Yet Another Reason to Tax Goods*

Carlos E. da Costa†
Getulio Vargas Foundation

March 30, 2004

Abstract

In this article we write a two period taxation model where: *i)* private information changes through time; *ii)* savings choices by an agent that are not observed, *iii)* affect preferences conditional on the realization of types. The simultaneous appearance of these three elements cause optimal commodity tax to depend on off-equilibrium levels of savings. As a consequence, separability no longer suffices for the uniform taxation prescription of Atkinson and Stiglitz to obtain. In what regards capital income taxation we show that, in the most 'natural' cases, return on capital ought to be taxed.

Keywords: Optimal Taxation; Non-observable savings; Dynamic Agency.
JEL Classification: H21, D82.

1 Introduction

The main result concerning the role of commodity taxation in the presence of an optimally designed non-linear income tax schedule is the uniform tax prescription of Atkinson and Stiglitz (1976) - henceforth AS. It says that, if preferences are

*I thank Luis Braido, Bev Dahlby and seminar participants at PUC-RJ, IBMEC and EPGE for their helpful comments. The usual disclaimer applies.

†Address: Graduate School of Economics at the Getulio Vargas Foundation - Praia de Botafogo 190/1105. Rio de Janeiro/RJ-BRAZIL 22450-900; E-mail: cdacosta@fgv.br. Phone: 55-21-2559.5493

separable between leisure and the other goods, there is no need for taxing goods - the non-linear tax schedule on income will fully implement the second best allocation.[1]

In this paper, we investigate AS in a two periods version of Stiglitz's (1982) representation of the optimal income taxation problem. The crucial dynamic element here is an evolving information set for the agents.

Elsewhere these types of dynamic features have been added to this framework - e.g. Cremer and Gahvari (1995), in a two type setting, Cremer and Gahvari (1999), da Costa and Werning (2000), Golosov et al. (2003), with a continuum of types. However, in all these cases, savings are directly controlled by the government, while here, I assume that government can only affect agents savings decisions by means of a linear tax on capital income.

On the one hand, this restriction on tax instruments incorporates in the model the sensible fact that, though the government can tax labor income and define net income as a function of gross income, it cannot control how agents distribute expenditures through time.[2] On the other, taxes on capital income that implements an allocation where savings are controlled by the government, are not measurable with respect to the information set available at the moment the investment is made. Here, I assume neither. The government can only use an anonymous tax on capital income that is known by the agent at the moment she chooses how much to save.

It turns out that this simple and compelling restriction on tax instruments is sufficient to overturn AS; a result that does not happen in all previous dynamic taxation models.

It is also worth noting that usual violations of AS, either in a static or a dynamic model, are caused by the introduction of an extra dimension of unobserved heterogeneity.[3] Here, however, heterogeneity is still unidimensional, in the form of differences in productivity.

[1] This result has been expressed in many different ways in Mirrlees (1976), Cooter (1978), Christiansen (1984) and Konishi (1995) and challenged in many others - e.g., Naito (1999), Saez (2001), Cremer et al. (2002), etc.

[2] We emphasize non-observability of savings, but other forms of smoothing consumption - like purchase of durables - render the control of consumption virtually infeasible.

[3] An important exception is Naito (1999), which, contrary to all papers cited before, considers a technology where labor inputs of different productivities are not perfect substitutes. Here, I maintain the hypothesis of perfect substitutability.

More to the point, in a recent contribution to the literature, Cremer et al. (2001) show how uniform taxation is usually not optimal, and how income effects become important, when another dimension of unobserved heterogeneity - namely, different endowments of some goods - is present.[4] What is surprising in our result is that we can generate tax prescriptions akin to theirs without the artifact of differences in endowments. As in their paper, the uniform taxation result breaks down in our setup and we homotheticity must be added for AS to obtain.

As we shall point out, the similarity of results is not a mere coincidence: there is a subtle way in which unobserved heterogeneity shows up in our framework. Agents off the equilibrium path have different 'endowments' from a second period perspective - as Cremer et al. (2001) - for they have different level of unobserved savings. This difference drives most of our results.

An important by-product of our discussion is a set of results related to the taxation of capital income. When taxes on capital income are allowed to depend on the realized type of each agent, and with separable utility, an inverse Euler equation describes the optimal intertemporal allocation. This implies that marginal utility of today's income is smaller than expected marginal utility of tomorrow's income. Though one might be tempted to interpret this inverse Euler equation as a prescription of positive tax on capital income, Albanesi and Sleet (2003) have shown that this is not the case. The expected tax on capital income is zero - and, consequently, no revenue is raised with these taxes. The inverse Euler equation is a consequence of a tax function that induces a positive correlation between after tax returns on asset and labor income.

In our paper, because the tax rate is constant across types, we are able to go unambiguously from the relationship between marginal utility of income today and tomorrow to the sign of marginal tax rates.

We show that when the incentive compatibility constraints bind in the 'usual direction' - where a high productivity agent pretends to be a low productivity one[5] - savings ought to be taxed. This result, is due to the fact that agents who 'intend to announce' falsely their types - i.e., announce low if they turn out to be high - save more then agents who intend to abide by the rules. Punishing off-equilibrium

[4] See their example in page 790, second paragraph.

[5] As we shall see single-crossing is not sufficient to guarantee that this will always be the case.

behavior helps relax IC constraints.

The remainder of this paper is organized as follows. The model of the economy is presented in section 2. Then, in section 2.1 the concept of equilibrium and the approach we adopt for tackling the problem is described. Optimal taxation is characterized in section 3, where Atkinson and Stiglitz (1976) uniform taxation result is discussed. Section 4 concludes. All results are proved in the appendices.

2 The Environment

The economy is populated by a continuum of 'ex-ante' identical agents who live for two periods and have preferences represented by

$$\upsilon\left(\mathbf{x}^0, l^0\right) + E[\upsilon(\tilde{\mathbf{x}}^1, \tilde{l}^1)], \tag{1}$$

where $\mathbf{x}^0, \mathbf{x}^1 \in \mathbb{R}^n$ are, respectively, first and second period consumptions and $l^0, l^1 \in \mathbb{R}$ are, respectively, first and second period labor supplies. We shall be using bold to represent vectors, and following the convention that prices are row and quantities are column vectors.

Uncertainty arises in this problem because in the first period agents do not know their 'adult' productivities, which we call their types. Preferences over bundles of l and $\mathbf{x}$ are identical; productivity is the only dimension of heterogeneity 'ex-post'. An agent of productivity w needs $l = Y/w$ hours to supply Y efficient units of labor. Hence, the higher the productivity, the more leisure one agent gets for a given output she produces.

To make the problem as simple as possible we follow Stiglitz (1982) in considering only two possible types: H (for high productivity) and L (for low). Shocks are assumed to be independent across agents so 'ex-post' distribution of types coincides with the 'ex-ante' one. For notational simplicity, we take types to be in equal proportion in the population.

Given our explicit intent to investigate Atkinson and Stiglitz's uniform taxation result, we take temporary utility to be separable between consumption and leisure/labor, $\upsilon(\mathbf{x}, l) \equiv u(\mathbf{x}) - \zeta(l)$, with $u'(\cdot), \zeta'(\cdot), \zeta''(\cdot) > 0$ and $u''(\cdot) < 0$.

With these assumptions, (1) becomes

$$u\left(\mathbf{x}^0\right) - \zeta\left(l^0\right) + \frac{1}{2}\left[u\left(\mathbf{x}^H\right) - \zeta\left(l^H\right) + u\left(\mathbf{x}^L\right) - \zeta\left(l^L\right)\right]. \tag{2}$$

As for the informational/transaction structure, the economy is divided in two sectors: a production sector and a consumption sector.

The production sector transforms efficiency units $Y \equiv lw$ in goods $\mathbf{x}$ with a linear technology. Units are chosen in such a way that production prices are also normalized to 1. Output may also be transferred inter-temporally using a linear technology. That is, the production sector transforms one unit of consumption today in one unit of consumption tomorrow.

Irreversibility poses a difficulty here. Efficient units in the first period may be used to produce a unit of consumption tomorrow, but the reverse does not hold. It is, then, quite possible that, at the optimum, agents are constrained in their willingness to smooth consumption across time. We sidestep this discussion by assuming that the constraint that consumption in the first period is no grater than production is not binding.[6]

Efficiency units may be sold to the production sector at no transaction cost, but within the consumption sector they may only be traded at prohibitively high costs. The trade of goods, including the trade of consumption across time, can be done between agents - outside the production sector - at no transaction costs.

A benevolent government who inhabits this economy adopts a tax system that maximizes the agents' expected utilities. Tax instruments, however are restricted by the informational/transaction structure of the economy.

As is standard in optimal taxation, we assume that once uncertainty is realized each agent's productivity is only observed by the agent herself. While productivity is not directly observable total output produced by each agent is observed by everyone; agents may sell to the production sector efficiency units.[7]

All transactions between sectors are assumed to be observable by the government while transactions within sectors are not.

Because all trade involving labor is observed by the government, the use a fully non-linear income tax schedule is possible. Negligible transaction costs for goods within the consumption sector mean, however, that on the one hand, the govern-

[6]The assumption is immaterial for the type of discussions that we aim at having. More importantly we want to think of this model as a reduced form of an overlapping generations economy, where the intertemporal nature of the resource constraint is removed.

[7]See Guesnerie (1995).

ment cannot directly control savings and, on the other it is not possible to define a non-linear tax schedule for other goods. Non-linear taxes would naturally lead to arbitrage opportunities in trades of goods.

2.1 The Direct Mechanism

To find the optimal tax schedule we define a truthful direct mechanism and derive the allocation that maximizes the government objective function. Given the optimal allocations, we invoke Hammond's taxation principle to map the optimal allocation into a tax system.

The problem here is, however, non-standard because the information set evolves in a non-trivial manner. Hence, we dedicate the next few pages to discuss the characterization of optimal allocations.

2.1.1 The Nature of the Game

Let y denote the after tax income of an agent - and Y, the gross income, as defined before. Then, the game played by the government and the agents is a Stackelberg game, where the Government, the Stackelberg leader, moves first by choosing (y, Y), a budget set $\left\{\left(y^L, Y^L\right), \left(y^H, Y^H\right)\right\}$, and tax rates $\boldsymbol{\tau} \equiv (\mathbf{p} - \mathbf{1})$ and $\theta = 1 - R$, where $\mathbf{p}$ are consumer prices and R is the after tax gross rate of return for this economy. It should also be clear in this case, that $T\left(Y\right) \equiv Y - y$.

The agent follows, with the next move, by deciding what and how much to consume (or, equivalently, how much to save), given (y, Y). This decision is made before nature defines the agent's type. Once individual productivities are realized each agent chooses her bundle from the budget set offered by the government.

The solution for this problem is potentially very complex. The fact that savings affect preferences over bundles of (y, Y) means that whether the IC constraints are satisfied or not depends on the level of savings. On the other hand, it is the expected marginal utility of income at the incentive compatible allocations that will determine the optimal level of savings, both on and off the equilibrium path.

The way we deal with this simultaneity is by first defining strategies as mappings from types to announcements $\sigma : \{H, L\} \rightarrow \{H, L\}$. We argue that, if an agent decides to adopt strategy σ, we need only to consider the incentive compatibility constraints for this strategy at the appropriate - that is, expected utility maximizing - level of savings.

A strategy is a rule that associates to each type a specific action. In this case, an announcement. The choice of an optimal announcement strategy, however, is not done in isolation: the agent is also choosing an optimal level of savings. In fact, because each strategy defines a strictly concave savings problem, we associate to each announcement strategy a unique optimal level of savings.

Formally, let $\sigma^k(j)$ be the announcement prescribed by strategy k if one realizes type j. With two types there are four possible strategies:[8]

$$\begin{array}{ll} \sigma^1(H) = H; & \sigma^1(L) = L \\ \sigma^2(H) = L; & \sigma^2(L) = L \\ \sigma^3(H) = H; & \sigma^3(L) = H \\ \sigma^4(H) = L; & \sigma^4(L) = H \end{array}$$

We shall use a special notation σ^* for strategy σ^1, the truthful announcement. Since this is the strategy we want to induce, we shall be comparing it with all other strategies.

The point we emphasize here is that, when looking for implementable allocations, we need only to compare the expected payoff for pairs of announcement and savings - $\left(\sigma^k, s^k\right)$ for $k = 2, 3, 4$ - with the expected payoff for the pair (σ^*, s^*) we want to induce.

We let y^i and Y^i be, respectively, the income available and the output produced by an agent who announces type i (for $i = H, L$) and y and Y the same variables in the first period of her life, that is, before she realizes her 'adult' productivity. Under our separability assumption, labor supply plays no role whatsoever in determining the optimal level of savings.

Let

$$v(\mathbf{p}, I) \equiv \max_{\mathbf{x}} u(\mathbf{x}) \text{ s.t. } \mathbf{p} \cdot \mathbf{x} \leq I,$$

with $\mathbf{x}(\mathbf{p}, I)$ as the corresponding (conditional) Marshallian demand. Then, we can

[8] Let Σ be the set of strategies, then if we let $\#(\Sigma)$ denote the cardinality of this set we have that $\#(\Sigma) = 2^{\#(\mathcal{I})}$ where $\#(\mathcal{I})$ is the cardinality of the set of types. Hence, the convenience of working with only two types.

define s^k by

$$s^k \equiv \arg\max_s \left\{2v\left(\mathbf{p}, y - s\right) + \sum_{i=H,L} v\left(\mathbf{p}, y^{\sigma^k(i)} + Rs\right)\right\}. \tag{3}$$

where $y^{\sigma^k(i)}$ is net income received by an agent of type i using strategy k.

At this point it should already be clear that it does not suffice to consider IC constraints at the equilibrium level of savings. We have to consider off-equilibrium savings choices because, though truthful announcement may be the optimal strategy at the equilibrium level of savings, there might be another level of savings that makes some other strategy's expected payoff higher than the equilibrium one.

2.1.2 The Redundancy of Second Period IC constraints

In principle, to set up the program to be solved by the government we should include: *i)* the two second period IC constraints that guarantee that each agent finds it in her best interest to announce truthfully *ii)* the three first period IC constraints that guarantee that the truthful announcement *strategy* is the chosen one. However, as we will see, only two constraints need to be considered.

To get to this point we first introduce some notation that will allow us to economize on space and improve aesthetically the paper. Let

$$U\left(\mathbf{p}, I, Y, w\right) \equiv v\left(\mathbf{p}, I\right) - \zeta\left(Y/w\right).$$

Then, we define

$$U^k\left(i\right) \equiv U\left(\mathbf{p}, y^i + Rs^k, Y^i, w^i\right)$$
$$U^k\left(i|j\right) \equiv U\left(\mathbf{p}, y^i + Rs^k, Y^i, w^j\right)$$
$$U^k \equiv U\left(\mathbf{p}, y^i - s^k, Y, 1\right)$$

for $k = *, 2, 3$, and 4.

We should also notice is that if we write $m\left(\cdot\right) \equiv -U_Y/U_y$, with subscripts denoting partial derivatives, our assumption on preferences imply

$$\frac{\partial m\left(\cdot\right)}{\partial w} < 0, \text{ and } \frac{\partial m\left(\cdot\right)}{\partial s} < 0.$$

The first consequence is single-crossing (or the Spence-Mirrlees condition), while the second one, due to normality of leisure, will turn out to be very important for many of our results.

Notice the dependence of $m(\cdot)$ on s. This means that the slope of indifference curves in the $y \times Y$ space are affected by the choice of savings - this is the sense in which first period choices affect second period ranking of (y, Y) bundles.[9] The ranking of bundles is not, in this case, independent of an agents' first period choices.

What we show next is that if first period IC constraints,

$$
\begin{aligned}
2U^* + U^*(H) + U^*(L) &\geq 2U^2 + U^2(L|H) + U^2(L), \qquad (4) \\
2U^* + U^*(H) + U^*(L) &\geq 2U^3 + U^3(H) + U^3(H|L) \text{ and} \qquad (5) \\
2U^* + U^*(H) + U^*(L) &\geq 2U^4 + U^4(L|H) + U^4(H|L), \qquad (6)
\end{aligned}
$$

are satisfied so are second period ones

$$
\begin{aligned}
U^*(H) &\geq U^*(L|H) \text{ and} \qquad (7) \\
U^*(L) &\geq U^*(H|L). \qquad (8)
\end{aligned}
$$

In fact assume that (4) holds. Because savings are optimally chosen in (4), the right hand side of this equation is no less than the same expression evaluated at any level of savings. In particular at $s = s^*$, that is,

$$
\begin{aligned}
2U^* + U^*(H) + U^*(L) &\geq 2U^* + U^*(L|H) + U^*(L) \\
&\therefore\ U^*(H) \geq U^*(L|H)
\end{aligned}
$$

In general, constraint (7) is satisfied with strict inequality.

An analogous argument can be used to show that (5) implies (8).

Hence, what we have shown is that, if the first period constraints are satisfied, so are the second period ones. It is apparent that second period IC constraints are usually satisfied with strict inequalities. As a consequence tax schedules will be interim inefficient,[10] in the sense that once the saving decision is made, agents would want the government to redesign the tax schedule.

[9] Notice that models of moral hazard followed by self-selection - which are often considered in similar discussions - do not generate this effect. Expected utility is affected by means of a change in the distribution, but preferences, conditional on a specific realization of uncertainty is not affected.

[10] Conversely, considering the case of (4), if we required

$$U^*(H) = U^*(L|H)$$

These results are akin to the ones found in the repeated moral hazard literature.[11] Once savings choices are made, agents would be better off if the government could redesign the tax schedule. The 'tax system' - for that matter, any deterministic implementable contract - is not renegotiation-proof in the sense of Dewatripont (1988).

2.1.3 The Relevant First Period IC Constraints.

After showing that if first period constraints are satisfied so are second period ones, we are left with the three first period IC constraints. What we show next is that, if there is a level of savings that, at the same time, makes strategy 4 optimal, and gives the same expected utility that the truthful announcement, then one of the other two constraints is violated.

In fact, assume that constraint (6) is binding at the optimal allocation. Then, there is a level of savings s^4 such that

$$2U^* + U^*(H) + U^*(L) = 2U^4 + U^4(H|L) + U^4(L|H).$$

Notice that, because this is the optimal strategy, it must be the case that

$$U^4(H|L) \geq U^4(L) \text{ and } U^4(L|H) \geq U^4(H).$$

However, it can be shown - see proof of claim 1 - that this requires $\left(y^H, Y^H\right) < \left(y^L, Y^L\right)$, in which case the allocation is not implementable according to the following.

Claim 1 *Monotonicity is necessary for an allocation to be implementable*

to be binding, we would have,

$$2U^* + U^*(H) + U^*(L) = 2U^* + U^*(L) + U^*(L|H)$$

But then, generically,

$$2U^* + U^*(H) + U^*(L) < 2U^2 + U^2(L) + U^2(L|H).$$

Which means that the allocation would not be implementable. The agent would choose to save an amount s^2 (not s^*) and always pick the allocation intended for type L.

[11] See Chiappori et al. (1995), for example.

Therefore, whenever constraints (4) and (5) are satisfied, (6) is satisfied as a strict inequality. Hence, we may always leave it in the background.

The maximization problem potentially includes both IC constraints (4) and (5). However, instead of investigating it in its most general format, we will concentrate on the case that is most likely to emerge for a reasonable parametrization of our model economy - the high type pretending to be a low type.

Notice that none of our results regarding AS depend on this being the case - though tax formulae would differ. However, the sign of the optimal tax rate on capital income will be different if it is constraint (5) rather than (4) that binds at the optimum.

3 Optimal Taxation

Before writing down the government's program, a word about tax rates on goods is needed. In principle, we could allow prices in the first period to differ from prices in the second period. That is, we could have two different price vectors $\mathbf{p}_t$ and $\mathbf{p}_{t+1}$ as instruments (in which case R would be redundant), instead of requiring $\mathbf{p}_t = \mathbf{p}_{t+1}$.

Because we want to think about this model as a reduced form for an overlapping generations economy - though we do not explicitly model it in order not to create distractions over the main issues we want to discuss - we restrict the government's instruments based on this idea.[12]

The problem the government solves is, in this case,

$$\begin{aligned} &\max_{\mathbf{p},R,\{(Y^i,y^i)\}_{i=H,L,0}} \quad 2U^* + U^*(H) + U^*(L) \\ &\text{s.t. } 2U^* + U^*(H) + U^*(L) \geq 2U^2 + U^2(L|H) + U^2(L) \qquad [\ \mu\ ] \\ &\text{s.t. } 2Y + Y^H + Y^L \geq 2\textstyle\sum_k x^{*k} + \sum_k \sum_{i=H,L} x^{*k}(i), \qquad [\ \lambda\ ] \end{aligned}$$

where $\mathbf{x}^*$, $\mathbf{x}^*(H)$, $\mathbf{x}^*(L)$, denote Marshallian demands and where R is the (inverse of) consumer price of future income in terms of today's consumption, i.e. gross interest rates. Notice also the Lagrange multipliers inside brackets to the right of each constraint.

[12] An earlier version of this paper does not impose this restriction. The same results obtain.

There are some important differences here with respect to a standard optimal taxation problem. First, there is an extra term in the utility function which is the first period utility. Second, the s term that appears not only in the objective function but also in the IC constraints. Most important, however, is the fact that the IC constraints are not there to guarantee that the agent chooses a certain *action*, but that she chooses a certain *strategy* (and corresponding savings).

Finally, it is worth remarking that first period labor supply does not appear in the IC constraint (it appears in both sides of $[\mu]$). This does not mean that labor supply is not distorted in the first period. In fact, it is easy to verify that, at the optimum, $v_y(\mathbf{p}, y) \neq \zeta'(Y)$.

In appendix A.2, we derive the tax prescription below

$$\begin{aligned} &2\mu \left[U_y^2 \left(x^{2j} - x^{*j}\right) + U_y^2 (L) \left(x^{2j} (L) - x^{*j} (L)\right)\right] = \\ &-\lambda \textstyle\sum_k \tau_k \left\{ \left(h_k^{*j} + \sum_{i=H,L} h_k^{*j} (i)\right) + R \sum_{i=H,L} x_y^{*k} (i) - x_y^{*k} \hat{s}_j \right\}. \end{aligned} \tag{9}$$

where $\hat{s}_j = s_j - \left(s_y x^j + s_{y^H} x^j (H) + s_{y^L} x^j (L)\right)$, is a form of 'compensated savings', $h_j^k = x_j^k - x_y^k x^j$, and $h_j^k (i) = x_j^k (i) - x_y^k (i) x^j (i)$ (for $i = H, L$.) are Hicksian demands, where it is state by state utility that is held fixed.

Let us consider the expression above. The right hand side is just the discouragement of consumption of good j, where by 'discouragement' one should understand the linear approximation of the reduction in compensated demand induced by the tax system.[13] The discouragement of consumption of good i has two components: the direct effect on compensated demands in each state of the world, and the indirect effect, through savings.

As for the left hand side, it is the change in consumption of good j, both in youth and in adulthood as caused by the choice of an optimal off-equilibrium level of savings.

3.0.4 On the Uniform Tax Prescription

In searching for conditions that deliver AS, our strategy is to suppose that it holds, i.e., taxes are uniform, and verify what conditions are needed for the derived expressions to be satisfied.

[13] See Cremer et al. (2001).

Along these lines, assume that $\mathbf{p}=\mathbf{1}$. Symmetry and homogeneity of Hicksian demands guarantee that, at $\mathbf{p}=\mathbf{1}$, $\sum_k \left(h_k^j+\sum_{i=H,L} h_k^j(i)\right)=0$, while $\sum_k x_y^k = \sum_k x_y^k(H)=\sum_k x_y^k(L)=1$ as a consequence of Engel's aggregation - which holds for each period and each state of nature, conditional on chosen level of income.

Expression (9), thus collapses to

$$2\mu\left[U_y^2\left(x^{2j}-x^{*j}\right)+U_y^2(L)\left(x^{2j}(L)-x^{*j}(L)\right)\right]=\lambda(R-1)\hat{s}_j$$

Notice also that, under strategy σ^2, the agent will get y^L, no matter what type she turns out to be. Then $U_y^2=RU_y^2(L)$, in which case,

$$R\left(x^{2j}-x^{*j}\right)+\left(x^{2j}(L)-x^{*j}(L)\right)=\frac{\lambda(R-1)}{2\mu U_y^2(L)}\hat{s}_j. \tag{10}$$

Therefore, we need $\left[R\left(x^{2j}-x^{*j}\right)+\left(x^{2j}(L)-x^{*j}(L)\right)\right]/\hat{s}_j$ to be independent of j, for expression (10) to hold.

We state this more formally in proposition 2, where an expression for $\hat{s}_j$ in terms of a more familiar object is also provided. It is in some sense, the most important result of this paper, in that it shows that separability alone is not sufficient to deliver the uniform tax prescription of Atkinson and Stiglitz.

Proposition 2 *There is a R that satisfies equation (10) only if*

$$\frac{R\left(x^{2j}-x^{*j}\right)+\left(x^{2j}(L)-x^{*j}(L)\right)}{U_y^*(H)\left(x_y^{*j}(H)-x_y^{*j}\right)+U_y^*(L)\left(x_y^{*j}(L)-x_y^{*j}\right)} \tag{11}$$

are constant across goods.

Were we in a traditional Mirrlees' setup and separability alone would do the job. Nonetheless, the condition required for uniform taxation to be optimal in Proposition 2 is *in addition* to separability. Preferences must be such that marshallian demands satisfy constancy across goods of (11).

To understand what this condition implies, we recall that $x^{2j}-x^{*j}$ is the difference in first period consumption of good for an agent who chooses strategy 2 and the analogous choice for an agent who chooses the truthful strategy. Because in the first period types have not been revealed, it is only through differences in savings that consumption choices are affected.

Similarly, $x^{2j}(L) - x^{*j}(L)$ is the effect of choosing different strategies on the consumption of good j for an agent who gets the after tax income of a low productivity agent. Separability guarantees that the amount of leisure an agent gets does not affect demand for goods *conditional on a given level of income*. The last observation is key in understanding our result. First, this means that consumption pattern of a low productivity agent and a high productivity agent who claims to be of low productivity are identical.

However, contrary to a situation where after tax income may be taken as identical to income available for consumption of goods, without loss in generality, in our model, savings are not controlled by the government and are added to after tax labor income to determine how much income the agent has to spend on consumption.

Because it is the differences in savings according with the different strategies that determine whether expression (11) is expected to be constant across goods, and ultimately determine the validity of AS, we shall now investigate how savings differ for the tow relevant strategies. Towards this end, we offer the following claim.

Claim 3 *If leisure is normal, then* $s^2 > s^*$.

As said before, the separability and convexity of the utility function guarantee that leisure is a normal good. Agents who anticipate that they will always announce to be of a low type, will in this case choose to save less than agents who opt for a truthful strategy.

Because they have higher disposable income for consumption in period two and lower in period one, expression (11) will also differ, in general, if income elasticities of demand differ across goods.

Intuitively, when an agent decides that she will announce a low type no matter what, she increases her savings according to claim 3. When uncertainty is finally revealed, she will announce L, but will have more income than a type L who saved the amount compatible with a truthful announcement strategy. If income effects differ across goods, the pattern of consumption will be altered by her savings decision, and will signal deviant behavior.

Increased savings means more second period (and less first period) consumption of at least one good, which signals the agent's lying. One may then wonder why this does not suffice for breaking down AS. The point here is that capital income taxation handles this part of the effects of deviant behavior.

Note also that an analogous expression obtains in the case where it is constraint (5) rather than (4) that binds at the optimum. The only difference being that 3 substitutes for 2 and H substitutes for L as superscripts and arguments, respectively. Hence, the violation of AS *is not* dependent on the particular assumption about the direction of binding IC constraints. Tax prescriptions *are*.

To summarize, if goods have different income elasticity of demand, independence condition (11) is (generically) not satisfied, despite our having imposed separability. Hence, AS is overturned.

The statement in the previous paragraph provides us with a hint concerning what condition on preferences we should expect to yield the optimality of uniform taxation: homotheticity. This is, indeed, the case.

Corollary 4 *If preferences are separable and homothetic, then uniform taxation of goods is optimal.*

At this point it is interesting to compare our results to the ones in Cremer et al. (2001, p. 709). In their paper it is assumed that agents have different endowments of a certain good k which is not observed by the government. However, because agents that have higher endowment are richer, they increase more (less) then proportionally the consumption of luxury goods (necessities) when compared to agents that have lower endowments. This helps the government in identifying deviant behavior. Homotheticity guarantees that all income elasticities are identical, and increase in consumption is proportional for all goods, delivering AS in their paper.

In our case, the 'higher endowment' only appears off the equilibrium path. Yet, it generates the same type of prescriptions that arise in their model. The exogenous extra dimension of heterogeneity is not needed in our model. In fact, agents are heterogeneous here only in what regards their productivity, as in Atkinson and Stiglitz (1976) and Mirrlees (1976).

There is, however, one sense in which it may be argued that we have introduced another dimension of heterogeneity. Along the equilibrium path agents only differ in their productivities. However, off-equilibrium agents add another dimension of heterogeneity in their second period 'endowments' very much like in Cremer et al. (2001). This explains why we get similar prescriptions in a model where an extra dimension of heterogeneity is not imposed; it is endogenously generated.

It must also be said that it is not the linearity restriction on commodity taxation - not imposed by Atkinson and Stiglitz (1976) - that drives the result, since uniform taxes *are linear*.

3.0.5 Capital Income Taxation

The marginal rate of transformation between current and future consumption has been assumed to be constant and equal to one, as a normalization. Despite its convenience in simplifying the problem this assumption poses a conceptual problem under the overlapping generations interpretation that we have been advocating.

The informational structure that underlies our model implies that, for any positive tax rate on capital income, the young and the elderly would transfer resources, and no aggregate savings would be observed in equilibrium. Positive taxation of capital income would not be feasible.

Notice that, strictly speaking, this is simply a consequence of the specific normalization we have adopted. For any value greater than one for the marginal rate of transformation, there would be an upper bound on the capital income tax rate. Even this, however, is overcome once the linear specification on technology for transferring goods across time is dropped. A concave production function would get around this issue at the cost of greatly increasing the algebraic burden.

We sidestep this issue and do not limit R to be greater than 1. Hence, assuming homotheticity, which we have shown guarantees uniform taxation in both periods, we shall now discuss the issue of optimal capital income taxation.

Unfortunately, because the two sides of (10) vanish at the optimum when preferences are homothetic we cannot use this expression to evaluate the sign of $R-1$.

In appendix A.2 we derive the first order conditions with respect to R, assume homotheticity, and show that the sign of $R-1$ is the same as the sign of $s^{*}-s^{2}$.

Next we recall claim 3 and the fact that our assumptions of preferences guarantee normality of leisure to show that $R<1$, in our model, as formally stated in the next proposition.

Proposition 5 *If preferences are separable and homothetic: i) goods are uniformly taxed, and; ii) marginal tax rate on capital income is positive.*

The above proposition summarizes the main results in this section. First, AS no longer holds in this setup. It is never too much to emphasize that this result does not hinge on the specific assumption concerning which constraint binds at the optimum. Second, the capital income tax rate is positive at the optimum. This is a novel result though related to other findings in multi-period Mirrlees' settings.

In what regards capital income taxation, the sign of the marginal tax depends on which specific IC constraint binds at the optimum. The logic is straightforward. The role of commodity and capital income taxation in the presence of an optimally designed non-linear labor income tax schedule, is to relax the IC constraints.

If it is the strategy of always announcing H, rather than always announcing L, that must be discouraged, the off-equilibrium level of savings is lower than that compatible with the truthful strategy. To relax the IC constraint one ought to subsidize savings, for these are 'under-consumed' by agents who opt not to abide by the rules.

It is worth also mentioning that, when the tax function is allowed to depend on information available only at the moment return on capital is realized, Albanesi and Sleet (2003) show that expected capital income taxation is zero. To create a bias towards current consumption and away from future consumption, according to the inverse Euler equation found in these papers, the tax function is such that the after tax return on capital is positively related to labor income, thus becoming a risky asset.

Here, however, tax rate on capital is positive and revenue is raised by means of this instrument.

4 Conclusion

In this paper, we investigate the properties of a tax system where the three main tax bases are explored; we have a non-linear labor income tax, and linear taxes on both goods and capital income.

The simultaneous appearance of three features, never put together in a single optimal taxation framework, to the best of our knowledge, results in the problem being quite non-standard. First, agents do not know their future productivities at the time savings decisions are made. Second, savings are not observable, so we are restricted to anonymous taxes on capital income as in Cremer et al. (2003). Unlike them, however, we add a third aspect.[14] We let savings affect 'ex-post' preferences in the $y \times Y$ space - i.e. change the (conditional) indirect utility.

These latter two elements: non-observability of first period choices and choices affecting preferences, render the problem a much harder one to characterize, since the set of implementable allocations, from a second period perspective, becomes endogenous.

This is easy to understand. Once preferences over net income y and supply of efficiency units, Y, are changed, the ranking of two bundles may be inverted, and what was incentive compatible for one specific 'ex-ante' choice may not to be for another. The fact that agents anticipate their pattern of announcements conditional on realized types, means that they will manipulate their preferences by means of convenient choices of savings, to take advantage of the possibilities made available by the second period budget set. Off-equilibrium savings become crucial in defining the set of feasible allocations, when viewed from a second period perspective. Implementation thus requires a different set of incentive constraints, which we impose in order to derive policy prescriptions.

We show that this form of non-observability generates a violation of Atkinson and Stiglitz' (1976) uniform taxation prescription. Homotheticity must be added to separability for the result to hold.

When capital income taxation is investigated sharp results are harder to come about. The problem here is that the assumptions usually adopted in the optimal taxation literature do not allow for determining which incentive constraints bind at the optimum in this model. Still, in the 'normal cases', taxation of savings (or subsidization of early consumption) is shown to be optimal.

From a purely theoretical perspective, this result resemble that found in Golosov

[14] The third aspect appears in all dynamic taxation literature where a Mirrlees' approach is used. However, savings are assumed to be observed in all papers, so far.

et al. (2003), Cremer and Gahvari (1995) and da Costa and Werning (2000). Unlike them, and because we only consider linear taxes on capital return, the encouragement of consumption early in life is done with a positive marginal tax rate, not a tax schedule that 'creates risk' in a risk free asset.

A Appendix

A.1 Proofs

Proof of Claim 1. At the equilibrium level of savings, define the following set, for each allocation (y, Y),

$$\mathbb{Z}_{+}^{H}(y,Y) \equiv \left\{ (y', Y') \in R_{+}^{2}; (x', y') \succsim_{H} (y,Y) \right\}$$

That is the set of bundles preferred to (y, Y) by agent type H.
Similarly,

$$\mathbb{Z}_{+}^{L}(y,Y) \equiv \left\{ (y', Y') \in R_{+}^{2}; (x', y') \succsim_{L} (y,Y) \right\}$$

Define also

$$\mathbb{Z}_{+}(y,Y) \equiv \left\{ (y', Y') \in R_{+}^{2}; (x', y') > (y,Y) \right\},$$

the set of bundles for which both quantities are at least as great as (y, Y), with at least one strictly greater. We can see that

$$\mathbb{Z}_{+}^{L}(y,Y) \cap \mathbb{Z}_{+}(y,Y) \subset \mathbb{Z}_{+}^{H}(y,Y) \cap \mathbb{Z}_{+}(y,Y) \tag{12}$$

Let

$$m^{k}(i|j) \equiv -\frac{U_{Y}\left(\mathbf{p}, y^{i} + Rs^{k}, Y^{i}, w^{j}\right)}{U_{y}\left(\mathbf{p}, y^{i} + Rs^{k}, Y^{i}, w^{j}\right)}$$

be the marginal rate of substitution for an agent of type j who announces to be of type i, given that she follows strategy k, and $m^{k}(\cdot|j)$ a generic allocation, for type j, at the level of savings corresponding to strategy k.
Without loss, take a path starting at (y, Y)such that

$$\begin{aligned} dU(H) &= U_{y}(\cdot|H)\left[dy - m(\cdot|H)\, dY\right] \geq 0 \Rightarrow \\ dy &= m(\cdot|H)\, dY \end{aligned}$$

In this case,

$$dU(L) = U_y(\cdot|L)\left[m(\cdot|H) - m(\cdot|L)\right]dY < 0$$

for $dY > 0$. If we define $\mathbb{Z}_-(y,Y) \equiv \left\{(y',Y') \in R_+^2; (x',y') < (y,Y)\right\}$, analogously, the same procedure shows that,

$$\mathbb{Z}_+^H(y,Y) \cap \mathbb{Z}_-(y,Y) \subset \mathbb{Z}_+^L(y,Y) \cap \mathbb{Z}_-(y,Y) \tag{13}$$

Notice from (12) and (13) that, at any level of savings, if an allocation is (y',Y') is preferred over an allocation (y,Y), by an agent of type H, then, the ordering can only be reversed by an agent of type L, for the same level of savings, $(y',Y') > (y,Y)$. Now take two allocations $\left(y^L,Y^L\right)$ and $\left(y^H,Y^H\right)$ such that $\left(y^H,Y^H\right) < \left(y^L,Y^L\right)$. From, (12) and (13), either $\left(y^L,Y^L\right) \in \mathbb{Z}_+^H(y,Y)$, or $\left(y^H,Y^H\right) \in \mathbb{Z}_+^H(y,Y)$. In either case, the allocation is not implementable. ■

Proof of Claim 3. Let

$$\mathbb{Z}_+^{H^*}\left(y^H,Y^H\right) \equiv \left\{\left(y',Y'\right) \in R_+^2; (x',y') \succsim_{H^*} \left(y^H,Y^H\right)\right\},$$

at the equilibrium level of savings, s^*, and

$$\mathbb{Z}_+^{H^2}\left(y^H,Y^H\right) \equiv \left\{\left(y',Y'\right) \in R_+^2; (x',y') \succsim_{H^2} \left(y^H,Y^H\right)\right\},$$

at the optimal level of savings for strategy 2, s^2.
Because the allocation is implementable we know form claim 1 that $\left(y^L,Y^L\right) \in \mathbb{Z}_-\left(y^H,Y^H\right)$ and $\left(y^L,Y^L\right) \notin \mathbb{Z}_+^{H^*}\left(y^H,Y^H\right)$. On the other hand, from the definition of strategy 2 it must be the case that $\left(y^L,Y^L\right) \in \mathbb{Z}_+^{H^2}\left(y^H,Y^H\right)$. However, if $s^2 < s^*$ normality of leisure implies that

$$\mathbb{Z}_+^{H^2}\left(y^H,Y^H\right) \cap \mathbb{Z}_-\left(y^H,Y^H\right) \subset \mathbb{Z}_+^{H^*}\left(y^H,Y^H\right) \cap \mathbb{Z}_-\left(y^H,Y^H\right).$$

A contradiction. ■

Proof of Corollary 4. If preferences are homothetic, $p^j x^{2j} = \omega^j\left(y - s^2\right)$, $p^j x^{*j} = \omega^j\left(y - s^*\right)$, $p^j x^{2j}(L) = \omega^j\left(y^L + Rs^2\right)$, $p^j x^{*j}(L) = \omega^j\left(y^L + Rs^*\right)$, and $p^j x_y^j(i) = \omega^j$, $i = H, L$, where ω^j is the (constant) proportion of income spent on good j.

In this case, $\hat{s}_j$ and the left hand side of (10) both vanish for all j, which guarantees that the condition is satisfied. ■

Proof of Proposition 2. If preferences are homothetic, $\mathbf{p} = \mathbf{1}$ and the first order condition with respect to R is

$$(1+\mu)\left(U_y^*(H)+U_y^*(L)\right)s^* - \mu\left(U_y^2(L|H)+U_y^2(L)\right)s^2 = \\ -\lambda \textstyle\sum_k \left[x_y^k s_R + \left(s^* - R s_R^*\right)\sum_{i=H,L} x_y^k(i)\right]$$

Multiply (17) and (18) by s^* and consider the fact that $U_y^2(L|H) = U_y^2(L)$, under separability, to get

$$\mu U_y^2(L)\left[s^* - s^2\right] = \lambda \textstyle\sum_k \left[-x_y^k + R\sum_{i=H,L} x_y^k(i)\right] \hat{s}_R^*$$

We have seen that homotheticity generates as optimal prescription $\mathbf{p} = \mathbf{1}$, as well as $x_y^k = x_y^k(H) = x_y^k(L) = 1$. Hence,

$$\mu U_y^2(L)\left[s^* - s^2\right] = \lambda\left[R-1\right]\hat{s}_R^*$$

We now invoke the fact that $\hat{s}_R$ is positive - proven in appendix A.3 - to show that $R < 1$ if and only if $s^* < s^2$.
Finally, we use claim 3 to get ii). ■

A.2 Optimal Commodity Tax Formulae

The first step towards evaluating AS is the derivation of optimal commodity tax formulae, which we start by writing the Lagrangian, $\mathcal{L} =$

$$(1+\mu)\left[2U^* + U^*(H) + U^*(L)\right] - \mu\left[2U^2 + U^2(L) + U^2(L|H)\right] + \\ \lambda\left[2Y + Y^H + Y^L - 2\textstyle\sum_k x^k - \sum_k\sum_{i=H,L} x^k(i)\right].$$

To derive the necessary conditions for the optimum, we start by differentiating the Lagrangian with respect to p_j and setting it equal to 0 ($\partial\mathcal{L}/\partial p_j = 0$.)

$$\left[2U_j^* + U_j^*(H) + U_j^*(L)\right](1+\mu) - \mu\left[2U_j^2 + U_j^2(L|H) + U_j^2(L)\right] - \\ \lambda\textstyle\sum_k\left[x_j^k + \sum_{i=H,L} x_j^k(i) + \left(-x_y^k + \sum_{i=H,L} R x_y^k(i)\right)s_j\right] = 0. \qquad (14)$$

where the derivatives of functions s and x^k $(k = 1, ..n)$ with respect both to prices and income are evaluated at the equilibrium choices associated to the truth telling strategy.

To interpret expression (14), we rewrite the left hand side of this expression by using the (conditional) Roy identity, $U_j = U_y x^j$.

That is, we note that

$$\begin{aligned}&\left[2U_j^* + U_j^*(H) + U_j^*(L)\right](1+\mu) - \mu\left[2U_j^2 + U_j^2(L|H) + U_j^2(L)\right] = \\ &\quad (1+\mu)\left[2U_y^* x^{*j} + U_y^*(H)\, x^{*j}(H) + U_y^*(L)\, x^{*j}(L)\right] \\ &\quad - \mu\left[2U_y^2 x^{2j} + U_y^2(L|H)\, x^{2j}(L|H) + U_y^2(L)\, x^{2j}(L)\right] \qquad (15)\end{aligned}$$

Now, consider the first order conditions, with respect to y,

$$2U_y^*(1+\mu) - \mu 2U_y^2 = \lambda \sum_k \left[x_y^k(1-s_y) + R\sum_{i=H,L} x_y^k(i)\, s_y\right], \qquad (16)$$

with respect to y^H,

$$2U_y^*(H)(1+\mu) = \lambda \sum_k \left[x_y^k s_{y^H}^* + x_y^k(H)\left(1 + R s_{y^H}\right) + R x_y^k(L)\, s_{y^H}\right], \qquad (17)$$

and with respect to y^L,

$$\begin{aligned}&2U_y^*(L)(1+\mu) - \mu\left[U_y^2(L|H) + U_y^2(L)\right] = \\ &\lambda \sum_k \left[x_y^k s_{y^L} + x_y^k(H)\, s_{y^L} + R x_y^k(L)\left(1 + R s_{y^L}\right)\right]. \qquad (18)\end{aligned}$$

Multiply (16) by x^{*j}, (17) by $x^{*j}(H)$ and (18) by $x^{*j}(L)$ to get, respectively,

$$\begin{aligned}2x^{*j}\left[U_y^*(1+\mu) - \mu U_y^2\right] &= \sum_k x_y^k x^{*j} + \\ &\quad \lambda \sum_k \left[-x_y^k + R\sum_{i=H,L} x_y^k(i)\right] s_y x^{*j}\end{aligned}$$

and

$$\begin{aligned}x^{*j}(H)\, U_y^*(H)(1+\mu) &= \lambda \sum_k x_y^k(H)\, x^j(H) + \\ &\quad \lambda \sum_k \left[-x_y^k + R\sum_{i=H,L} x_y^k(i)\right] s_{y^H} x^j(H),\end{aligned}$$

and

$$\begin{aligned}&x^{*j}(L)\left[2U_y^*(L)(1+\mu) - \mu\left[U_y^2(L|H) + U_y^2(L)\right]\right] = \\ &\lambda \sum_k x_y^k(L)\, x^{*j}(L) + \lambda \sum_k \left[-x_y^k + R\sum_{i=H,L} x_y^k(i)\right] s_{y^L} x^{*j}(L),\end{aligned}$$

Now, add the three to obtain

$$
\begin{aligned}
&(1+\mu)\left[2U_y^* x^{*j} + \sum\nolimits_{i=H,L} U_y^*(i)\, x^{*j}(i)\right] - \\
&\mu\left[2U_y^2 x^{*j} + \left(U_y^2(L|H) + U_y^2(L)\right) x^{*j}(L)\right] = \\
&\lambda \sum\nolimits_k \left(s_y x^j + \sum\nolimits_{i=H,L} s_{y^i} x^j(i)\right)\left[R\sum\nolimits_{i=H,L} x_y^k(i) - x_y^k\right] + \\
&\lambda \sum\nolimits_k x_y^k x^j + \sum\nolimits_{i=H,L} x_y^k(i)\, x^j(i) \qquad (19)
\end{aligned}
$$

We can now substitute (19) in (14), using also (15) to get

$$
\begin{aligned}
&\mu[2U_y^2\left(x^{2j} - x^{*j}\right) + U_y^2(L|H)\left(x^{2j}(L|H) - x^{*j}(L)\right) + \\
&U_y^2(L)\left(x^{2j}(L) - x^{*j}(L)\right) = \lambda\sum\nolimits_k \left(h_j^k + \sum\nolimits_{i=H,L} h^k(i)\right) + \\
&\lambda \hat{s}_j \sum\nolimits_k \left[-x_y^k + R\sum\nolimits_{i=H,L} x_y^k(i)\right] \qquad (20)
\end{aligned}
$$

where $\hat{s}_j = s_j - \left(s_y x^j + s_{y^H} x^j(H) + s_{y^L} x^j(L)\right)$, $h_j^k = x_j^k - x_y^k x^j$, and $h_j^k(i) = x_j^k(i) - x_y^k(i)\, x^j(i)$ (for $i = H, L.$).

Finally we use the fact that separability implies $U_y^2(L|H) = U_y^2(L)$ to write (20) as

$$
\begin{aligned}
&2\mu\left[U_y^2\left(x^{2j} - x^{*j}\right) + U_y^2(L)\left(x^{2j}(L) - x^{*j}(L)\right)\right] = \\
&\lambda\sum\nolimits_k\left(h_j^k + \sum\nolimits_{i=H,L} h_j^k(i)\right) + \lambda\sum\nolimits_k\left[-x_y^k + R\sum\nolimits_{i=H,L} x_y^k(i)\right]\hat{s}_j.
\end{aligned}
$$

finally notice that, homogeneity and symmetry of hicksian demand implies

$$
\sum\nolimits_k\left(h_j^k + \sum\nolimits_{i=H,L} h_j^k(i)\right) = -\sum\nolimits_k\left(\tau_k h_k^j + \sum\nolimits_{i=H,L} \tau_k h_k^j(i)\right).
$$

A.3 The effects of price variations in Savings

We cannot understand what is necessary for this condition to be satisfied without, first, understanding $\hat{s}_j$. Let us, then, explore the its meaning.

Consider the following problem:

$$
\max_s \left\{2U(\mathbf{p}, y - s, Y, w) + \sum\nolimits_{i=H,L} U\left(\mathbf{p}, y^i + Rs, Y^i, w^i\right)\right\} \qquad (21)
$$

The first and second order conditions for this problem are, respectively

$$
2U_y - R\left[U_y^*(H) + U_y^*(L)\right] = 0,
$$

and

$$\Theta \equiv 2U_{yy} + R^2 \left[U^*_{yy}(H) + U^*_{yy}(L)\right] < 0.$$

We may then use this to find s_j :

$$\begin{aligned}\Theta \frac{ds}{dp_j} &= 2U_{yj} - R\left[U^*_{yj}(H) + U^*_{yj}(L)\right] \\ &= -2U_{yy}x^j + \sum\nolimits_{i=L,H} RU_{yy}(i)\, x^j(i) - U_y x^j_y + R\sum\nolimits_{i=L,H} U_y(i)\, x^j_y(i)\end{aligned}$$

where, once again, Roy's identity was used.

The same procedure for s_y, s_{y^H} and s_{y^L}, yields

$$\frac{ds}{dy} = -\frac{-2U_{yy}}{\Theta}, \quad \frac{ds}{dy^H} = -\frac{RU_{yy}(H)}{\Theta}, \text{ and } \frac{ds}{dy^L} = -\frac{RU_{yy}(L)}{\Theta}, \text{ respectively.}$$

Combining these, it is readily seen that

$$\begin{aligned}\frac{ds}{dp_j} &= -\frac{ds}{dy}x^j - \frac{ds}{dy^H}x^j(H) - \frac{ds}{dy^L}x^j(L) \\ &\quad -\frac{2U_y x^j_y - R\left[U_y(H)\, x^j_y(H) + U_y(L)\, x^j_y(L)\right]}{\Theta}.\end{aligned}$$

Hence,

$$\hat{s}_j = -\frac{2U_y x^j_y - R\left[U_y(H)\, x^j_y(H) + U_y(L)\, x^j_y(L)\right]}{\Theta}$$

Because $2U_y = R\left[U_y(H) + U_y(L)\right]$, we can finally write

$$\hat{s}_j = R\frac{U_y(H)\left(x^j_y(H) - x^j_y\right) + U_y(L)\left(x^j_y(L) - x^j_y\right)}{\Theta} \tag{22}$$

Back to program (21), we adopt the same procedure we have been adopting so far to get

$$\frac{ds}{dR} = \frac{\left[U^*_y(H) + U^*_y(L)\right] + R\left[U^*_{yy}(H) + U^*_{yy}(L)\right]s}{\Theta}.$$

Next, we define $\hat{s}_R \equiv s_R - s_{y^H}s - s_{y^L}s$.

$$\hat{s}_R = \frac{U^*_y(H) + U^*_y(L)}{\Theta} > 0$$

References

[1] Albanesi, S. and C. Sleet "Dynamic optimal taxation with private information," Discussion Paper - Institute for Empirical Macroeconomics 140, Federal Reserve Bank of Minneapolis (2003).

[2] Atkinson A. and J. Stiglitz "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6, (1976) 55-75.

[3] Brito, D., J. Hamilton, S. Slutsky and J. Stiglitz "Dynamic Optimal Income Taxation with Government Commitment" *Journal of Public Economics, 44,* (1991) *15-35.*

[4] Chiappori, P.-A., I. Macho, , P. Rey and B. Salanié, "Repeated Moral Hazard: the Role of Memory, Commitment and the Access to Credit Markets", *European Economic Review*, 38 (1994) 1527-53.

[5] Christiansen, V. "Which Commodity Taxes Should Supplement the Income Tax?" *Journal of Public Economics*, 24 (1984) 195-210.

[6] Cooter, R. "Optimal Tax Schedules and Rater: Mirrlees and Ramsey". *American Economic Review,* 68 (1978) 756-768.

[7] Cremer, H and F. Gahvari, "Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy." *The Economic Journal* 105, (1995) 1165-1179

[8] Cremer, H and F. Gahvari, "Uncertainty, Commitment and Optimal Taxation." *Journal of Public Economic Theory*, 1 (1999) 51-70

[9] Cremer, H., Pestieau, P. and Rochet, J.-C. "Capital Income Taxation when Inherited Wealth is not Observable" *Journal of Public Economics*, 87: (2003) 2475-2490.

[10] Cremer, H., P. Pestieau, and J.-C. Rochet, "Direct versus Indirect Taxation: the Design of the Tax Structure Revisited" *International Economic Review*, 42 (2001) 781-799.

[11] da Costa, C. "Two Essays on Supplementary Commodity Taxation" PhD. Thesis, University of Chicago, 2001.

[12] da Costa, C. and I. Werning, 'A Note on Pre-Committed Goods'. mimeo., University of Chicago, 2000.

[13] da Costa, C. and I. Werning "On The Optimality of the Friedman Rule with Heterogeneous Agents and Non-linear Taxation" mimeo., University of Chicago, 2001.

[14] Diamond, P. and J. Mirrlees "A Model of Social Insurance with Variable Retirement" *Journal of Public Economics* 10 (1978) 295-336.

[15] Deaton, A. and J. Muellbauer, *Economics and Consumer Behavior* (Cambridge: Cambridge University Press, 1980)

[16] Dewatripont, M. "Commitment through Renegotiation-proof Contracts with Third Parties" *Review of Economic Studies* 55, (1988) 377-90.

[17] Guesnerie, R. "A Contribution to the Pure Theory of Taxation" (Cambridge: Cambridge University Press, 1995)

[18] Golosov, M., N. Kocherlakota, and A. Tsyvinski "Optimal Indirect and Capital Taxation" *Review of Economic Studies* 70, (2003) 569-587.

[19] Golosov, M., and A. Tsyvinski "Designing Optimal Disability Insurance" *Federal Reserve Bank of Minneapolis Research Department Working Paper* 293, 2003.

[20] Kocherlakota, N. "A Mirrlees Approach to Dynamic Optimal Taxation: Implication for Wealth Taxes and Asset Prices" *Federal Reserve Bank of Minneapolis Research Department Staff Report*, 2003.

[21] Konishi, H. "A Pareto-improving Commodity Tax Reform Under a Smooth Nonlinear Income Tax". *Journal of Public Economics*, 56 (1995) 413-446.

[22] Mirrlees, J. "An Exploration on the Theory of Optimum Income Taxation" *Review of Economic Studies* 38 (1971) 175-208.

[23] Mirrlees, J. “Optimal Tax Theory: A Synthesis,” *Journal of Public Economics*, **6** (1976) 327-58.

[24] Naito, H. ”Reexamination of Uniform Commodity Taxes Under a Non-Linear Tax System and Its Implications For Production Efficiency” *Journal of Public Economics*, 71 (1999) 165-188.

[25] Rogerson, W. ”Repeated Moral Hazard” *Econometrica*, 58 (1985) 69-76.

[26] Saez, E. ”The Desirability of Commodity Taxes with Non-linear Income Taxation and Heterogeneous Tastes” *Journal of Public Economics*, 83 (2002) 217-230.

[27] Stiglitz, J. E. ”Self-selection and Pareto Efficient Taxation,” *Journal of Public Economics*, 17 (1982) 213-240.

[28] Stiglitz, J. E. “Pareto Efficient and Optimal Taxation and the New New Welfare Economics” in *Handbook of Public Economics*, vol. II, (Amsterdam: North Holland, 1987), 991-1042.