe household incomes as well as anonymous transactions in the goods market. It can therefore implement a non-linear income tax as well as proportional commodity taxes.[3] As is well-known, only the structure of commodity taxes, and not their level, constitutes an independent policy instrument: proportional commodity taxes can be replicated by an appropriate adjustment in the income tax schedule. Therefore, we can normalize the commodity tax rate on good x to be zero, and treat the tax rate on z as the policy instrument reflecting the differential commodity tax structure. Let t be the per unit tax on purchases of good z. If $t = 0$ in the optimum, the redistributive objectives of government can be achieved by an income tax alone. Goods prices are normalized to unity, and we define the consumer price of good z to be $q \equiv 1 + t$.

To facilitate our analysis, we disaggregate household decision-making into two stages.[4] In the first stage, the household chooses labor supply, earns income, pays income taxes, and ends up with disposable income. In the second stage, disposable income is allocated between the two goods. Consider the second stage first. Let c_i be disposable income, where $c_i = x_i + qz_i$. Given the separable utility function, a household of type i solves the following problem:

$$\max_{\{z_i\}} \quad g\left(c_i - qz_i, z_i\right)$$

where c_i and q are given. From the first-order conditions, $g_z^i / g_x^i = q$,[5] we obtain the demand function $z(q, c_i)$ and the value function $h(q, c_i)$. Applying the envelope function, we obtain:

$$h_q^i = -g_x^i z_i, \qquad \text{and} \qquad h_c^i = g_x^i$$

In the first stage, the household chooses labor supply, given the income tax chosen by the government and the anticipated outcome of stage 2. Effectively, the household is choosing earned income y_i and, via the income tax, c_i. For this stage, we follow the standard procedure of optimal income tax analysis initiated by Stiglitz (1982) of allowing the government to choose y_i and c_i implicitly by its choice of an income tax schedule. Individual utility functions are reformulated in terms of what the government can observe as follows:

$$v^i\left(h\left(q, c_i\right), \frac{y_i}{w_i}\right) \equiv u\left(h\left(q, c_i\right), \ell_i\right)$$

The government is assumed to maximize a utilitarian objective function, although any quasi-concave function in individuals utilities would give the same results. The Lagrange expression for the optimal income and commodity tax problem of the government can then be written as:

$$\mathcal{L} = \sum_{i=1,2} n_i v^i\left(h\left(q, c_i\right), \frac{y_i}{w_i}\right) + \lambda \sum_{i=1,2} n_i\left(y_i + tz^i\left(q, c_i\right) - c_i\right)$$
$$+\gamma\left[v^2\left(h\left(q, c_2\right), \frac{y_2}{w_2}\right) - v^2\left(h\left(q, c_1\right), \frac{y_1}{w_2}\right)\right]$$

The first constraint reflects the government budget constraint, and assumes no net revenue requirement. The second constraint is the incentive constraint and reflects the fact that this will only be binding for type-2 households. The Lagrange multipliers associated with the two constraints are λ and γ, respectively.

The relevant first-order conditions for our purposes are those with respect to c_1, c_2 and q:[6]

$$n_1 v_h^1 h_c^1 - \lambda n_1 \left(1 - t\frac{\partial z^1}{\partial c_1}\right) - \gamma \widehat{v}_h^2 \widehat{h}_c^2 = 0 \tag{1}$$

$$n_2 v_h^2 h_c^2 - \lambda n_2 \left(1 - t\frac{\partial z^2}{\partial c_2}\right) + \gamma v_h^2 h_c^2 = 0 \tag{2}$$

$$\sum_{i=1,2} n_i v_h^i h_q^i + \lambda \sum_{i=1,2} n_i \left(z_i + t\frac{\partial z^i}{\partial q}\right) + \gamma \left(v_h^2 h_q^2 - \widehat{v}_h^2 \widehat{h}_q^2\right) = 0 \tag{3}$$

where the 'hat' refers to a type-2 household who is mimicking a type-1. Multiplying (1) by z_1, (2) by z_2, and adding both equations to (3), we immediately obtain the A-S Theorem:

$$t \sum_{i=1,2} n_i \frac{\partial \widetilde{z}^i\left(q\right)}{\partial q} = \frac{\gamma}{\lambda} \widehat{v}_h^2 \widehat{h}_c^2 \left(z_1 - \widehat{z}_2\right) = 0 \tag{4}$$

where we have used the envelope condition on q from the second-stage of the household's problem, $h_q^i + g_x^i z_i = 0$, which also applies to the mimicking type-2 household. The function $\widetilde{z}^i\left(q\right)$ represents the compensated demand for z_i, where the compensation takes the form of disposable income. The second equality follows from the fact that type-1 households and the mimicking type-2 households have the same disposable income c_1, but differ in their labor supplies. By separability, they will consume the same bundle of goods, so $\widehat{z}_2 = z_1$. Therefore, when the income tax is being set optimally, $t = 0$, so no differential commodities taxes should be applied. This demonstrates the A-S Theorem.[7]

Next we turn to two sorts of extensions to the above analysis. In the first, taken up in the following section, we modify the manner in which goods enter the subutility function $g\left(x, z\right)$ by allowing households to have different basic needs or, equivalently, different endowments of one of the goods. In the subsequent section, we consider different specification for labor supply.

3. Needs and Endowments

Suppose, in the manner of the Stone-Geary utility function, that households have some basic non-discretionary expenditures that must be made on one of the goods, say, z. The separable utility function can then be written $u(g(x, z-b), \ell)$, where b is non-discretionary expenditures on z. One interpretation that can be given to b, following Rowe and Woolley (1999), is that of a basic *need* for good z, such as sustenance, health spending, etc. Alternatively, b might be interpreted as an initial endowment, as in Cremer et al (2001), in which case it takes a negative value.[8] The only difference between the two approaches is that in case of initial endowments, these enter the overall resource constraint of the economy by adding to net output. Note that b might enter into the utility function in other ways, such as multiplicative. Since this would not affect our basic results, the additive form is adopted for simplicity. Note also that there may be a needs parameter associated with good x as well. Since the same analysis would apply to this case, we analyze only the case of non-discretionary expenditures in good z.[9]

If b were the same for all persons, it would obviously have no effect on the A-S Theorem derived in the previous section. The non-discretionary expenditures would simply be an element of the common utility function faced by all households, which would remain separable. Instead, we assume that b can differ across households. For expositional purposes, we assume that b can take on two values b_j, $j = 1, 2$. This implies that, in principle, there can now be four household types, $\{w_i, b_j\}$, $i, j = 1, 2$. In analyzing government policy, two informational settings are considered. In one, following Cremer et al (2001), the government can able observe neither w nor b. This is consistent with interpreting b as an endowment. In the other, the government can observe b, but not w. This is the assumption adopted by Rowe and Woolley (1999) in their analysis of needs. We consider these two cases in turn.

3.1 Government Does Not Observe Needs

With both w and b unobservable, the government faces a two-dimensional screening problem. This is the case analyzed by Cremer et al (2001). As is well-known, the analysis is complex and the results ambiguous, mainly because the direction in which the self-selection constraints bind is no longer unambiguous. We can simplify the analysis considerably without affecting the main results by assuming that each ability-type is associated with a given need type. Thus, a household of type w_i has a need of b_i. This leaves us with at most one binding self-selection constraint which, unlike in the previous section, can bind in either direction even under a utilitarian objective function. For example, if high-wage households also have high needs, the government may want to redistribute from the low-wage to the high-wage types. But for our purposes, that does not affect the results. Therefore, we proceed by assuming that the self-selection constraint applies downwards as in the previous section.

As before, we can proceed in a two-stage manner, assuming that in the first stage, labor supply and income are chosen, while in the second stage, disposable income is allocated between the two goods. The analysis of the second stage is identical to earlier. A type-i household chooses z_i to maximize $g(c_i - qz_i, z_i - b_i)$. This yields the demand function $z(q, c_i, b_i)$, and the value function $h(q, c_i, b_i)$. The envelope theorem again yields $h_q^i + g_x^i z_i = 0$.

In the first stage, the Lagrangean expression for the government's choice of $\{c_i, y_i, q\}$ is exactly as before, and the first-order conditions on c_i and q can be used to obtain the analog of (4):

$$t \sum_{i=1,2} n_i \frac{\partial \widetilde{z}^i(q)}{\partial q} = \frac{\gamma}{\lambda} \widehat{v}_h^2 \widehat{h}_c^2 (z_1 - \widehat{z}_2) \tag{5}$$

Unlike in the previous section, the right-hand side is generally not zero: it will only be so if $b_1 = b_2$. It can be shown that $t > 0$ if $b_2 > b_1$, and vice versa. That is, if high-wage households also have higher needs, the tax on z should be higher (assuming, of course, that the self-selection constraint on the high-wage types is binding).

This result can be illustrated using Figure 1, which depicts preferences and the budget constraint for a type-1 person and a mimicking type-2. Define the net (after-needs) consumption of good z by $\overline{z}_i \equiv z_i - b_i$. Then, the sub-utility function for the two types of individuals is identical in x and $\overline{z}$, and preferences over x and $\overline{z}$ are independent of labor supply. When $t = 0$, the budget constraints for each household are given by $c_i - b_i = x_i + \overline{z}_i$. The figure shows the choices of x_i and $\overline{z}_i$ for the two types of households when $b_2 > b_1$. As can be seen, $\overline{z}_1 - \widehat{\overline{z}}_2 < b_2 - b_1$ (recall that the 'hat' refers to the mimicker), which implies that $\widehat{z}_2 > z_1$. Since the mimicker purchases more of good z, the self-selection constraint can be weakened by imposing a tax on that good.

Quite clearly if instead we had the higher productivity workers having the higher needs, the self-selection constraint could apply in the other direction. This is surely the case when productivity differences are small and differences in needs huge. Then, both $t > 0$ or < 0 are also possible. With four types, the pattern of self-selection constraints becomes quite complex, but as shown by Cremer et al. (2001) the case for a non-zero tax, positive or negative, is very strong.

The upshot of this discussion is that if persons have different needs or endowments, the A-S Theorem will fail to be satisfied even if the utility function is weakly separable.

3.2 Government Observes Needs

Suppose now that the government can observe household needs b_j, but it cannot observe wage rates w_i. There are now four household types, and we denote the government's policy instruments by $\{c_{ij}, y_{ij}, q\}$, $i, j = 1, 2$. However, since needs are now observable, the population can be divided into the two identifiable need types $\{w_1, b_1; w_2, b_1\}$ and $\{w_1, b_2; w_2, b_2\}$. The second stage of the household problem is analogous to above, the only difference being that household demands and functions are now indexed by 'ij' rather than simply i.

The government can now condition its policies on need, and that simplifies matters considerably. In particular, it need worry only about incentive compatibility within each need type. The Lagrangean expression for the government can be written:

$$\mathcal{L} = \sum_i \sum_j n_{ij} v^{ij} \left(h\left(q, c_{ij}\right), \frac{y_{ij}}{w_i} \right) + \lambda \sum_i \sum_j n_{ij} \left(y_{ij} + t z^{ij}\left(q, c_{ij}\right) - c_{ij} \right)$$

$$+ \sum_j \gamma_j \left[v^{2j} \left(h\left(q, c_{2j}\right), \frac{y_{2j}}{w_2} \right) - v^{2j} \left(h\left(q, c_{1j}\right), \frac{y_{1j}}{w_2} \right) \right]$$

The first-order conditions on disposable income and q are:

$$n_{1j} v_h^{1j} h_c^{1j} - \lambda n_{1j} \left(1 - t \frac{\partial z^{1j}}{\partial c_{1j}} \right) - \gamma_j \widehat{v}_h^{2j} \widehat{h}_c^{2j} = 0 \qquad j = 1, 2$$

$$n_{2j} v_h^{2j} h_c^{2j} - \lambda n_{2j} \left(1 - t \frac{\partial z^{2j}}{\partial c_{2j}} \right) + \gamma_j v_h^{2j} h_c^{2j} = 0 \qquad j = 1, 2$$

$$\sum_i \sum_j n_{ij} v_h^{ij} h_q^{ij} + \lambda \sum_i \sum_j n_{ij} \left(z_{ij} + t \frac{\partial z^{ij}}{\partial q} \right) + \sum_j \gamma_j \left(v_h^{2j} h_q^{2j} - \widehat{v}_h^{2j} \widehat{h}_q^{2j} \right) = 0$$

It should be apparent that by combining these conditions, we obtain the analog of (4) derived earlier:

$$t \sum_i \sum_j n_{ij} \frac{\partial \widetilde{z}^{ij}\left(q\right)}{\partial q} = \sum_j \frac{\gamma_j}{\lambda} \widehat{v}_h^{2j} \widehat{h}_c^{2j} \left(z_{1j} - \widehat{z}_{2j} \right) = 0$$

The last equality comes about because within each need group j, a type-1 household and a type-2 mimicker has the same disposable income and the same value of b_j, so by separability, they have the same demand for good z, or $z_{1j} = \widehat{z}_{2j}$. Therefore, the A-S Theorem applies in this case. It ought also to be obvious that this result extends to other formulations of need, such as multiplicative. Provided the government can classify households by need, and utility functione are separable, the A-S Theorem applies.

The optimal income tax system is obviously more complicated in this case, since there is a different schedule for each need type.

4. Multiple Forms of Labor

In this section, we consider the robustness of the A-S Theorem when household labor supplies are disaggregated into more than one type. For simplicity, we assume that households can supply two types of labor, say ℓ_c and ℓ_d, whose interpretations will be discussed for various cases considered below. As in the previous section, the informational restrictions that face the government will be key in determining whether the A-S Theorem applies.

4.1 Two Types of Market Labor

Suppose each household supplies two types of labor to the market, receives a wage rate for each, and uses the proceeds to purchase goods. In this case, the utility function becomes $u\left(g\left(x,z\right),\ell_c,\ell_d\right)$. The two types of labor supply could be two different jobs, or the problem could be given an intertemporal interpretation, with ℓ_c and ℓ_d interpreted as present and future labor supply (where x and z can then be interpreted as present and future consumption). In this case, the applicability of the A-S Theorem depends on whether or not incomes from the two forms of labor supply y_c and y_d are observable.

If both y_c and y_d are observable either individually or in the aggregate, the analysis of Section 2 goes through with virtually no modification. The government's selection of an optimal tax policy involves selecting consumption levels and disposable incomes for the two types of households, as well as the commodity tax on z. The conditions on c_i and q are the same as before. Moreover, since the mimicker has the same disposable income as a type-1 person, separability ensures that $\widehat{z}_2 = z_1$, so the optimal commodity tax rate is zero ($t = 0$).

On the other hand, suppose that, say, y_c is observable, but y_d is not. This might correspond with the case in which labor supply ℓ_d is to the underground economy, as in Boadway et al (1994). Of course, for this interpretation to apply, one ought to model explicitly the penalty and detection technologies associated with the underground sector. However, that would serve only to complicate the story without affecting the main result. That result is that the A-S Theorem generally no longer applies if one source of income is not observable to the government.

The intuition is straightforward, even without a formal analysis. If only y_c is observable, the government can control only that part of disposable income that comes from ℓ_c. Assuming that the wage rate of the mimicker is higher that that of a type-1 household in the unobserved sector, it will generally be the case that $\widehat{y}_{d2} \neq y_{d1}$. That implies that $\widehat{c}_2 \neq c_1$, so that even with separable preferences, $\widehat{z}_2 \neq z_1$. So, for example, if $\widehat{y}_{d2} > y_{d1}$ because of the higher productivity of a type-2 person, $\widehat{z}_2 > z_1$, and it will be optimal to impose a tax on good z.

More generally, suppose the subutility function $g(x, z)$ is homothetic. In this case, the proportions in which the two goods are consumed by the two persons will be the same. Even in this case it will be optimal to tax good z. In fact, as Boadway et al (1994) show, the optimal commodity tax system is a proportional one on the two goods x and z. The point is that in the absence of full observability of income, a proportional income tax is no longer a perfect substitute for a proportional commodity tax. In the optimum, there needs to be a mix of the two taxes.

4.2 Household Production

Suppose that the second form of labor supply ℓ_d represents non-market or household production with no disposable income that can be used to purchase x and z. All disposable income comes from y_c, which is observed by the government. Assume that the utility function still takes the form $u(g(x, z), \ell_c, \ell_d)$, where the argument ℓ_d incorporates both the disutility of the non-market work as well as the the product of that work. In this case, the A-S Theorem still holds regardless of whether non-market labor is observed by the government. Indeed if all households share the same preferences, that will not be relevant.

The analysis is a straightforward application of that used in Section 2. The government controls y_c, and therefore disposable income that is used to purchase x and z. The existence of non-market labor complicates things slightly because it conditions the structure of the optimal non-linear income tax, and might in principle affect the direction in which the incentive constraint is binding. Suppose, for example, that the self-selection constraint is binding on type-2 households in the optimum. Households of type 2 who mimick those of type 1 will earn the same market income y_{c1} and obtain the same disposable income c_1. By the same analysis as above, $\widehat{z}_2 = z_1$, and so $t = 0$ in the optimum. This logic still applies if the self-selection constraint binds on type-1 households.

4.3 Different Preferences for Leisure

Household might differ not only by ability but according to their preferences for leisure. This adds another important and difficult dimension to redistributive policy. For one thing, governments are unlikely to be able to differentiate persons according to their preferences for leisure, that is, their laziness or diligence. For another, even if they could, it is not obvious how redistributive policies ought to differentiate among preference types. There is a school of thought that suggests that households are responsible for their own preferences, and redistributive policies ought only to compensate for ability differences.[10] On the other hand, as stressed by Cuff (2000), preferences for leisure might be viewed as being partly determined by not just one's attitude to work, but also to the degree of difficulty individuals face.

Consider the simple case in which there are two ability-types of households and two preference types. A convenient way to formulate the utility function when there are differences in preferences is as $u\left(g\left(x,z\right),\alpha\ell\right)$.[11] In this formulation, α can take on the values α_1 and α_2. If $\alpha_1 > \alpha_2$, preference type-1 has a greater preference for leisure than preference type-2 households. In the unlikely event that the government could distinguish between high and low preference for leisure types, it could simply design two separate non-linear income tax systems for the two types, exactly analogously to the case of different needs for goods considered earlier. In this case, it is obvious that the A-S Theorem applies, since within each preference type, the high-wage mimicking person would have the same disposable income as the low-wage person, and by separability would consume the same bundle of goods.

If the government cannot distinguish preference types, it is again faced with a two-dimensional screening problem. Depending on the relative welfare weights attached to the two preference types, the pattern of binding self-selection constraints can vary (cf. Boadway et al, 2001). However, regardless of what the pattern might be in the optimum, the separability of the utility function combined with the commonality of the sub-utility function $g(x, z)$ implies that the A-S Theorem still holds. Mimickers will have the same income and disposable income as those they are mimicking regardless of the type of either. Therefore, they will consume the same bundle of goods, implying that a differential commodity tax cannot be used to separate the two types. Differences in preferences for leisure merely serve to complicate the form of the optimal non-linear income tax.

Finally, note that differences in preference for leisure could reflect differences in need, analogous to the case of differences of need for different goods. For example, utility functions might take the form $u(g(x, z), \ell + a)$, where a reflects need and can vary from one household to another. By similar reasoning to above, the A-S Theorem continues to apply in this case regardless of whether the government can observe household needs.

4.4 Becker-Gronau Household Production

A final case to consider is the case where consumption of goods itself requires the allocation of some time, following Becker (1965), Gronau (1977) and Jacobsen Kleven (2000). One way of formulating the utility function in this case is as $u(g(X, Z), \ell)$, where X and Z are commodities produced by household production functions $f^x(x, \ell_x)$ and $f^z(z, \ell_z)$, where ℓ_x and ℓ_z are labor inputs into the production of the home-produced commodities. Assuming that the household production functions are the same for both households, the A-S Theorem applies directly. Mimicking households will have the same income and disposable incomes of those being mimicked, and given these assumptions will purchase the same quantities of the two goods x and z to produced the same quantities of household goods. On the other hand, if the households had different productivities in home production, which might be a reasonable assumption, the same disposable incomes would generally give rise to different demands for x and z by mimicking type-2's and type-1's.

5. Restricted Instruments: Linear Income Taxation

Up to now we have assumed that there was no restriction on the income tax schedule. Yet there are a number of reasons for not having an unrestricted non-linear income tax. To illustrate one type of restriction, we consider commodity taxation in a model where there is an optimal linear income tax. Our approach allows us to characterize clearly the restrictions on the utility function that will be sufficient to rule out differential commodity taxation. For this purpose, we revert to the basic model in which all consumers have identical utility functions, now specified to take the general form $u(x, z, \ell)$. We retain the commodity tax on z and now introduce a uniform lump-sum subsidy T and a constant marginal income tax m applicable to all households.

The household's problem is to maximize $u(x, z, \ell)$ subject to the budget constraint $x + qz = \omega\ell + T$, where $\omega = (1 - m)w$ is the after-tax wage rate. Maximizing $u(\omega\ell + T - qz, z, \ell)$ leads to the demand functions $z(\omega, q, T)$ and $\ell(\omega, q, T)$, and the indirect utility $v(\omega, q, T)$ with

$$v_T = u_x; \qquad v_\omega = u_x\ell \qquad \text{and} \qquad v_q = -u_x z$$

The government's revenue constraint is now simply:

$$\sum n_i[(w_i - \omega_i)\ell_i(\cdot) - T + (q-1)z_i(\cdot)] \geq 0$$

The government's optimal tax problem can be expressed by the Lagrangean:

$$\mathcal{L} = \sum n_i v^i(\omega_i, q, T) + \lambda \sum n_i[(w_i - \omega_i)\ell_i(\cdot) - T + (q-1)z_i(\cdot)]$$

We thus obtain the first-order conditions:

$$(m) \qquad -\sum n_i u_x^i \ell_i w_i + \lambda \sum n_i \left[w_i \ell_i m w_i^2 \frac{\partial \ell_i}{\partial \omega_i} + t \frac{\partial z_i}{\partial \omega_i} w_i \right] = 0$$

$$(T) \qquad \sum n_i u_x^i - \lambda \sum n_i \left[1 - m w_i \frac{\partial \ell_i}{\partial T} + t \frac{\partial z_i}{\partial T} \right] = 0$$

$$(q) \qquad -\sum n_i u_x^i z_i + \lambda \sum n_i \left[z_i + m w_i \frac{\partial \ell_i}{\partial q} + t \frac{\partial z_i}{\partial q} \right] = 0$$

These three first-order conditions can be combined by the following operations: $(T) \cdot \sum n_i w_i \ell_i + (m)$ and $(T) \cdot \sum n_i z_i + (m)$ to yield:

$$-\text{cov}(b_i, w_i, \ell_i) - \sum n_i \left[m w_i^2 \frac{\partial \tilde{\ell}_i}{\partial \omega_i} - t w_i \frac{\partial \tilde{z}_i}{\partial \omega_i} \right] = 0$$

$$-\text{cov}(b_i, z_i) + \sum n_i \left[m\, w_i \frac{\partial \tilde{\ell}_i}{\partial q} + t \frac{\partial \tilde{z}_i}{\partial q} \right] = 0$$

where b_i is the standard expression for the net social marginal utility of income:

$$b_i = \frac{u^i x}{\lambda} + \sum n_i \left(m w_i \frac{\partial \ell_i}{\partial T} - t \frac{\partial z_i}{\partial q} \right)$$

It is generally specified that b_i is negatively correlated with gross earnings $w_i \ell_i$.

From the first-order condition (T), $\sum n_i b_i = 1$. When t is equal to 0, the first equations yields:

$$m = \frac{-\text{cov}(b_i, w_i \ell_i)}{\sum n_i w_i^2 \frac{\partial \tilde{\ell}_i}{\partial \omega_i}}$$

This is the standard formula for the optimal linear income tax rate, with the equity term in the numerator and the efficiency term in the denominator.

More generally, it is clear from these two equations that if $\text{cov}(b_i, w_i \ell_i) + m \sum n_i w_i^2 \partial \tilde{\ell}_i / \partial \omega_i$ is proportional to $\text{cov}(b_i, z_i) + m \sum n_i w_i \partial \tilde{\ell}_i / \partial q$, then $t = 0$. It can be shown that this will be the case when goods are separable from leisure in the utility function, and the subutility function $g(x, z)$ is quasi-homothetic; that is, Engel curves relating goods consumption and disposable income are linear.[12]

Proof:

Note first that quasi-homotheticity makes the consumption of z a linear function of $w\ell$.

As well, quasi-homotheticity combined with separability make $w_i \partial \tilde{\ell}_i / \partial \omega_i$ proportional to $\partial \tilde{\ell}_i / \partial q$. To see this, consider the expenditure minimization problem of the household, letting p be the price of x:

$$\min_{x,z,\ell} \; px + qz - \omega \ell \qquad \text{s.t. } \; u(g(x,z), \ell) \succeq u$$

This problem's solution gives the compensated functions $\tilde{x}(p, q, \omega, u)$, $\tilde{z}(p, q, \omega, u)$ and $\tilde{\ell}(p, q, \omega, u)$, and the expenditure function $e(p, q, \omega, u)$. The envelope theorem then yields

$$e_p = \tilde{x}(p, q, \omega, u); \quad e_q = \tilde{z}(p, q, \omega, u); \quad e_\omega = -\tilde{\ell}(p, q, \omega, u)$$

Differentiating, we obtain:

$$e_{\omega p} = e_{p\omega} = -\frac{\partial \tilde{\ell}}{\partial p} = \frac{\partial \tilde{x}}{\partial \omega}; \quad e_{\omega q} = e_{q\omega} = -\frac{\partial \tilde{\ell}}{\partial q} = \frac{\partial \tilde{z}}{\partial \omega}; \quad e_{\omega\omega} = -\frac{\partial \tilde{\ell}}{\partial \omega}$$

Compensated demand functions, and therefore e_ω, are homogeneous of degree zero in prices. Applying Euler's Theorem, we have $pe_{\omega p} + qe_{\omega q} + \omega e_{\omega\omega} = 0$. Using the symmetry condition, this may be written:

$$p\frac{\partial \tilde{x}}{\partial \omega} + q\frac{\partial \tilde{z}}{\partial \omega} - \omega\frac{\partial \tilde{\ell}}{\partial \omega} = 0$$

Because of separability, preferences over $\{x, z\}$ are defined by the sub-utility function $g(x, z)$, which is independent of ℓ. A compensated increase in the wage rate causes labour supply to rise. In turn, disposable income rises, and the budget constraint in (x, z)–space moves parallel upward. By quasi-homotheticity, we have that for any change in dispsable income, $\Delta\tilde{x}/\Delta\tilde{z} = k$, where k is a constant that depends on relative prices and is the same for all households. Therefore, for a change in ω, we have $\partial\tilde{x}/\partial\omega = k\partial\tilde{z}/\partial\omega$. The Euler condition can then be written:

$$(pk + q)\frac{\partial \tilde{z}}{\partial w} - \omega\frac{\partial \tilde{\ell}}{\partial \omega} = 0 = -(pk + q)\frac{\partial \tilde{\ell}}{\partial q} - \omega\frac{\partial \tilde{\ell}}{\partial \omega}$$

So, normalizing $p = 1$, we have the proportionality we are seeking:

$$(k + q)\frac{\partial \tilde{\ell}}{\partial q} = -\omega\frac{\partial \tilde{\ell}}{\partial \omega}$$

Now, since $\omega\ell = x + qz - T$ by the budget constraint, $\partial(\omega\ell)/\partial z = k + q$. Therefore,

$$\left(\frac{\partial \omega\ell}{\partial z}\right)\frac{\partial \tilde{\ell}}{\partial q} = -\omega\frac{\partial \tilde{\ell}}{\partial \omega}$$

This implies that the factor of proportionality relating $w_i \partial\tilde{\ell}_i/\partial\omega_i$ to $\partial\tilde{\ell}_i/\partial q$ is the same as the one relating $\text{cov}(v_i, w_i, \ell_i)$ to $\text{cov}(b_i, z_i)$. Therefore, $t = 0$. ■

This result can be seen as an adaptation of the A-S Theorem to a setting in which there are restrictions on the structure of the income tax. By using the same logic as before, differences in need and endowments for goods will have the same effect on the applicability this modified A-S Theorem, as in the A-S Theorem under non-linear optimal income taxation. That is, if households have different unobservable needs for one of the goods, it will generally be desirable to impose a tax or subsidy on it. On the other hand, if needs are observable, different tax schedules will apply to persons of different needs classes. Similarly, if households supply two types of labor, the modified A-S Theorem applies if both types are observable, but not otherwise. As well, the modified version still applies if there is unobserved household production or if preferences for leisure differ.

6. Conclusions

When looking at real life tax systems one finds almost everywhere a mix of direct and indirect taxes, or more precisely of consumption and income taxes. What are the reasons for such an apparent violation of the A-S proposition? Ignorance of basic public economics and thus bad fiscal engineering? Huge compliance costs in income taxation relative to consumption taxation? Reasons developed in this paper and elsewhere for infirming the A-S proposition? Unwillingness to implement an optimal income tax? Lack of separability of the utility function?

As usual the answer is 'a bit of everything'. It is clear that in developing countries, the compliance and administration costs of income taxation are so high that tax authorities have to rely on friendlier indirect taxation. The arguments developed above have also some empirical relevancy. For example, it is natural to think that needs differ across individuals and are not always observable. The issue of separability is also far from being settled: most econometric studies do not lend support to such separability. It is possible that some public finance experts and policy makers miss the point of A-S proposition and believe that taxes are like eggs: you do not put them in the same basket as income taxation. Finally, there is an issue with the willingness to implement an optimal income tax. The A-S proposition assumes that one starts with such a tax. It is far from being granted that existing income tax systems correspond to such a scheme, and without optimal income taxation there is no AS proposition.

ENDNOTES

1. Cremer et al. (2001). See below Naito (1999) who shows that if production consists of several sectors using in variable proportion the different types of workers, then it pays to tax the sectors employing a relatively high proportion of skilled labor, Saez (2002) who shows that Naito's objection disappears in the long run, and Cremer and Gahvari (1995) who underline the desirable insurance effect of commodity taxation.

2. This assumption has been questioned on empirical grounds. See on this Christiansen (1984) and Browning and Meghir (1991).

3. Revesz (1986) has shown that if the government could levy license fees alongside proportional commodity taxes, it might be optimal to do so even if optimal proportional commodity tax rates are zero because of separability. In this paper, we assume that license fees cannot be enforced because of the possibility of resale.

4. A similar procedure has been used by Edwards et al (1994), Nava et al (1996), and Cremer et al (2001).

5. In what follows, variables applying to households of type i are denoted by a subscript, while functions for household i are denoted by a superscript. Function subscripts refer to partial derivatives.

6. The first-order conditions on incomes y_i can be used to characterize the structure of the optimal income tax. The characterization is standard, and we suppress it here. In what follows, the government is always taken to be applying

the optimal income tax.

7. Note that $\widehat{z}_2 > z_1$ implies that the tax on z should be positive (using the negativity of own substitution effects). This corresponds with the case in which z and leisure are complements: type-2 mimickers take more leisure than type-1's. See Edwards et al (1994) and Nava et al (1996).

8. An interpretation of unobserved endowments that gives rise to a rationale for differential commodity taxation is the case of bequests, analysed in Boadway et al (2000) and Cremer et al (2002). In this case, the analysis is intertemporal, and the differential taxation applies to future versus present consumption. These authors treat capital income taxation as the policy instrument for taxing future consumption.

9. Differences in needs or endowments are similar to heterogeneous tastes. See on this Saez (2000).

10. See, for example, Roemer (1998) and Fleurbaey and Maniquet (1999).

11. This is the formulation for the preferences for leisure used by Boadway et al (2002), who study the design of the optimal redistributive income tax when households differ in both ability and preferences.

12. On this, see Deaton (1979).

REFERENCES

Atkinson, Anthony B. and Joseph E. Stiglitz. "The Design of Tax Structure: Direct versus Indirect Taxation". *Journal of Public Economics* 6 (1976): 55–75.

Becker, Gary S. "A Theory of the Allocation of Time". *The Economic Journal* 75 (1965): 493–517.

Boadway, Robin, Maurice Marchand and Pierre Pestieau. "Towards a Theory of Direct-Indirect Tax Mix". *Journal of Public Economics* 55 (1994): 71–88.

Boadway, Robin, Maurice Marchand and Pierre Pestieau. "Redistribution with Unobservable Bequests: A Case for Taxing Capital Income". *Scandinavian Journal of Economics* 102 (2000): 253–67

Boadway, Robin, Maurice Marchand, Pierre Pestieau and Maria del Mar Racionero. "Optimal Redistribution with Heterogeneous Preferences for Leisure". *Journal of Public Economic Theory* (2002), forthcoming.

Browning, Martin and Costas Meghir. "The Effects of Male and Female Labor Supply on Commodity Demands". *Econometrica* 59 (1991): 925–951.

Corlett, W.J. and D.C. Hague. "Complementarity and the Excess Burden of Taxation". *Review of Economic Studies* 21 (1953–54): 21–30.

Cremer, Helmuth and Firouz Gahvari. "Uncertainty and optimal taxation: in defense of commodity taxes". *Journal of Public Economics* 56 (1995): 291–310.

Cremer, Helmuth, Pierre Pestieau and Jean-Charles Rochet. "Capital income taxation when inherited wealth is not observable". *Journal of Public Economics* (2002), forthcoming.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet. "Direct Versus Indirect Taxation: The Design of the Tax Structure Revisited". *International Economic Review* 42 (2001): 781–99.

Cuff, Katherine. "Optimality of Workfare with Heterogeneous Preferences". *Canadial Journal of Economics* 33 (2000): 149–74.

Christiansen, Vidar. "Which Commodity Taxes Should Supplement the Income Tax?". *Journal of Public Economics* 24 (1984): 195–220.

Deaton, Angus. "Optimally Uniform Commodity Taxes". *Economics Letters* 2 (1979): 357–361.

Edwards, Jeremy, Michael Keen and Matti Tuomala. "Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide". *Finanzarchiv* 51 (1994): 472–97.

Fleurbaey, Marc and Francois Maniquet. "Compensation and Responsibility". (1999), mimeo.

Gronau, Ruben. "Leisure, Home Production and Work — The Theory of the Allocation of Time Revisited". *Journal of Political Economy* 85 (1977): 1099–123.

Guesnerie, Roger. *A Contribution to the Pure Theory of Taxation.* Cambridge, UK: Cambridge University Press, 1995.

Jacobsen Kleven, Henrik. “Optimal Taxation and the Allocation of Time”. (2000), mimeo.

Mirrlees, James A. “An Exploration in the Theory of Optimal Income Taxation”. *Review of Economic Studies* 38 (1971): 175–208.

Naito, Hisahiro. “Re-examination of Uniform Commodity Taxes under a Nonlinear Income Tax System and its Implication for Productive Efficiency”. *Journal of Public Economics* 71 (1999): 165–188.

Nava, Mario, Fred Schroyen and Maurice Marchand. “Optimal Fiscal and Public Expenditure Policy in a Two-Class Economy”. *Journal of Public Economics* 61 (1996): 119–37.

Revesz, John T. “On Some Advantages of Progressive Indirect Taxation.” *Public Finance* XXXXI (1986): 182–99.

Roemer, John E. *Equality of Opportunity.* Cambridge, US: Harvard University Press, 1998.

Rowe, Nicholas and Frances Woolley. “The Efficiency Case for Universality”. *Canadian Journal of Economics* 32 (1999): 613–29.

Sandmo, Agnar. “Optimal Taxation: An Introduction to the Literature”. *Journal of Public Economics* 6 (1976): 37–54.

Saez, Emmanuel. “The Desirability of Commodity Taxation under Nonlinear Income Taxation and Heterogeneous Tastes”. NBER Working Paper (2000), *Journal of Public Economics*, forthcoming .

Saez, Emmanuel. “Direct or Indirect Tax Instruments for Redistribution: Short-run versus Long-run”. NBER Working Paper, (2002).

Stiglitz, Joseph E. “Self-Selection and Pareto Efficient Taxation”. *Journal of Public Economics* 17 (1982): 213–40.

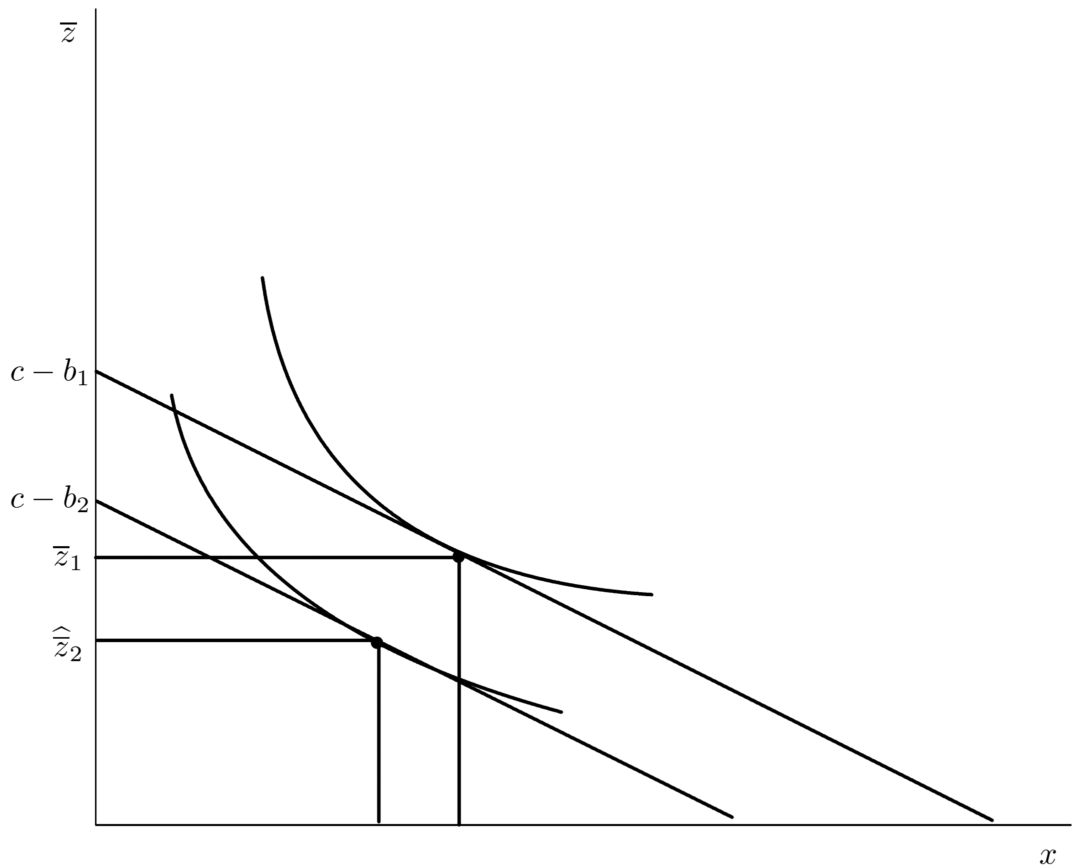

Figure 1