



ELSEVIER

Journal of Public Economics 83 (2002) 217–230

JOURNAL OF
PUBLIC
ECONOMICS

www.elsevier.com/locate/econbase

The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes

Emmanuel Saez*

Harvard University and NBER, Department of Economics, Littauer Center, Cambridge, MA 02138, USA

Received 15 May 2000; received in revised form 14 November 2000; accepted 20 November 2000

Abstract

This paper revisits the Atkinson–Stiglitz result on uselessness of commodity taxation in the presence of optimal non-linear income taxation in a more general setup, namely when tastes are heterogeneous. This general analysis displays the key economic assumptions under which the Atkinson–Stiglitz result is robust. A small tax on a given commodity is desirable if high income earners have a relatively higher taste for this commodity or if consumption of this commodity increases with leisure. An application to the case of savings suggests that, even in the presence of optimal non-linear earnings taxation, there is a role for a supplemental capital income tax in the standard overlapping generation model. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Optimal nonlinear income tax; Commodity taxation

JEL classification: H21; H23

1. Introduction

Taxation is a key instrument in the hands of the government to redistribute among individuals. Assessing the power of differential commodity taxation versus nonlinear income taxation for redistribution is a central tax policy question which has attracted much attention in the literature on optimal taxation. The role of

*Corresponding author. Tel.: +1-617-495-0586; fax: +1-617-495-8570.

E-mail address: saez@fas.harvard.edu (E. Saez).

differential commodity taxation has been severely undermined by the influential paper of Atkinson and Stiglitz (1976). They showed that, under a condition of separability of leisure and consumption choices, optimal non-linear income taxation makes commodity taxation useless. This result, applied to dynamic models, has provided a strong theoretical argument against the use of capital income taxation in the presence of nonlinear taxation of wage income.

Atkinson and Stiglitz derived their result in a way that made economic interpretation difficult. A number of studies by Mirrlees (1976), Christiansen (1984), and Konishi (1995), have tried to understand in more depth the Atkinson–Stiglitz result. Mirrlees (1976) showed that commodity taxation is desirable on goods that are relatively more preferred by the high skilled individuals. Christiansen (1984) showed that goods that are complementary with leisure should be taxed. However, all these studies have considered the Atkinson–Stiglitz result in a context of strong homogeneity of preferences for consumption goods. Namely, they considered models where all individuals would buy exactly the same bundle of goods when provided with the same amount of disposable income.

This strong homogeneity in tastes for goods is clearly unrealistic and the goal of this paper is to investigate how the Atkinson–Stiglitz theorem can be adapted to the case of heterogeneous consumption preferences. This improves previous findings on two grounds. First, considering the general case of heterogeneous tastes displays the conditions under which the Atkinson–Stiglitz result is robust. Second, the stringent homogeneity assumptions considered by previous studies obscure the economic mechanism behind the Atkinson–Stiglitz result. Considering the general case clarifies the key economic assumptions necessary to obtain the result and allows perhaps a more transparent interpretation.

A number of studies have examined optimal income and commodity taxation in the context of two skill types and have obtained a simpler economic interpretation of the Atkinson–Stiglitz result in that particular case. Stiglitz (1982) showed that when leisure and goods are separable, differential taxation of commodities cannot be used as a basis of separation of the two types and is thus sub-optimal. Naito (1999) obtained a similar interpretation and showed that the Atkinson–Stiglitz result breaks down when wages and prices are endogenous.

Recently, Cremer et al. (2000) have investigated the role of commodity taxation in a discrete type model with optimal income taxation. Their important innovation is to consider a situation where individuals also differ along their initial endowment. This second source of heterogeneity can be seen as a first step toward the general case considered in the present paper. Cremer et al. (2000) note that, in their model, separability is no longer enough to obtain the Atkinson–Stiglitz result. They do not, however, investigate the conditions necessary to restore the Atkinson–Stiglitz result which is the focus of the present paper. They center instead their analysis on the size of optimal commodity taxes. Their study can therefore be considered as complementary to the present paper.

The present paper proceeds as follows. Section 2 presents the model. Section 3

investigates and discusses the conditions under which no commodity taxation is desirable. The present method of analysis generalizes the original method of Christiansen (1984). A brief application to the case of the taxation of savings is presented in Section 4. Section 5 offers concluding remarks.

2. The model

Individuals are indexed by h belonging to a possibly multi-dimensional index set \mathcal{H} . To simplify notation, I use the symbol \sum_h to denote summation over all individuals even though \mathcal{H} should be thought as a continuum. The total number of individuals is normalized to one. There are K consumption goods and labor supply. Individual utility is denoted by $U^h(c, z)$ where $c = (c_1, \dots, c_K)$ is the vector of consumption goods and z is earnings.¹ I denote by p the vector of before tax prices of goods. The government sets a non-linear income tax on earnings $T(z)$. I denote by t the vector of tax rates on goods and $q = p + t$ the after tax commodity prices. Individuals maximize utility $U^h(c, z)$ subject to the budget constraint $q \cdot c \leq z - T(z)$. The individual choice of consumption goods on the one hand and the labor supply choice on the other hand play a key role in the problem we investigate. Therefore, as in Christiansen (1984), I decompose this maximization into two parts: first, the choice of consumption goods c and second, the labor choice z .

Consumption choice

Assume that a given individual h supplies labor in order to earn a given income level z and is given an amount R to spend on the K goods. That individual would choose c to maximize $U^h(c, z)$ subject to $q \cdot c \leq R$. I denote by $v^h(q, R, z)$ the corresponding indirect utility function and $c_k^h(q, R, z)$ the demand for good k for a labor supply level generating earnings z , and a disposable income level R .

Labor choice

In fact, disposable income R and earnings z are related through the income tax by the equation $R = z - T(z)$. Individual h chooses z so as to maximize $v^h(q, z - T(z), z)$. The optimal earnings level z^h is a function of q and the tax schedule $T(\cdot)$.

Assume that the government does not use commodity taxation ($p = q$) and implements an income tax $T(\cdot)$ so as to maximize a weighted sum of individual utilities $W = \sum_h \alpha^h v^h(p, z^h - T(z^h), z^h)$ subject to $\sum_h T(z^h) \geq E$ where E is an

¹In the original formulation of the optimal non-linear income tax problem by Mirrlees (1971), h is scalar and positive and $U^h(c, z) = u(c, z/h)$. z/h is labor supply and h is the wage rate.

exogenous revenue requirement for the government and subject to the fact that each individual chooses z^h to maximize utility. I denote by λ the multiplier of the government budget constraint. λ is thus equal to the marginal value of public funds. I denote by $g^h = \alpha^h v_R^h / \lambda$ the social marginal weight on individual h . These weights summarize the redistributive tastes of the government. At the optimum, the government is indifferent between giving one more dollar to individual h or getting g^h more dollars of public funds.

The following analysis relies critically on small variations around the optimal tax schedule. Therefore, I assume that the number of individuals is large enough so that average consumption demands by earnings levels are smooth functions of earnings.

More importantly, in order to apply standard calculus of variations, it is necessary to assume that the optimal income tax schedule is regular and that there is no bunching nor gaps in the optimal schedule. Without this assumption, a small variation in the income tax could produce discrete jumps in quantities consumed or supplied and the first order analysis would be invalidated. Christiansen (1984) and Konishi (1995) also made their analysis assuming smooth income tax schedules. The conditions necessary to obtain a smooth income tax schedule in the one-dimensional skill distribution model they consider has been extensively analyzed (see notably Mirrlees, 1976; Seade, 1977). The case of multi-dimensional skill characteristics considered in this paper has been much less studied. The small literature on multi-dimensional screening models has shown that it is difficult to obtain general assumptions insuring that the optimal tax schedule is smooth (see for example the analysis of Rochet and Choné, 1998). The present paper will nevertheless ignore completely this issue and make the strong assumption of a smooth schedule.

3. When does commodity taxation improve welfare?

I assume from now on that the income tax schedule $T(\cdot)$ is optimal and smooth as described previously. In order to see whether a commodity taxation improves welfare, I consider, as in Christiansen (1984), the introduction of a small tax dt_1 on (say) commodity 1. This small tax reform has three effects on welfare and tax revenue.

First, the tax raises mechanically $dM_1 = \sum_h c_1^h dt_1 = C_1 dt_1$ additional taxes where C_1 denotes aggregate consumption of good 1.

Second, the tax has a negative welfare effect on individuals consuming good 1. Using Roy's identity, the welfare effect on individual h is equal to $v_{q_1}^h dt_1 = -v_R^h c_1^h dt_1$. The total welfare effect expressed in terms of the value of public funds is

$$dU_1 = -\frac{1}{\lambda} \sum_h \alpha^h v_R^h c_1^h dt_1 = -\sum_h g^h c_1^h dt_1. \quad (1)$$

Last, changing price q_1 produces a behavioral labor supply response $dz_{t_1}^h = dt_1 \partial z^h / \partial q_1$ and therefore changes the amount of taxes collected by the income tax by

$$dB_1 = -\sum_h T'(z^h) dz_{t_1}^h. \tag{2}$$

The total welfare effect of the small commodity tax dt_1 is $dW/dt_1 = dM_1/dt_1 + dU_1/dt_1 + dB_1/dt_1$. When dW/dt_1 is not equal to zero, imposing a small tax (or subsidy) on good 1 is desirable.

To check whether $dW/dt_1 = 0$, I use, as in Christiansen (1984), the fact that any small income tax reform has no first order effect on welfare because the income tax is optimal. The natural small income tax reform dT that mimics the small commodity tax is such that $dT(z) = C_1(z) dt_1$ for every z , where $C_1(z)$ denotes average consumption of good 1 for individuals earning z . As above, the effects of this tax change can be decomposed into mechanical, welfare and behavioral effects.

The mechanical effect due to the income tax change is equal to $dM_T = \sum_h dT(z^h) = \sum_h C_1(z^h) dt_1 = C_1 dt_1$. The commodity tax change dt_1 and the income tax change dT induce the same mechanical effect on tax revenue: $dM_T = dM_1$.

The welfare effect for individual h is equal to $-v_R^h dT(z^h) = -v_R^h C_1(z^h) dt_1$. The total welfare effect due to the income tax change is therefore

$$dU_T = -\frac{1}{\lambda} \sum_h \alpha^h v_R^h C_1(z^h) dt_1 = -\sum_h g^h C_1(z^h) dt_1. \tag{3}$$

I denote by dz_T^h the behavioral response of earnings due to the income tax change. Thus the total loss in tax revenue due to the behavioral response to the income tax change is

$$dB_T = -\sum_h T'(z^h) dz_T^h. \tag{4}$$

Because the income tax is optimal, the small income tax change has no first order effect on welfare: $dM_T/dt_1 + dU_T/dt_1 + dB_T/dt_1 = 0$. Therefore, the total welfare effect dW/dt_1 of the small commodity tax can be rewritten as $dW_1/dt_1 = dU_1/dt_1 - dU_T/dt_1 + dB_1/dt_1 - dB_T/dt_1$. Using Eqs. (1) to (4)

$$\frac{dW}{dt_1} = -\sum_h g^h [c_1^h - C_1(z^h)] + \sum_h T'(z^h) \left[\frac{dz_T^h}{dT(z^h)} \cdot \frac{dT(z^h)}{dt_1} - \frac{dz_{t_1}^h}{dt_1} \right]. \tag{5}$$

The total welfare effect on the commodity tax is decomposed into two terms. The first term is the pure welfare effect and the second term is the behavioral effect. Commodity taxation can improve welfare if either of these two terms is non zero.

Pure welfare effect

As $C_1(z)$ is average consumption of good 1 for individuals earning $z^h = z$, the pure welfare term in (5) is zero if conditional on any given z , g^h and c_1^h are uncorrelated. This is the case when, conditional on income, the government does not systematically give higher (or lower) social weights to individuals with higher tastes for good 1.

Assumption 1. Conditional on any income level z , g^h and c_1^h are uncorrelated.

Under what conditions can the government not set the same welfare weights on individuals with the same income? First, the government might want to advantage some classes of individuals over others, such as the young at the expense of the old. If these different groups do not have the same consumption pattern then g^h and c_1^h might be correlated conditional on income. However, one might think that, in a liberal society, the government should not set judgment values on the citizens based on their consumption. Second, if the good is consumed mostly by a particular class of individuals that are disadvantaged even conditioning on income, then g^h and c_1^h might be positively correlated. An example of such a good could be medical expenses. People with high medical expenses have less income left for consumption of other goods and thus could be given higher weights than healthy people with the same total income but with no medical expenses.²

In the model considered by Atkinson and Stiglitz (1976) and by subsequent studies, there is only one individual at each income level and thus no heterogeneity within income levels. As a result, Assumption 1 is always satisfied. Note finally that g^h and c_1^h are endogenous parameters and thus it would be difficult to specify general primitive conditions on intrinsic parameters insuring that Assumption 1 is true at the optimum. Nevertheless, if we want to model a government that does not want to discriminate between different consumption patterns, it seems reasonable to assume that the primitive conditions on utilities and social weights have been specified so that Assumption 1 is true at the optimum.³

Behavioral effect

The second reason why commodity taxation might be desirable is when the behavioral term in (5) is non zero. An important result in optimal income taxation

²Deaton and Stern (1986), in the context of optimal linear taxation, discuss the same issue from a different perspective.

³In other words, one should not put too much stress on the intrinsic parameters of the model. We should rather assume that these parameters are chosen so that the social weights at the optimum satisfy the conditions one wants to impose on government preferences such as neutrality with respect to consumption patterns as in Assumption 1.

states that, under weak assumptions, the optimal marginal tax rate is non-negative (that is, $T'(z) \geq 0$ for all z). Mirrlees (1971) presented this result and Seade (1982) clarified the conditions under which it is valid. This result is valid when leisure is a non-inferior good and when the government has redistributive tastes. We assume here that $T'(z) \geq 0$ for every z . Therefore, to sign the behavioral term, we must compare the behavioral responses $dz_{t_1}^h$ and dz_T^h induced respectively by the commodity tax dt_1 and the small income tax change dT . Suppose that $dt_1 > 0$, if the reduction of labor supply is smaller with the commodity tax dt_1 than with the mimicking income tax change dT (that is, $dz_{t_1}^h > dz_T^h$), then commodity taxation is more efficient than the income tax and is thus desirable (granted Assumption 1 holds). Note that in this case, desirability of commodity taxation is unrelated to the redistributive tastes of the government and is uniquely due to the difference in labor responses to commodity versus income taxation.

Let us compute successively the change in earnings induced by dt_1 and by $dT(z) = C_1(z) dt_1$. To obtain the change in earnings due to the commodity tax dt_1 , let us prove first the following result,

Lemma 1. *The change in earnings $dz_{t_1}^h$ for individual h induced by the small commodity tax dt_1 is equal to the change in earnings induced by a small income tax reform specific to individual h equal to $dT^h(z) = c_1^h(p, z - T(z), z) dt_1$.*

Proof. Let us define $dT^h(z)$ to be such that for any z , $v^h(q + dt_1, z - T(z), z) = v^h(q, z - T(z) - dT^h(z), z)$. Because these two functions are identical for all z , it must be the case that the value of z maximizing these two functions is identical. By assumption, $z + dz_{t_1}$ maximizes $v^h(q + dt_1, z - T(z), z)$. Therefore, it follows that $z + dz_{t_1}$ maximizes $v^h(q, z - T(z) - dT^h(z), z)$. In other words, dz_{t_1} is also the earnings response to the change $dT^h(z)$.

Applying a first order Taylor expansion for each of these two functions around $(q, z - T(z), z)$, and using Roy's identity, $v_{q_1}^h dt_1 = -v_R^h c_1^h dt_1$, one obtains $dT^h(z) = c_1^h(p, z - T(z), z) dt_1$. Therefore, dt_1 has the same effect on labor supply as a small individual h specific change in the income tax $dT^h(z) = c_1^h(p, z - T(z), z) dt_1$. □

In general, the functions of z , $C_1(z)$ and $c_1^h(p, z - T(z), z)$ are different implying that the behavioral responses induced by the changes $dT(z)$ and $dT^h(z)$ (or equivalently dt_1 by Lemma 1) are also different. I denote by $D_z C_1$ and $D_z c_1^h$ the (total) derivatives with respect to z of the two functions of z , $C_1(z)$ and $c_1^h(p, z - T(z), z)$.

A change in the tax schedule affects earnings z through income and substitution effects. For a given individual, I denote by $z^h(q, 1 - \tau, R)$ the earnings level he would supply when facing prices q and a linear budget constraint with tax rate τ and virtual income R . I denote by $z_R^h = \partial z^h / \partial R$ and $z_{1-\tau}^h = \partial z^h / \partial(1 - \tau)$ the income

and price effects. Using the Slutsky equation, I denote by $z_c^h = \partial z^h / \partial(1 - \tau) - z^h \partial z^h / \partial R$ the substitution effect. With the non-linear income tax in place, we have $\tau = T'(z^h)$ and $R = z^h - T(z^h) - z^h[1 - T'(z^h)]$.

The following result allows us to compare the labor supply responses to $dT(z)$ and $dT^h(z)$,

Lemma 2. *The response of earnings dz^h of individual h due to an arbitrary small change in the tax schedule $P(z)$ dt is such that,*

$$dz^h = -z_c^h[P'(z^h) dt + T''(z^h) dz^h] - z_R^h P(z^h) dt. \tag{6}$$

Proof. An arbitrary small income tax change $P(z) dt$ produces a behavioral response dz^h in earnings due to income and substitution effects: $dz^h = -z_{1-\tau}^h d\tau + z_R^h dR = -z_c^h d\tau + z_R^h(dR - z^h d\tau)$. The change in marginal tax rate $d\tau$ is equal to $-P'(z^h)dt + T''(z^h)dz^h$ and routine computation shows that the income shock $dR - z^h d\tau$ is equal to $-P(z^h) dt$. Hence, Eq. (6) is obtained. \square

Applying Lemma 2 to the income tax changes $dT^h(z) = c_1^h dt_1$ and $dT(z) = C_1(z)dt_1$, we obtain the following expressions for $dz_{t_1}^h$ and dz_T^h ,

$$dz_{t_1}^h = -z_c^h[D_z c_1^h dt_1 + T''(z^h) dz_{t_1}^h] - z_R^h c_1^h dt_1, \tag{7}$$

$$dz_T^h = -z_c^h[D_z C_1 dt_1 + T''(z^h) dz_T^h] - z_R^h C_1(z) dt_1. \tag{8}$$

We want to sum these two equations over all individuals with income z . Let $E[\cdot]$ denote expectation over all individuals h with income $z^h = z$. Eqs. (7) and (8) imply,⁴

$$E[dz_{t_1}^h] = -dt_1 \left\{ E \left[\frac{z_c^h}{1 + T''(z)z_c^h} D_z c_1^h \right] + E \left[\frac{z_R^h}{1 + T''(z)z_c^h} c_1^h \right] \right\}, \tag{9}$$

$$E[dz_T^h] = -dt_1 \left\{ E \left[\frac{z_c^h}{1 + T''(z)z_c^h} D_z C_1(z) \right] + E \left[\frac{z_R^h}{1 + T''(z)z_c^h} C_1(z) \right] \right\}. \tag{10}$$

Granted Assumption 1 holds, commodity taxation is useless if for any z , $E[dz_{t_1}^h] = E[dz_T^h]$. Comparing Eqs. (9) and (10), two conditions are enough to obtain this equality,

⁴ $1 + T''(z^h)z_c^h > 0$ is equivalent to the second order condition for the individual maximization problem being satisfied. That is, the curvature of the indifference curve at the optimum labor supply choice is greater than the curvature of the budget constraint. Saez (2000) shows that this assumption is satisfied everywhere when the optimum income tax schedule is smooth which is the key assumption we made earlier on.

Assumption 2. Conditional on each income level z , behavioral responses z_c^h and z_R^h are independent of consumption patterns c_1^h and $D_z c_1^h$.

This assumption ensures that the expressions c_1^h and $D_z c_1^h$ in (9) can be averaged and pulled out of the expectation operator. Intuitively, it is clear that if conditional on income, labor supply responses are highest for those with no taste for good 1, then taxation of good 1 does not affect these highly responsive individuals and is more efficient than the indiscriminate income tax change $dT(z)$.

Assumption 2 is obviously satisfied in the usual case considered by previous studies because they consider models with a single individual at each income level. This assumption does not hold in the general heterogeneous case. However, for most goods, there are no reasons to think that conditional on income, consumption patterns should be related systematically to substitution or income effects parameters. In any case, checking this condition empirically is difficult and thus it seems reasonable to assume that this condition holds.

Assumption 3. For any income level z , $E[D_z c_1^h | z^h = z] = D_z C_1(z)$.

This is the key assumption needed for the Atkinson–Stiglitz result. When Assumption 2 holds and as, by definition, $E[c_1^h | z^h = z] = C_1(z)$, it is clear that Assumption 3 is enough to ensure that $E[dc_{t_1}^h] = E[dc_T^h]$. In order to understand Assumption 3, it is useful to note that,

$$D_z C_1(z) = \lim_{dz \rightarrow 0} \frac{E[c_1^h | z^h = z + dz] - E[c_1^h | z^h = z]}{dz}.$$

That is, $D_z C_1(z)$ captures the *cross-sectional* variation in consumption of good 1 when income increases whereas $E[D_z c_1^h | z^h = z]$ captures the *individual* variation in consumption of good 1 when earnings supply and disposable income change.

Thus Assumption 3 can be restated as follows. Consider all individuals h earning $z^h = z$ and call this group A . By definition, group A consumes on average a quantity $C_1(z)$ of good 1. Consider then all individuals h with income $z^h = z - dz$ and call them group B . Group B has less income than group A and consumes on average $dc_1 = D_z C_1(z) dz$ less good 1 than group A . Suppose you force group A individuals to decrease their labor supply so that their earnings fall to $z - dz$. I call group A' these individuals, group A' has on average the same earnings (namely $z - dz$) and same disposable income as group B . Compared to group A , group A' individuals reduce on average their consumption of good 1 by $dc'_1 = E[D_z c_1^h | z^h = z] dz$.

Assumption 3 states precisely that group A' and group B should consume on average the same quantity of good 1, that is, $dc_1 = dc'_1$. In the standard Atkinson–Stiglitz model that has been studied in the literature, each individual is characterized by a scalar skill level h . Individual utility is of the form $U(v(c_1, \dots, c_K), z, h)$ where the function $v(\cdot)$ is common to all individuals. The separability assumption

means that, given a level of after tax income R , individuals maximize $v(c)$ subject to $q \cdot c \leq R$. Therefore, consumption of the goods depends only on after tax income $R = z - T(z)$ available for consumption and not on the labor choice per se. As a result, the functions $c_1^h(z)$ and $C_1(z)$ are identical.

Group A' and group B have the same disposable income but individuals in group A' have been forced to reduce their labor supply. Consumption of good 1 can differ across these two groups for two reasons.

First, group A' is different from group B because they have higher incomes when they can freely choose their labor supply. If higher income earner individuals have relatively lower tastes for good 1, independently of disposable income, then group B consumes more good 1 than group A' . An example of such a good could be cigarettes because, higher incomes tend to smoke less⁵ and this clearly cannot be due to the mechanical fact that they have higher disposable income.

Second, group A' is different because it has been forced to reduce earnings and thus presumably enjoys more leisure than group B .⁶ As a result, if consumption of good 1 increases when leisure time increases, as could be the case for holiday travels for example, then group A' would consume more of this good than group B .

The first interpretation has been put forward by Mirrlees (1976) and the second one by Christiansen (1984). However, these studies used the Atkinson–Stiglitz framework and did not use the expression in Assumption 3.

When group A' consumes more good 1 than group B (i.e. when $E[D_z c_1^h | z^h = z] < D_z C_1(z)$), then using Eqs. (9) and (10) and the fact that the substitution effect z_c^h is positive, we see that for $dt_1 > 0$, $E[dz_T^h] < E[dz_{t_1}^h]$. This is the case when high income earners have a relatively stronger taste for good 1 than lower income earners or when good 1 consumption increases when leisure time increases. In that case, the ‘equivalent’ marginal rate effect of the commodity tax dt_1 given by $D_z c_1^h dt_1$ is smaller on average than the marginal effect of the income tax change $dT(z)$ given by $D_z C_1(z) dt_1$. This happens when the individual consumption demand $c_1^h(z)$ is flatter on average than the aggregate consumption demand $C_1(z)$ by income level. In that case, the distortionary effect of the commodity tax is lower than the income tax change. Commodity taxation then reduces earnings less and thus is more efficient than the small income tax change $dT(z)$. When Assumptions 1 and 2 hold, this implies that taxation of good 1 is desirable.

The shape of $C_1(z)$ can be easily measured empirically using a cross-sectional survey of household consumption. Measuring the average individual consumption patterns $c_1^h(q, R = z - T(z), z)$ is more difficult because one has to disentangle the

⁵See e.g., Evans et al. (1999) for empirical evidence of this fact in the U.S.

⁶This is of course true in the standard model where individuals differ only through their marginal productivity of labor. In general, this is expected to be true if earnings are correlated with skills. Note that this would not be true in a model where all individuals had the same wage rate or skill and differed only through their taste for leisure.

income effect $R = z - T(z)$ from the leisure effect z . The income effect might be measured using longitudinal data but it is harder to see how to estimate the leisure effect. We can now state the generalized Atkinson–Stiglitz result,

Proposition 1.

- If Assumptions 1,2 and 3 are satisfied for good 1, then commodity taxation of good 1 cannot improve welfare.
- If Assumptions 2 and 3 are satisfied but not Assumption 1, then taxation (resp. subsidization) of good 1 is desirable if marginal welfare weights g^h are negatively (resp. positively) correlated to consumption c_1^h of good 1 at each income level.
- If Assumptions 1 and 2 are satisfied but not Assumption 3, then taxation of good 1 is desirable if $E[D_z c_1^h | z^h = z] < D_z C_1(z)$ for all z (i.e. when high income earners have relatively more taste for good 1 or when good 1 consumption increases with leisure). Subsidization of good 1 is desirable if $E[D_z c_1^h | z^h = z] > D_z C_1(z)$ for all z .

4. Application to the taxation of savings

The most fruitful and influential application of the Atkinson–Stiglitz theorem has been in the case where different goods are interpreted as consumption at different dates. When consumption and labor enter the utility function in a separable way (that is, $u^h = u(c_1, \dots, c_T), z, h$), then there is no need to supplement a non-linear tax on earnings with a tax on interest income. This point was originally made by Atkinson and Stiglitz (1976), and developed in Ordober and Phelps (1979). It has been a major theoretical argument against the use of capital income tax.

To simplify the discussion, let us consider the standard two period model where individuals work only in period 1 and live out of their savings in period 2. Utility functions are defined by $u(c_1, c_2, z)$. Following the long tradition in macroeconomics, it tempting to specify a separable utility function such as,

$$u^h(c_1, c_2, z) = u(c_1) + \frac{u(c_2)}{1 + \delta} - v(z/h), \quad (11)$$

where δ is the discount rate and $v(\cdot)$ is a function capturing the disutility of effort. Specification (11) implies that commodity taxation (or equivalently interest income taxation) is useless in the presence of an optimal income tax. However, the key assumption embodied in specification (11) is not so much the separability assumption but rather that δ (as well as $u(\cdot)$ and $v(\cdot)$) are common to all individuals, that is, savings behavior is the same for every individual and independent of skills.

Using the framework of the present paper, in the presence of optimal nonlinear earnings taxation, the desirability of taxing interest income (or equivalently, taxing consumption in period 2) hinges on whether Assumption 3 is true.⁷ However, propensities to save vary widely across the population and empirical studies have shown that savings rates are correlated with education even controlling for income (see for example, Lawrance, 1991). Therefore, there is a strong presumption that higher income individuals save more not only because they have more income to save but also because they might have a better financial education and be more aware of the need to save for retirement. As a result, it is plausible to think that if individuals with high earnings z levels are forced to work less and thus earn $z' < z$ they would still have a higher taste for savings and thus save more than individuals with income z' . In terms of specification (11), the discount rate δ is probably negatively correlated with skills. This suggests that interest income ought to be *taxed* even in the presence of a non-linear optimal earnings tax. Trying to quantify the optimal tax rate on interest income using data on tastes for savings by skill or income levels is an important but difficult task left for future research.

5. Conclusion

The key element to assess the desirability of commodity taxation in the presence of optimal income taxation is whether a small commodity tax can be replicated by a small income tax change. When this is not the case, commodity taxation is a tool that allows the government to expand its taxation power and is therefore desirable. This can happen either when the government uses social weights correlated with consumption patterns even conditioning on income or when patterns of consumption are related to intrinsic earning power or leisure choices. In that latter case, the desirability of commodity taxation hinges not on whether the commodity is consumed disproportionately by high earners but rather on whether individual demand for that good is less elastic than the cross-sectional demand pattern.

The analysis shows that the separability property might be misleading because what is central in the Atkinson–Stiglitz assumption is not so much that consumption and labor enter the individual utilities separately but rather the fact that all individuals share the same subutility of consumption. In contrast, separability does not intervene in the analysis of the general heterogeneous case and the present analysis shows directly the key economic assumptions needed. As an important caveat, it should be noted again that the present analysis is valid only when the optimal income tax schedule is assumed to be smooth. Singularities in the optimal

⁷There is no reason to think that Assumptions 1 and 2 should not be approximately true for consumption at each date.

income tax might change the present conclusions, though it is not clear whether results based on singularities could be of much practical relevance.⁸

It would of course be extremely useful to obtain optimal commodity tax formulas when the assumptions insuring the Atkinson–Stiglitz result are not satisfied. Cremer et al. (2000) take an important step in that direction in the context of a discrete type model with a simple structure. Obtaining such formulas expressed in terms of empirically estimable magnitudes in the general model considered here is an important task left for future research.

Acknowledgements

I thank Roger Gordon, David Spector, and two anonymous referees for very helpful comments and discussions.

References

- Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. *Journal of Public Economics* 6, 55–75.
- Christiansen, V.A., 1984. Which commodity taxes should supplement the income tax? *Journal of Public Economics* 24, 195–220.
- Cremer, H., Pestieau, P., Rochet, J.-C., 2000. Direct versus indirect taxation: the design of the tax structure revisited. Forthcoming, *International Economic Review*.
- Deaton, A., Stern, N., 1986. Optimally uniform commodity taxes, taste differences and lump-sum grants. *Economic Letters* 20, 263–266.
- Evans, W., Ringel, J.S., Stech, D., 1999. Tobacco taxes and public policy to discourage smoking. In: Poterba, J. (Ed.), *Tax Policy and the Economy*, Vol. 13. MIT Press, Cambridge.
- Konishi, H., 1995. A Pareto-improving commodity tax reform under a smooth non-linear income tax. *Journal of Public Economics* 56, 413–446.
- Lawrance, E.C., 1991. Poverty and the rate of time preference: evidence from panel data. *Journal of Political Economy* 99 (1), 54–77.
- 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.
- Naito, H., 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics* 71, 165–188.
- Ordover, J., Phelps, E., 1979. The concept of optimal taxation in the overlapping-generations model of capital and wealth. *Journal of Public Economics* 12, 1–26.

⁸Cremer et al. (2000) note in their multi-type discrete model that determining the set of binding incentive compatibility constraints is difficult and affects in a complex and interesting way the economic interpretation of optimal taxes. Whether this characteristic of discrete models and its relevance for optimal tax analysis goes away in the continuum case considered here is not yet understood.

- Rochet, J.-C., Choné, P., 1998. Ironing, sweeping, and multi-dimensional screening. *Econometrica* 66, 783–826.
- Saez, E., 2000. Using elasticities to derive optimal income tax rates. NBER working paper no. 7628, forthcoming *Review of Economic Studies*.
- Seade, J.K., 1977. On the shape of optimal tax schedules. *Journal of Public Economics* 7, 203–236.
- Seade, J.K., 1982. On the sign of the optimum marginal income tax. *Review of Economic Studies* 49 (4), 637–643.
- Stiglitz, J.E., 1982. Self-selection and Pareto efficient taxation. *Journal of Public Economics* 17, 213–240.