Capital income taxation when inherited wealth is not observable[¤]

Helmuth Cremer

IDEI and GREMAQ, University of Toulouse
and Institut Universitaire de France

Pierre Pestieau

CREPP, University of Liège, CORE and Delta

Jean-Charles Rochet

IDEI and GREMAQ, University of Toulouse

December 1999, revised March 2001

Abstract

This paper extends the Atkinson-Stiglitz model of direct and indirect taxation to a dynamic setting with two unobservable characteristics: productive ability and inherited wealth. Bequests are motivated by the "joy of giving". A child's inheritance is a random variable with a probability distribution that depends on his parent's investment in a "bequest technology". Public borrowing is assumed and implies the modi...ed golden rule. We study the optimal tax policy when two instruments are available: a non-linear (wage) income tax and a proportional tax on capital income. We show that the second instrument ought, in general, to be used but that the tax rate is not necessarily positive. However, a positive tax rate is more likely when there is a positive correlation between inherited wealth and innate ability.

1 Introduction

One of the most celebrated propositions in the optimal taxation literature is the property that income taxation does not need to be supplemented by commodity taxation, under a mild separability assumption. This proposition, which is due to Atkinson and Stiglitz (1976), has several important implications for the design of tax policy. In particular, it undermines the role played by a number of tax instruments, including capital income taxation. When cast within the framework of an overlapping-generations model, where taxes on both wage earnings and interest income are available, the Atkinson and Stiglitz analysis implies that there should be no interest income tax if the economy converges to the (modi...ed) golden rule.[1].

Not surprisingly the Atkinson and Stiglitz proposition has been challenged with the concern of justifying the often observed coexistence of direct and indirect taxation. Some authors have questioned the relevance of separability between consumption and leisure. Others have introduced arguments of di¤erential compliance to explain why the two taxes are needed.[2] Naito (1999) and Pirtilla and Tuomala (1999), on the other hand, have shown that indirect taxation may be desirable with non-linear technologies. Recently Cremer et al. (2001) have presented an alternative objection to the Atkinson and Stiglitz proposition. They consider a setting where individuals di¤er in several unobservable characteristics (productivity and endowments) and show that di¤erential commodity taxation remains a useful instrument of tax policy, even if preferences are separable between labor and produced goods.

This paper builds on Cremer et al. (2001) and studies the implications

---

[1] See Atkinson and Sandmo (1980). This argument that is generally developed in a setting of optimal income taxation à la Mirrlees (1974) can readily be extended to the case of linear wage income taxation. One then needs a stronger separability assumption (such as implied, e.g., by the Stone-Geary utility function); see Atkinson and Stiglitz (1980, p. 433).

[2] Boadway et al. (1994). See also Cremer and Gahvari (1995) who develop a direct-indirect tax structure in a setting of uncertainty.

of multi-dimensional heterogeneity for the taxation of savings. Our starting point is the observation that in reality, and for all sorts of reasons, not only ability but also part of inherited wealth may not be observable. If this is the case, then an interest income tax might become desirable even with separability. To be more precise, such a tax may then be an indirect way of screening for the part of inherited endowment that is not public knowledge. However, this essentially static argument is incomplete and leaves out some crucial elements when it comes to the taxation of savings. To introduce those, we depart from Cremer et al. (2001) in several respects. Most significantly, we account for the fact that inherited wealth (unlike endowments in a static model) cannot be assumed to be exogenous and independent of the tax policy. Our analysis shows that this is a very crucial feature; it does not in general tend to make the taxation of capital income a redundant instrument. However, it will a¤ect the level and possibly even the sign of the optimal tax on interest income. In some cases, a negative tax on capital income may be called for.

We consider a two-generations overlapping growth model in which individuals draw utility from present and future consumption, from leisure and from the prospect of leaving their children a certain amount of wealth. Each individual is characterized by two parameters: his productivity and his initial endowment. To keep the presentation simple, we assume that each of these parameters can only take two values.[3] Assuming that an optimal non-linear income tax is implemented, we examine whether or not savings ought to be taxed. In other words, taking ...rst period consumption as a numeraire, we study the optimal taxation of interest income, given that a non-linear tax is imposed on wage earnings. We focus on the steady-state solution.

Capital accumulation is equal to saving and saving is motivated by two concerns: second period consumption and bequests. We adopt a particular bequest technology. There are two possible levels of bequest, low and high.

---

[3] Our analysis can be extended to to an arbitrary (but discrete) distribution of types.

The more the parent invests in this technology, the likelier his heirs will inherit the high level. Bequests are motivated by a joy of giving argument also called the "warm glow e¤ect" (Andreoni (1990)), as opposed to dynastic altruism or strategic motivations.[4] We shall come back on this assumptions in the concluding section. Compared to Cremer et al. (2001), we endogenize the prices, here the interest rate and the wage level, and we close the model by providing a source of inherited endowment.[5]

Anticipating the results, we …nd two reasons for departing from the zero capital income tax rule. The …rst is the same as in Cremer et al. (2001): by taxing (or subsidizing) capital income, we indirectly reach inherited wealth which by assumption escapes taxation. The second normally goes towards subsidizing capital income because in our model the e¤ort toward bequeathing generates additional resources to the economy and henceforth ought to be encouraged. Finally, our paper makes a methodological contribution which goes beyond the considered context of capital income taxation. The endogeneity of inheritance means that we are e¤ectively studying an optimal income tax problem where the proportions of types are endogenous and depend on the tax instruments.[6]

The rest of the paper is organized as follows. Section 2 introduces the basic framework of the study. Section 3 then provides the formulas for optimal non linear income tax and optimal ad valorem interest income tax. Section 4 considers an extension to the case where ability and wealth are correlated.

---

[4] This assumption is not uncommon in the literature on the dynamics of wealth distribution and on endogenous growth; see e.g. Glomm and Ravikumar (1992). Empirically, that type of bequests appears to be at least as realistic as the dynastic bequests; see Arrondel et al (1997).

[5] Ludden (2000) looks at a model close to Cremer et al. (2001) in a two-period setting.

[6] Other reasons for which proportions may be endogenous include mobility (tax competition) and human capital accumulation.
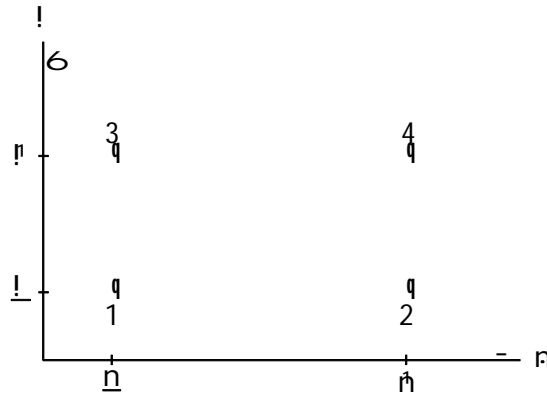
!

⑥

n̄ ┼   3   4

!̲ ┼   q   q
      1   2

                                    n

      n̲   m̄

Figure 1 : The distribution of types

2   The framework

2.1   Households, production and capital accumulation

Consider an overlapping generation model where individuals live two periods. They supply labor in the ...rst period and consume a composite good in both periods. Assume for simplicity that there is no population growth. In other words, each individual has just one child. Each generation consists of households that have di¤erent productivity in the labor market and different inherited endowment. We consider two levels of productivity $\underline{n}$ and $\bar{n}$ with $\bar{n} > \underline{n} > 0$ and two levels of inherited endowment $\underline{!}$ and $\bar{!}$ with $\bar{!} > \underline{!} > 0$. Consequently, there is a total of four types of households. The proportion of type $i = 1;:::4$ in the population is denoted $\frac{1}{4}^i$ and is assumed to be time-invariant for the time being; the indexing of the types is illustrated on Figure 1.

All individuals have the same strictly quasi concave utility,

$$u^i \left(c^i; d^i; x^i \right) i v^i \left(L^i \right) \tag{1}$$

where c^i is ...rst period consumption, d^i second period consumption, x^i the investment in the bequest technology, and L^i the labor supply. Separability is assumed for the sake of simplicity but also to keep in line with the Atkinson and Stiglitz result.

The production side of economy is assumed to exhibit constant returns

4

to scale with respect to labor and capital. Under perfect competition, ...rms do not make any pro...ts.[7] The production utilizes two factors: capital K and labor L; with:

$$L = \sum_i \mu^i n^i L^i:$$

We write the production function as $Y = F(K; L)$ and thus production prices are given by:

$$1 + r = F_K \quad \text{and} \quad w = F_L:$$

The rate of interest is r and the wage rate per e¢ciency unit is w: In other words, the wage rate of an individual with productivity n^i is $w^i = n^i w:$ Full depreciation after one period is assumed.

It is interesting to write the resource constraint of the economy for each period. To do so, we have to introduce time index. We thus have at date t:

$$F(K_t; L_t) + \sum_i \mu_t^i !^i = K_{t+1} + \sum_i \mu_t^i c_t^i + \sum_i \mu_{t_i 1}^i d_t^i + x_t^{i\mathbb{C}}: \quad (2)$$

We will also need an expression for the capital accumulation equation:

$$K_{t+1} = \sum_i \mu_t^i s_t^i {}_i B_t \quad (3)$$

where s_t^i is saving of individual i at date t, and B_t is public debt at date t. Returns from saving are used to ...nance second period consumption and spending on the bequest technology to be de...ned below. These returns come from the productive use of capital or from the interests on public debt.

As shown by Atkinson and Sandmo (1980), there are two reasons why a tax-transfer on capital income can be desirable in a traditional overlapping generation setting; see also Stiglitz (1987), Pestieau (1974) or Atkinson and Stiglitz (1980). The ...rst one is the complementarity between leisure and

---

[7] We assume an homogeneous production sector so that the objection to zero commodity taxation made by Naito (1999) does not hold. In his model, imposing a commodity tax on skilled labor-intensive goods is Pareto-improving.

savings. The second one is the desire to achieve an optimal path of capital accumulation which, in the steady-state, means satisfying the modi...ed golden rule. In this paper, we present additional arguments pro or con capital income taxation. To make our points crisper we abstract from these two traditional arguments. Consequently, we assume separability between leisure and saving. Furthermore, our setting includes a device to secure optimality of capital accumulation, namely the public debt. This is one of the mechanisms considered in the literature to obtain, in a decentralized setting, the level of capital that would be chosen in a fully controlled economy. Alternative devices include direct control of K_t, lump sum transfers across generations and pay-as-you-go social security. The public debt in our setting could be replaced by either of these mechanisms procedures without a¤ecting the results. Without such a device our formula for capital income taxation would include a term depending on the gap between the rate of capital return and the rate of economic growth (which is here 0 in the steady-state).[8]

It is important to realize that to induce the optimal capital accumulation, one may have to transfer resources "downward", namely from the old to the young generation. This is the case if the laisser-faire (Diamond) equilibrium implies under-accumulation (marginal productivity of capital lower than the rate of economic growth). Concretely, such downward transfers mean negative social security bene...ts or a negative public debt. In the latter case, the government saves instead of borrowing. This is of no relevance here as we only use public borrowing as a way of securing the modi...ed golden rule in the long run.

---

[8]What is crucial for the absence of such a term in our formula is not the modi...ed golden rule per se but simply the fact that the (steady state) interest and wage rates do not depend on the tax rate on capital. Our results would go through without any modi...cation if the interest rate were set at an arbitrary level, e.g. the worldwide return on capital in a small open economy setting.

2.2 Tax instruments and household's problem

Let us come back to the consumer's problem. In the tradition of the optimal income taxation literature, we assume that an individual's productivity, n^i; and his labor supply, L^i; are not observable by the tax administration. Yet, his before tax labor income, $I^i = w \, n^i \, L^i$; is public knowledge. This rules out ...rst-best taxation of types while nonlinear (labor) income taxation is available. Furthermore, neither personal spending levels, c^i; d^i; x^i; nor personal net consumption $z^i = c^i \, ¡ \, !^i$ are publicly observable. The tax authority has however information on anonymous transactions, in particular regarding the payment of interest income.[9] Under this circumstance, only linear taxation on interest income is available. Finally, we assume that inherited wealth, $!^i$, is not observable. This extreme assumption captures the idea that a large fraction of ...nancial wealth is not reported at death in many countries. We will take ...rst period consumption as the numeraire. Therefore, the linear tax on interest income is equivalent to a linear tax on second period spending.[10]

To sum up, the tax policy consists of a non linear tax $T(I)$ on labor income and a linear tax on saving at rate $¿$. The problem of individual of type i is given by:

---

[9] Many European countries use (anonymous) withholding taxes on interest incomes with rates around 15%.

[10] Consequently, the formal problem is the same as if the tax authority could directly distinguish ...rst period consumption from second period expenditures and tax them at di¤erent rates. Let us brie‡y elaborate on the informational assumptions regarding capital income and inherited wealth. Capital income is not observable at the individual's level but can be subject to a withholding tax, based on anonymous transactions. Inherited wealth, on the other hand, is not observable, neither at an aggregate (or anonymous) level, nor at the individual level. We also assume that the government cannot distinguish (even anonymously) the two usages of savings: future consumption and bequest.

These assumption are inter-related. Non observability of inherited wealth explains why individual saving is not common knowledge. Admittedly, in the real world, a part of inherited wealth, particularly real estate, cannot be hidden. However, for our purpose it is su¢cient that some inherited wealth (e.g., of ...nancial nature), cannot be observed. In the logic of our model, the government would impose a 100% tax on the observable wealth and the rest of the results would hold through for the non observable part of wealth.

$$\max \quad u\left(c_t^i,\, d_{t+1}^i;\, x_{t+1}^i\right) + v\left(L_t^i\right)$$

$$\text{s.t.:} \qquad c_t^i + s_t^i = l_t^i - T\left(l_t^i\right) + \iota^i \equiv R_t^i; \tag{4}$$

$$s_t^i = \left(d_{t+1}^i + x_{t+1}^i\right) p_{t+1}; \tag{5}$$

$$l_t^i = w_t\, n^i\, L_t^i \tag{6}$$

where $p_t^{-1} = (1 + r_t(1 - \tau))$ and R_t^i denotes the disposable income of individual i, obtained from gross income by subtracting the tax on labor income and adding inherited wealth.

Because of the separability of the preferences, the objective can be rewritten as:

$$\max \quad V\left(p_{t+1};\, R_t^i\right) + v\left(L_t^i\right);$$

$$\text{s.t.:} \qquad R_t^i = w_t\, n^i\, L_t^i - T\left(w_t\, n^i\, L_t^i\right) + \iota^i$$

where V is the indirect utility function associated with u, obtained by substituting the demand functions $c^i\left(p_{t+1};\, R_t^i\right)$, $d^i\left(p_{t+1};\, R_t^i\right)$ and $x^i\left(p_{t+1};\, R_t^i\right)$ into (1).

2.3 Bequest technology and the determination of each type's proportion

Up to now, we have taken as given the proportion of type i individuals, μ_t^i. We have thereby neglected one crucial feature of the model namely that these proportions are endogenously determined. We assume that they are determined by the parents' bequest behavior and specifically by their choice of x (which in turn is affected by the tax policy). Consider for the time being the simplest case where earning ability and bequest are independent random variables (and thus not correlated). The proportion of able individuals, that is those with productivity $\bar{n}$, is denoted by $\tilde{A}$ and that of individuals with high endowment is ϕ_t. We assume that $\tilde{A}$ is given but that ϕ_t results from the bequest effort x_{t-1}. More precisely, the probability that the child of a

8

parent i (in generation t) receives a high endowment, $\bar{r}$, is given by $h\left(x_t^i\right)$, with $0 \le h\left(x_t^i\right) \le 1$ and $h^0\left(x_t^i\right) > 0$.[11] The proportion of high wealth individuals in the next generation will then be given by

$$'_{t+1} = \sum_i \mu_t^i h\left(x_t^i\right):$$ (7)

The proportions of the types in generation $t + 1$ are then determined according to:

$$
\begin{array}{llll}
\mu_{t+1}^1 = \left(1 i\ '_{t+1}\right)(1 i\ \tilde{A}) & \text{for} & \underline{n}; \underline{!}; \\
\mu_{t+1}^2 = \left(1 i\ '_{t+1}\right) \tilde{A} & \text{for} & \bar{n}; \underline{!}; \\
\mu_{t+1}^3 = '_{t+1}\ (1 i\ \tilde{A}) & \text{for} & \underline{n}; \bar{!}; \\
\mu_{t+1}^4 = '_{t+1}\ \tilde{A} & \text{for} & \bar{n}; \bar{!};
\end{array}
$$ (8)

where $'_{t+1}$ is de…ned by (7). Substituting (7) in (8) yields …rst-order di¤erence equations in μ_t^i. Recall that n and $!$ are, for the time being, assumed to be independent random variables. In the steady-state, $'$, and hence μ^i's become time invariant. However, they remain endogenous and depend on the tax policy via the individual's choices of x.

2.4 The government's problem

We are now going to move to the government's problem. Its objective is to maximize an intertemporal social welfare function subject to a budget constraint at each period of time. This budget constraint can be written as:

$$\sum_i \left[\mu_t^i T_t\left(l_t^i\right) + \mu_{t i 1}^i \dot{c}\ r_t s_{t i 1}^i\right] + B_t i\ B_{t i 1}(1 + r_t) = 0:$$ (9)

By assumption the only purpose of taxation is redistributive. Using the CRS property, the pricing equations, and (3)–(5), we check that (9) implies (2). Consequently, budget constraint and resource constraint are equivalent and welfare maximization can be considered subject to either of these constraints.

[11] Recall that x^i enters the parent's utility function. In particular, we can think about u as depending on $h\left(x^i\right)$ i.e., the probability that the child of individual i inherits $\underline{!}$. For instance, the utility of a parent could be speci…ed as: $u\left(c^i; d^i; x^i\right) = \tilde{u}\left(c^i; d^i\right) + h\left(x^i\right)$.

9

In addition, the incentive compatibility constraints have to be satis…ed. Types being private information, one has to be sure that no individual mimics another. These constraints imply that individual $i = 1; :::; 4$ when he consumes $^i I_t^i; R_t^i$ has a utility

$$U_t^i = V \left(^i p_{t+1}; R_t^i \right) \text{¡ } v \left(\frac{I_t^i}{n^i w_t} \right) ;$$

that is at least equal to the utility level he would obtain by choosing $^i I_t^h; R_t^h$, $h = 1; :::; 4$. Consequently we must have:

$$U_t^i > U_t^{ih} = V \left(^3 p; R^{ih} \right) \text{¡ } v \left(\frac{I_t^h}{n^i w_t} \right) ;$$

where

$$R_t^{ih} = R_t^h + \text{!}^i \text{ ¡ !}^h:$$

The objective of the social planner is to maximize the discounted sum of utilities U_t^i, the discount factor $°$ (< 1) re‡ecting the social rate of time preference. We can now write the Lagrangean expression with the multipliers $_{s t}$ associated with the resource constraints, 1_t^{ih}; with the self-selection constraints, $´·_t$ with constraint (7) and $´_t^i$ with constraint (8).

$$\$ = \sum_t °t \left(\sum \left[¼_t^i U_t^i \right] \right.$$
$$+ _{s t} F \left(\sum_i ¼_{t_i 1}^i \left[^i d_t^i + x_t^i \right] p_{t} \text{ ¡ } B_{t_i 1}; \sum \left[¼_t^i n^i L_t^i \right] + \sum_i ¼_t^i !^i \right.$$
$$\sum_i ¼_t^i R_t^i + B_t \text{ ¡ } \sum_i ¼_{t_i 1}^i \left[^i d_t^i + x_t^i \right]$$
$$+ \sum_{i;h} 1_t^{ih} \left(^3 U_t^i \text{ ¡ } U_t^{ih} \right) + ´·_t \left(\sum_i ¼_t^{ih} \left[^i x_t^i \text{ ¡ } '_{t+1} \right] \right.$$
$$+ ´_t^1 ^i ¼_t^1 \text{ ¡ } (1 \text{ ¡ } '_t)(1 \text{ ¡ } \tilde{A}) + ´_t^2 ^i ¼_t^2 \text{ ¡ } (1 \text{ ¡ } '_t) \tilde{A}$$
$$+ ´_t^3 ^i ¼_t^3 \text{ ¡ } '_t (1 \text{ ¡ } \tilde{A}) + ´_t^4 ^i ¼_t^4 \text{ ¡ } '_t \tilde{A} : \tag{10}$$

10

This expression is to be maximized with respect to R_t^i, L_t^i, B_t, $\frac{1}{4}_t^i$, $'_t$ and p_t. Note that controlling L_t^i or $I_t^i = w_t n^i L_t^i$ is equivalent in the present setting. Recall that $d_t^i = d^i \left(p_{t+1}; R_t^i \right)$ and $x_t^i = x^i \left(p_{t+1}; R_t^i \right)$. We ...rst di¤erentiate $\$$ with respect to B_t and obtain:

$$\frac{@\$}{@B_t} = {}^\circ{}_{,t}^t \, i \, {}^\circ{}_{,t+1}^{t+1} (1 + r_{t+1}) = 0: \tag{11}$$

This equation determines the e¢cient amount of capital accumulation along the growth path. The same condition would be obtained if K_t were directly controlled by a social planner. As already mentioned, this condition can very well imply negative values of B_t. In the steady-state (11) leads to the modi...ed Golden rule:

$$(1 + r) = {}^\circ{}^{i\,1} > 1; \tag{12}$$

given that population does not grow. With this rule, we make sure that both producer prices r and w are constant in the steady-state. We also make sure that our tax instruments are not used to achieve the desired level of capital accumulation.

3   The optimal tax on capital income

To study the determination of the tax on capital income, we now turn to the optimality conditions with respect to R_t^i and p_{t+1}.[12] For the steady-state

---

[12] The ...rst-order condition with respect to L_t^i is not directly relevant for our purpose and is therefore omitted.

they are as follows:

$$\frac{\partial \$}{\partial R^i} = \left(\mu^i + \sum^h \gamma_{1ih} \frac{\partial V}{\partial R}\left(p, R^i\right) \right.$$

$$\left. - \mu^i \left(1 - \zeta rp\right)\left(\frac{\partial d^i}{\partial R}\left(p, R^i\right) + \frac{\partial x^i}{\partial R}\left(p, R^i\right)\right) \right.$$

$$\left. - \sum_h \gamma_{1hi}\frac{\partial V}{\partial R}\left(p, R^{hi}\right) + \lambda \mu^i h^0\left(x^i\right)\frac{\partial x^i}{\partial R^i} = 0; \qquad (13)$$

$$\frac{\partial \$}{\partial p} = \sum_i \left(\mu^i + \sum_{1ih}\frac{\partial V}{\partial p}\left(p, R^i\right) \right.$$

$$\left. + \sum_i \mu^i\left(d^i + x^i\right)(1 + r) + \zeta rp\left(\frac{\partial d^i}{\partial p} + \frac{\partial x^i}{\partial p}\right) \right.$$

$$\left. - \sum_{h,i} \gamma_{1hi}\frac{\partial V}{\partial p}\left(p, R^{hi}\right) + \lambda \sum_i \mu^i h^0\left(x^i\right)\frac{\partial x^i}{\partial p} = 0; \qquad (14)$$

where we use the equality $\zeta rp = p(1 + r) - 1$. Di¤erentiating $\$$ with respect to μ^i $(i = 1, \ldots, 4)$ and γ yields the following additional …rst-order conditions:

$$\frac{\partial \$}{\partial \mu^i} = U^i + \left(T^i\left(l^i\right) + \frac{\zeta rs^i}{1 + r} \right) + \gamma^i + \lambda h^i\left(x^i\right) = 0 \quad i = 1, \ldots, 4; (15)$$

$$\frac{\partial \$}{\partial \gamma} = \gamma^1(1 - \tilde{A}) + \gamma^2\tilde{A} - \gamma^3(1 - \tilde{A}) - \gamma^4\tilde{A} - \lambda(1 + r) = 0. \qquad (16)$$

We now combine (13) and (14) by taking $\sum_i \frac{\partial \$}{\partial R_i}\left(d^i + x^i\right)(1 + r) + \frac{\partial \$}{\partial p}$. This gives:

$$\zeta rp\sum_i \mu^i\left(\frac{\partial d^i}{\partial p} + \frac{\partial x^i}{\partial p}\right) + \sum_{h,i}\gamma_{1hi}\frac{\partial V}{\partial R}\left(p, R^{hi}\right)\left(l^h - l^i\right)$$

$$+ \lambda \sum_i \mu^i h^0\left(x^i\right)\frac{\partial x^i}{\partial p} = 0;$$

or

$$\zeta rp = \frac{\sum_{h,i}\gamma_{1hi}\frac{\partial V}{\partial R}\left(p, R^{hi}\right)\left(l^h - l^i\right) + \lambda \sum_i \mu^i h^0\left(x^i\right)\frac{\partial x^i}{\partial p}}{\sum_i \mu^i\left(\frac{\partial d^i}{\partial p} + \frac{\partial x^i}{\partial p}\right)}. \qquad (17)$$

12

In this formula, x^i and d^i denote the compensated demand for bequest investment and second period consumption. The derivatives with respect to p are negative. The multiplier $\acute{}$ is the shadow price of $'$; the expected proportion of wealthy heirs. One can expect that $\acute{} > 0$ as an increase in $'$ implies an increase in aggregate wealth in this steady-state economy; see below.

The left hand side of (17) is an increasing function of $¿$, the tax rate on capital income. Observe that one can interpret $¿rp$ as the per unit tax on x and d; recall that $¿rp = p(1+r) ¡ 1$.[13]. This interpretation is useful to facilitate the comparison with the results in Cremer et al. (2001).

The right-hand side of (17) consists of three different terms of which two have a familiar ‡avor. First, the denominator, which is positive, re‡ects the substitution effects and, hence, the distortions created by the tax on capital income (and thus on second period expenditures). If the two (compensated) demands are highly inelastic, there is a good case for a high tax or subsidy depending on the sign of the numerator. This is similar to the inverse elasticity rule derived in the Ramsey model.

The …rst term of the denominator is the weighted sum of the difference between the mimicker's endowment and the mimicked's endowment. This term is positive if the incentive constraints bind from high to low wealth individuals. In that case the taxation of second period expenditures is used as an imperfect substitute for the taxation of (unobservable) wealth. As Cremer et al. (2001) have shown, this depends on a number of factors including the correlation between ability and wealth. In the case of zero correlation considered up to now, it appears to be tempting to conjecture that the utilitarian solution would imply that the incentive constraint are binding downward according to wealth and ability (even though not all of these constraints would be binding). However, within a multi-dimensional setting such intuition can be misleading and no general result as for the sign

---

[13] We also have $p(x+d) = s$, so that the tax on interest income $¿rs$ is equal to $¿rp(x+d)$.

if this term can be obtained. We shall show below how the results change if wealth and ability are allowed to be correlated.

Turning now to the second term of the numerator of (17) it is negative when $\nu > 0$, which can be expected (as long as $h^0 > 0$). To see this let us have a closer look at the interpretation of this Lagrange multiplier. First, it is interesting to observe that the utilitarian full information solution would imply:

$$\nu = \gamma(\bar{T} - \underline{T}): \tag{18}$$

To obtain this expression, we have used (15)–(16) and the property that types 1 and 3 on the one hand and 2 and 4 on the other hand are treated identically at a ...rst-best utilitarian optimum.[14] Expression (18) shows that in a ...rst best world, the impact on welfare of an increase in the proportion of wealthy individual is simply equal to the social value of the di¤erence in endowments; recall that γ is the multiplier of the government's budget constraint. In the second best setting with unobservable types and binding incentive constraints considered here, the expression for ν is more complicated. Combining (15)–(16) one obtains:

$$\nu(1 - a) = \tilde{A}\left[U^4 + \gamma\left(T^4 + \frac{\dot{c}rs^4}{1+r}\right) - U^2 + \gamma(T^2 + \dot{c}rs^2) \right] +$$
$$(1 - \tilde{A})\left[U^3 + \gamma\left(T^3 + \frac{\dot{c}rs^3}{1+r}\right) - U^1 + \gamma\left(T^1 + \frac{\dot{c}rs^1}{1+r}\right) \right]; \tag{19}$$

with $0 < a < 1$. In words, ν is proportional to the average di¤erence between the contribution of high wealth and low wealth individuals to social welfare.[15] Using the incentive constraint along with some weak normality conditions one can show that (19) implies $\nu > 0$.

---

[14] Recall that preferences are separable. It is than easy to verify that the ...rst-best utilitarian solution yields: $l^1 = l^3$; $l^2 = l^4$; $R^1 = R^2 = R^3 = R^4$; $x^1 = x^2 = x^3 = x^4$; $U^1 = U^3$ and $U^2 = U^4$.

[15] Recall that types 4 and 2 on the one hand and 3 and 1 on the other hand have the same productivity but di¤er in wealth. An increase in $'$ increases the proportion of types 4 and 2 at the expense of types 3 and 1.

It follows from these arguments that the presence of this second term tends to decrease the tax rate on capital income. This is because an increase in ζ brings about a decrease in the individuals (compensated) investment in the bequest technology. This reduces the proportion of high wealth individuals and thus makes society poorer. Interestingly, this wealth e¤ect is not associated with an impact on capital accumulation. The steady state capital stock is determined according to (12) and is thus independent of the tax policy. This second term captures the idea that x not only provides some utility to bequeathers but also creates a positive externality on the economy. Consequently, it can be viewed as a Pigouvian subsidy.[16] Note that in a …rst-best setting, this Pigouvian subsidy would only apply to x and not to both x and d as here.

To further clarify the interpretation of (17) it is useful to look at two special (extreme) cases:

- Case 1: $\mu^{hi} = 0$; $\forall h, i$ such that $\omega^h \neq \omega^i$. In words, incentive constraints bind only between types who have identical wealth. This would be the case for instance if inherited wealth were observable (see the Conclusion for further discussion of such a setting). Then the …rst term on the RHS of (17) vanishes, and we are left with the second term which calls for a subsidization of savings. The taxation of savings has no redistributive (traditional optimal tax) role to play here, and only the Pigouvian (externality correcting) e¤ect is relevant.

- Case 2: $h'(x) = 0$: parents cannot a¤ect the probability distribution of their child's endowment (and would then of course set x at zero). Here we are left with the traditional optimal tax considerations and

---

[16] We have de…ned the …rst- and the second-best objective of the social planner as the utilitarian sum of individuals' utilities, including the terms pertaining to x. Consequently, bequests contribute twice to welfare: …rst through their e¤ect on individuals' utility and second through their e¤ect on aggregate inherited resources $\gamma \underline{L} + (1 - \gamma)\overline{L}$. If instead we had chosen to purge individuals' utilities from their altruistic component, the social desirability of bequests would have been smaller and the case for a Pigouvian subsidy weaker than in the current speci…cation.

the possibility to have a positive tax on savings.

Summing up, we do not obtain what would be the counterpart of Atkinson and Stiglitz's result: the tax on capital income (or equivalently on second period expenditures) is not made redundant by the general income tax. Put di¤erently, the optimal tax rate on capital income is not in general equal to zero. However, at this point, there does not appear to be a compelling case for imposing a positive tax on capital income. The possibility that the optimal tax rate implied by (17) is negative cannot be ruled out.

4 Correlation between ability and wealth

So far we have assumed that ability and wealth were independent random variables. This has allowed us to make the description of the bequest technology and of the determination of the types' proportions as simple as possible. We shall now show how this setting can be generalized to allow for correlation between the two individual characteristics. Let $\frac{1}{2} \in [-1, 1]$ denote the coe¢cient of correlation between n and $!$, and assume that it is exogenously given and constant over time. The case of a positive correlation is of course the one which appears empirically the most appealing, but our formal model allows also for negative correlation.

For simplicity we continue to assume that the conditional probabilities of $\underline{n}$ and $\overline{n}$ are independent of wealth.[17]. Consequently, the proportion of able individual will be $\bar{A}$, both among the high wealth and the low wealth individuals. This is exactly similar to the case with $\frac{1}{2} = 0$ considered above. The new feature which is introduced here is that the conditional probability of inheriting a high level of wealth now depends on n. Let us de…ne $'^L$ and $'^H$ as the probability of inheriting a high level of wealth given that

---

[17]This essentially amounts to assuming that while parents' have an impact on their children's inheritance, they cannot a¤ect abilities. Alternatively, we could have assumed two levels of (conditional) probabilities for $\overline{n}$, $\bar{A}^L$ and $\bar{A}^H$, depending on the inherited wealth. This complicate the expression for $\frac{1}{2}$, (22), while leaving the main conclusions una¤ected.

productivity is low or high respectively. With these additional, the steady state version of (8) can be rewritten as:

$$
\begin{aligned}
\mu^1_{t+1} &= \left(1 - \pi^L_{t+1}\right)(1 - \theta) \\
\mu^2_{t+1} &= \left(1 - \pi^H_{t+1}\right)\theta \\
\mu^3_{t+1} &= \pi^L_{t+1}(1 - \theta) \\
\mu^4_{t+1} &= \pi^H_{t+1}\,\theta .
\end{aligned}
\tag{20}
$$

Now, the proportion of individuals inheriting $\top$ is given by $\pi_t = \pi^L_t (1 - \theta) + \pi^H_t \theta$. We continue to assume that the probability of any given child of inheriting $\top$ is given by the level of $h\left(x^i_t\right)$ chosen by its parent. This implies that like before π_{t+1} is equal to the average level of $h\left(x^i_t\right)$, and (7) is replaced by:

$$
\pi_{t+1} = \pi^L_{t+1}(1 - \theta) + \pi^H_{t+1}\theta = \sum \mu^i_t\, h\left(x^i_t\right) .
\tag{21}
$$

Observe that the coefficient of correlation can easily be expressed as a function of π^H and π^L and we have:

$$
\rho = \frac{\left(\pi^H_{t+1} - \pi^L_{t+1}\right)(1 - \theta)\,\theta}{\pi_{t+1}\left(1 - \pi_{t+1}\right) + \theta(1 - \theta)} .
\tag{22}
$$

With the assumption that ρ is given (and constant over time) expression (20)–(22) then completely determine the dynamics and the steady state values of proportions μ^i's and π.

The interesting feature is that all this has very little impact on the specification of the Lagrangean expression for the government's problem. Starting from (10), the Lagrangean in Section 2.4, one just has to replace the last four constraints by their counterpart obtained from (20) while adding (22) as a new constraint. It then immediately follows that the expression for the optimal tax on capital income is not affected and continues to be given by (17). The actual level of τ as well as the interpretation of (17) do however depend on the degree of correlation.

17

To see this, observe that the case where incentive constraints are binding from high to low wealth individuals becomes increasingly likely as the correlation increases; see Cremer et al. (2001). At the limit, when there is perfect positive correlation ($\frac{1}{2} = 1$), it is easy to see that this is necessarily true. This is because under perfect correlation, we return to a standard single dimensional setting for which one easily determines that the downward incentive constraints bind at the utilitarian solution. In this case, the ...rst term in the denominator of (17) is thus necessarily positive. Furthermore, the case of perfect correlation is only possible if $h^0(x) ´ 0$ so that the second term must also be zero.[18].

To sum up, when $\frac{1}{2} = 1$, a positive tax on capital income is necessarily optimal. Here the case for taxing capital income as a substitute for unobservable bequest is of course strongest. We can expect the same type of conclusion to emerge (by continuity) when correlation is not perfect, but su¢ciently high. And we know from the previous section that $\frac{1}{2} = 0$ may or may not qualify as "su¢ciently high" from this perspective. Finally, when $\frac{1}{2} < 0$ a negative tax on capital income (i.e., a subsidy) becomes increasingly likely. We can then have the case that both terms in the numerator of (17) are negative. However, it has to be pointed out that this is not guaranteed either, even in the case where $\frac{1}{2} = ¡ 1$. In such a situation the direction of binding self-selection constraints would depend on the relative importance of wealth and productivity di¤erentials.

5 Conclusion

The starting point of this paper was the now classic Atkinson and Stiglitz's proposition that with separability between consumption and leisure, there is no need to tax capital income; non linear income tax is su¢cient both to raise

---

[18] If the parent knows that he cannot in‡uence the probability of his child receiving !' because this probability only depends on his child's ability, why would he bother leaving him any bequest ?

revenue and to redistribute resources across households. While keeping the separability assumption, we have added one feature to Stiglitz's [1985; 1987] overlapping generations model, namely the desire of parents to leave some bequests to their heirs. Consequently, individuals are characterized not only by their unobservable level of productivity but also by an equally unobservable level of inherited wealth. Even though we realize that a non negligible part of inheritance can be observed by taxing authorities, the fact that part of it currently escapes tax control is sufficient to justify such an assumption. In our model, there are two sources of output: a standard CRS technology which is subject to the modified golden rule through public debt and a linear technology which transforms a certain amount of bequests into some inherited wealth. Such a specification makes it desirable no to tax but to subsidize saving. There is however a good reason for taxing saving; it can indeed be an indirect way of taxing inherited wealth and thus of redistributing resources across households. This reason is particularly persuasive when there is some correlation between labor productivity and inherited financial endowment.

Throughout this paper we have assumed that inherited wealth, was not observable. If, instead, ω^i's were observable, the results would of course be quite different. The first term in the denominator of the RHS of (17) would then vanish because incentive constraints between types with different ω's are irrelevant. Consequently, one is then left with the second term pushing for a subsidy on savings. Observe that a 100 % tax on bequests would then be desirable.[19] We thus return to the Atkinson and Stiglitz setting with a single source of heterogeneity. The tax on savings then has no redistributive role to play. However, the external effect of x, namely its impact on the proportion of types remains relevant so that a Pigouvian subsidy is called for.

---

[19] To show this, on can derive the first order condition of the governments problem with respect to I^i and combine it with (13). From the resulting expression it directly follows that $n^i = n^j$ implies $I^i = I^j$, and $R^i = R^j$, so that $T^i - T^j = \omega^i - \omega^j$; $i, j = 1, \ldots, 4$.

Finally, let us come back to our bequest motive. We use a joy of giving speci...cation (also called paternalistic altruism) rather than an approach based on accidental or purely altruistic bequests. This raises two questions. First, one may wonder whether it is legitimate for the social planner to include this altruistic component in its objective? Harsanyi and Hammond believe that it should be dropped.[20] Then our optimal tax system would be less biased towards saving and the case for a positive capital income taxation would be strengthened. In this paper, we have decided not to follow Harsanyi and Hammond and not to "launder" individuals' utilities.

The second question concerns the implications of considering alternative assumptions, such as a speci...cation based on pure altruism à la Barro or an accidental inheritance setting. Pure altruism makes the household problem quite di¢cult. Without randomness, we know from Chamley (1986) and Lucas that capital income taxation should not be taxed. Given our informational structure, what would be the behavior of an agent having an in...nite horizon but facing a bequest technology that can yield a zero inheritance if the low endowment is zero? This problem is close to that studied by Gevers and Michel (1998). It is not impossible that the government has no role in such an economy. Even if there is room for government intervention, its implication will be surely di¤erent from the ones we have studied. Accidental bequests constitute yet another alternative setting. They have been considered in a recent paper by Boadway, Marchand and Pestieau (2000) whose conclusions are quite close to those obtained in our paper.

---

[20]See e.g., Hammond (1988).

References

[1] Andreoni, J.,(1990), Impure altruism and donations to public goods: A theory of warm-glow giving, Economic Journal, 38, 209-227

[2] Arrondel, L., A. Masson and P. Pestieau, (1997), Bequests and inheritance : empirical issues and French-US comparison, in G. Erreygers and T. Vandevelde (eds), Springer-Verlag Berlin, 89-125.

[3] Atkinson, A.B. and A. Sandmo, (1980), Welfare implications of the taxation of savings, Economic Journal, 90, 529-549.

[4] Atkinson, A.B. and J. Stiglitz, (1976), The design of tax structure: direct and indirect taxation, Journal of Public Economics, 6, 55-75.

[5] Atkinson, A.B. and J. Stiglitz, (1980), Lectures on public economics, McGraw Hill, New York.

[6] Boadway, R., M. Marchand and P. Pestieau, (1994), Towards a theory of direct-indirect tax mix, Journal of Public Economics, 55, 71-88.

[7] Boadway, R., M. Marchand and P. Pestieau, (2000), Redistribution with unobservable bequests: a case for taxing capital income, Scandinavian Journal of Economics, 102, 253-267.

[8] Chamley, C, (1986), Optimal taxation of capital income in general equilibrium with in...nite lifes, Econometrica, 54, 607-622.

[9] Cremer, H. and F. Gahvari, (1995), Uncertainty, optimal taxation and the direct versus indirect tax controversy, Economic Journal, 105, 1165–1179.

[10] Cremer, H., P. Pestieau and J.-Ch. Rochet, (2001), Direct versus indirect taxation. The design of the tax structure revisited, International Economic Review, (forthcoming).

[11] Gevers,L and P. Michel (1998), Economic dynastics with intermissions, Games and Economic Behavior, 25, 251–271.

[12] Glomm, G. and B. Ravikumar, (1992), Public versus private investment in human capital: endogenous growth and income inequality, Journal of Political Economy, 100, 813–834.

[13] Hammond, P, (1988), Altruism, in J. Eatwell, M. Milgate and P. New-manthe (eds) The New Palgrave: A Dictionnary of Economics, Mac Millan Press, London.

[14] Ludden, U., (2000), Optimal capital income taxation and redistribution, unpublished.

[15] Mirrless, J., (1971), An exploration in the theory of optimum income taxation, Review of Economics Studies, 38, 175-208.

[16] Naito, H., (1999), Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production e¢-ciency, Journal of Public Economics, 71, 165-188.

[17] Ordover, J. and E. Phelps, (1979), The concept of optimal taxation in the overlapping-generations model of capital and wealth, Journal of Public Economics, 12, 1-26.

[18] Pestieau, P., (1974), Optimal taxation and discount factor for public investment in a growth setting, Journal of Public Economics, 217-235.

[19] Pirtilla, J. and M. Tuomala, (1999), Non linear taxation and public good provision in an overlapping generations economy, Journal of Public Economics, (forthcoming).

[20] Stiglitz, J.E., (1985), Inequality and capital taxation, IMSS Technical Report n$^\pm$ 457, Stanford University.

[21] Stiglitz, J.E., (1987), Pareto-e¢cient and optimal taxation and the new welfare economics, in A. Auerbach and M. Feldstein (eds), Handbook of Public Economics, vol. 2, Elsevier.