Yet Another Reason to Tax Goods*

Carlos E. da Costa†

Getulio Vargas Foundation

February 3, 2006

Abstract

An important finding of the new dynamic public finance literature is the validity of Atkinson and Stiglitz' uniform commodity tax prescription in a dynamic Mirrleesian setting. However, this need not apply to the taxation of goods across time, i.e., the taxation of savings. We model an overlapping generations economy where, following the new dynamic public finance literature, we assume that information regarding agents' productivities is private and changes through time, but depart from the rest of the literature in assuming that the government does not have full control of agents' savings. Optimal commodity taxes are shown to depend on off-equilibrium savings, thus overturning Atkinson and Stiglitz' result. With regards to the taxation of savings, the inverse Euler equation found in the literature becomes here a positive marginal tax rate prescription.

Keywords: Optimal Taxation; Non-observable savings; Multi-period agency.

JEL Classification: H21, D82.

*I thank Luis Braido, Bev Dahlby, Lucas Maestri, Danilo Coelho, Tiago Severo and seminar participants at PUC-RJ, IBMEC, EPGE and the Latin American and European Meetings of the Econometric Society 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

1 Introduction

Golosov et al. (2003) have set in motion what is now known as the dynamic public finance literature[1] by analyzing optimal tax policies in a dynamic Mirrleesian setting where agents' skills are private information and evolve stochastically over time. They show that, if the government controls agents' savings, the uniform tax prescription of Atkinson and Stiglitz (1976) — henceforth, AS — extends to this economy. I.e., when preferences are separable between leisure and the other goods, there is no need for taxing goods: the income tax schedule will fully implement the second best allocation.

However, despite their considering agents with strongly separable temporary utility, one of the consequences of Atkinson and Stiglitz's (1976) result, the zero taxation of savings, need not survive in this more complex setting. Instead, an inverse Euler equation that implies a wedge in inter-temporal marginal rates of substitution characterizes optimal policies regarding taxation of savings.

In this paper we examine the problem of optimal supplementary commodity and inter-temporal taxation in a dynamic setting. We model an overlapping generations economy where, following the new dynamic public finance literature, we include an evolving information set for the agents as the crucial dynamic element. We depart from most of this literature, however, in assuming that private savings are not directly controlled by the government.[2] This simple and compelling restriction on policy instruments is sufficient to overturn AS, which is in contrast with the rest of the literature. As for taxation of goods across time, we show that the discouragement of savings, found in the new dynamic public finance literature, takes here the form of a positive tax rate on the returns of savings.

To understand the rationale of the result, we recall the derivation of an optimal non-linear income tax schedule. In a Mirrlees world, and without loss—as assured by the revelation principle—a direct mechanism is used to derive the (constrained) optimal allocations. Agents are asked their productivities, and are assigned a corresponding bundle comprised of gross income which they must supply and net income to which they are entitled. Truthful announcement of productivities is guaranteed if incentive

[1] The term was coined by Narayana Kocherlakota in his 2004 plenary speech at the Society for Economic Dynamics (SED) meeting.

[2] Golosov and Tsyvinski (2005) is a notable exception.

compatibility constraints are satisfied. Commodity taxation are only useful in such a world if it allows for relaxing the incentive compatibility constraints, which will be the case if the consumption pattern of an agent signals whether she is telling the truth or not. Separability rules out this possibility by making conditional demands independent of labor supply, hence, identical for liars and abiders.

The reason why this result survives in the dynamic setting of Golosov et al. (2003) is easy to understand. Separability rules out differences in choices *conditional on available income,* which is the relevant difference between mimickers and the agents they mimic when the government directly controls savings.

In this paper, however, agents who announce falsely do not have the same available income as the agents they pretend to be. The subtlety here is that agents anticipate deviating behavior, i.e., false announcements of productivities, and change their savings choice when compared to agents who intend to abide by the rules. Therefore, deviating behavior will be accompanied by different consumption choices generated by off-equilibrium savings. Contrary to Golosov et al. (2003), in our case, differential taxation of goods may help separating those who announce truthfully from those who announce falsely.

So far, we have not been specific about how off-equilibrium announcements affect savings. The point is that the violation of AS only depends on recognizing that choices do differ and *not* on how they differ. Knowing how they differ is, however, of paramount importance in determining how savings are to be taxed. We are able to prove that, if goods are uniformly taxed within each period, the only relevant deviating strategy is that of always announcing to be a low productivity agent. The consequence is that off-equilibrium savings are always greater than equilibrium savings. Punishing deviant behavior is thus accomplished by taxing savings.[3]

It is now well established that, in a dynamic mirrleesian setting with evolving information sets, optimal policies are characterized by the existence of a wedge between the agents' expected marginal utility of income and the marginal utility of income at youth.[4] This wedge, however, need not be associated with taxation of savings, as one

[3] In a parallel and independent work, Golosov and Tsyvinsky (2005) also consider a multi-period version of this problem and investigate optimal capital income taxation in a decentralized economy. If constraints bind in the 'usual direction', they show that it is optimal to tax capital income.

[4] This wedge, associated with an inverse Euler equation, was first derived in a repeated moral hazard

might be lead to think. In fact, it is possible to show—e.g., Kocherlakota (2005)—that the marginal tax rates are dependent on the realization of an agent's productivity, hence stochastic, and that the expected tax rate on savings need not be positive. In our model, because tax rates on savings are not state-dependent, the wedge is necessarily created by a positive marginal tax rate on savings.

The remainder of this paper is organized as follows. The economy is presented in Section 2. In Section 3, the government's program is written as a mechanism design problem and solved. The first order conditions found in Section 3 are evaluated in the steady-state in Section 4, where Atkinson and Stiglitz (1976) uniform taxation result and inter-temporal taxation are discussed. Section 5 concludes. The formal derivations of the results of section 3 are found in appendix A.1, and some auxiliary results for Section 4, in A.2.

2 The Environment

We consider an overlapping generations economy where each generation is populated by a continuum of ex-ante identical expected utility maximizing agents who live for two periods: youth and adulthood.

Uncertainty. At youth, productivity, w, is identical for all agents and normalized to 1, while at adulthood, following Stiglitz (1982), we consider only two possible types: H (for high productivity) and L (for low), with $w_H > w_L$.

Uncertainty arises in this world because, in the first period, i.e., at youth, agents do not know their 'adult' productivities, w_i, $i = H, L$. They do know, however, the probability associated with any realization, which defines a state of nature for the agent since we shall be be ruling out any form of aggregate uncertainty.

That is, we take shocks to productivity to be i.i.d. and invoke the law of large numbers to equate the cross-sectional distribution of types with the probability distribution faced by each agent in the first period. Hence, π^i, $i = H, L$, will denote both the probability that an agent attributes to her being of productivity w^i when 'adult' and the share of adult agents with this productivity in the economy.

context by Rogerson (1985). In a two period Mirrlees setup, Cremer and Gahvari (1999) and da Costa and Werning (2000) are early references, while a fully dynamic characterization is first presented by Golosov et al. (2003).

Preferences and Technology. Preferences are represented by the von-Neumann Morgenstern utility function,[5]

$$u(\boldsymbol{x}_t) - \zeta(l_t) + \sum\nolimits_{i=H,L} \pi^i \left[u\left(\boldsymbol{x}_t^i\right) - \zeta\left(l_t^i\right)\right],$$

where π^i is the probability associated with state i, i.e., the realization of productivity w^i.

Consumption vectors for first period and state i of second period, are represented by $\boldsymbol{x}_t$ and $\boldsymbol{x}_t^i \in \mathbb{R}^n$, respectively, while l_t, $l_t^i \in \mathbb{R}$ represent the related labor supplies. Temporary utility, $u(\cdot) - \zeta(\cdot)$, is the same for both periods and all states of the world, with (additive) separability imposed for us to investigate Atkinson and Stiglitz's uniform taxation result as well as its consequence for the taxation of savings. Functions u and ζ are strictly increasing while u and $-\zeta$ are strictly concave. Both functions are smooth.

Technology is very simple. Output is produced with a linear fixed coefficient technology that uses efficiency units of labor, Y to produce the various goods. We represent the associated marginal costs with the vector $\boldsymbol{p}$. Allowing for a more general technology with capital accumulation would make the discussion of taxation of savings identical to the taxation of capital. The cost of doing so is that we would have to follow part of the literature in assuming which IC constraints bind at the optimum.

Informational/Transaction Structure. Asymmetric information arises in this model because, once uncertainty is realized, each agent's productivity is only observed by the agent herself.

Despite the fact that firms cannot determine agents' productivities, they can observe how many efficiency units are supplied. I.e., they can observe $Y = lw$ but not l and w separately. Efficiency units are the relevant inputs in the production function, therefore, this type of asymmetric information does not create any inefficiency in production.

Efficiency units, supplied by the agents, are hired by firms to produce goods. Goods are sold back to agents with competition driving (producer) prices to marginal costs—thus the notation $\boldsymbol{p}$ for the marginal costs.

There is a key difference between trade in goods and trade in efficiency units. Agents are assumed to be able to conduct (non-observed) side trade of goods at no transaction

[5] Time subscript, t, when attached to elements of an allocation, represents the generation to which the allocation is associated. When attached to a price, however, it represents the period in which these prices prevail. Bold is used to represent vectors, where prices are row and quantities, column vectors.

cost, while for efficiency units the transaction costs are assumed do be prohibitively high. Savings are also assumed not to be observed at an individual level. However, we follow Cremer et al. (2003) in assuming that the government has access to information on anonymous transactions regarding capital income.

The Government. A benevolent government who inhabits this economy maximizes a utilitarian social welfare function. However, the informational structure restricts the set of instruments that are available for its pursuing this objective. While labor income may be taxed non-linearly, side trade of goods rules out any form of non-linearity in their taxation. Similarly, savings cannot be individually taxed, however they may be subject to a withholding tax based on anonymous transaction.[6]

Therefore, a Government policy is a sequence of commodity tax vectors, $\{\boldsymbol{\tau}_t\}_{t=0}^{\infty}$, taxes on capital return, $\{\theta_t\}_{t=0}^{\infty}$, and (possibly) age dependent labor income tax functions $\left\{T_t(\cdot), \hat{T}_t(\cdot)\right\}_{t=0}^{\infty}$, and a corresponding sequence for the public debt that satisfies the government's flow budget constraint and a no-ponzi condition. While $T(\cdot)$ represents the income tax schedule for young agents, $\hat{T}(\cdot)$ is the relevant tax schedule for agents at their earnings capacity peak.

The Equilibrium Given a government policy, an equilibrium for this economy is a sequence of consumer prices, $\{\boldsymbol{q}_t\}_{t=0}^{\infty}$, after tax gross interest rates, $\{R_t\}_{t=0}^{\infty}$, and an allocation $\{\boldsymbol{x}_t, (\boldsymbol{x}_{i,t})_i, l_t, (l_{i,t})_i\}_{t=0}^{\infty}$ such that:
i) agents maximize utility, $\left\{\boldsymbol{x}_t, (\boldsymbol{x}_{i,t})_i, l_t, (l_{i,t})_i, s_t\right\}_t \in$

$$\begin{array}{ll} \arg\max u(\boldsymbol{x}) - \zeta(l) + \sum_i \pi^i \left[u(\boldsymbol{x}_i) - \zeta(l_i)\right] & \\ \text{s.t.} \quad \begin{cases} \boldsymbol{q}_t \boldsymbol{x} \leq l - T_t(l) - s, \text{ and} \\ \quad \boldsymbol{q}_{t+1}\boldsymbol{x}_i \leq l_i w_i - \hat{T}_{t+1}(l_i w_i) + R_t s \quad i = H, L \end{cases} \end{array} \tag{1}$$

ii) firms maximize profits;
iii) resource constraints are satisfied,

$$\boldsymbol{p}\boldsymbol{x}_t + \sum_i \pi^i \boldsymbol{p}\boldsymbol{x}_{i,t-1} \leq Y_t + \sum_i \pi^i Y_{i,t-1}, \forall t \tag{2}$$

and;
iv) the government's budget constraint is satisfied.

Walras' law guarantees that if *ii* and *iii* are satisfied so is *iv*. We may, therefore, leave it in the background.

[6] The same type of instrument found in Cremer et al. (2001, 2003), Golosov and Tsyvinsky (2005).

3 Optimal Taxation

The problem the government solves is to choose tax functions, $\left\{T_t(\cdot), \hat{T}_t(\cdot)\right\}_{t=0}^{\infty}$, consumer price sequences $\{\boldsymbol{q}_t \equiv \boldsymbol{p} + \boldsymbol{\tau}_t\}_{t=0}^{\infty}$, and net interest rate sequences, $\left\{R_t \equiv \tilde{R}_t + \theta_t\right\}_{t=0}^{\infty}$, that solve

$$\max \sum\nolimits_t \beta^t \left\{u(\boldsymbol{x}_t) - \zeta(l_t) + \sum\nolimits_i \pi^i \left[u\left(\boldsymbol{x}_t^i\right) - \zeta\left(l_t^i\right)\right]\right\},$$

subject to (1) and the resource constraint, (2).

Solving this problem directly may be a formidable task. Instead, we take advantage of the revelation principle to concentrate on a direct truthful mechanism. In the first period, an agent of generation t must supply Y_t efficiency units and is left with after-tax income y_t. All this before nature defines the agent's type. In the second period, after her individual productivity is realized, the agent chooses her bundle of net and gross income—y_t and Y_t, respectively—among those contained in the budget set, $\mathbb{B}_t \equiv \{(y_{i,t}, Y_{i,t})\}_{i=H,L}$, made available by the government for generation t.

Announcing her productivities is not all that an agent does, however. In the first period, in possession of y_t she chooses her preferred consumption bundle, given prices $\boldsymbol{q}_t = \boldsymbol{p} + \boldsymbol{\tau}_t$ and decides how much to save, s. These choices are represented by the indirect utility functions

$$v(\boldsymbol{q}_t, I_t) \equiv \max_{\boldsymbol{x}} u(\boldsymbol{x}) \text{ s.t. } \boldsymbol{q}_t \boldsymbol{x} \leq I_t, \tag{3}$$

with $\boldsymbol{x}(\boldsymbol{q}_t, I)$, denoting the corresponding (conditional) Marshallian demand, and by the inter-temporal optimization problem

$$\max_s v(\boldsymbol{q}_t, y_t - s) + \sum\nolimits_{i=H,L} \pi^i v\left(\boldsymbol{q}_{t+1}, \hat{y}_{i,t} + Rs\right),$$

where $\hat{y}_{i,t}$ is the after tax income corresponding to her choice from the set $\mathbb{B}_t$ when she turns out to be of productivity w_i.

We say that an allocation is implementable by a direct truthful mechanism, which we shall call simply *implementable*, if any agent—who has freely chosen how much to save in the first period—finds it in her best interest to always choose the bundle associated with her type, from the budget set, $\mathbb{B}_t$. This is guaranteed if each allocation satisfies the associated incentive compatibility, henceforth IC, constraints.

The problem here is that only considering IC constraints at the equilibrium level of savings will not suffice. Off-equilibrium savings must be taken into account because, even 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. This is the so called *double-deviation* problem.[7]

We handle this by noting that, when choosing how much to save, agents anticipate their announcements, conditional on the realization of types. If we, then, define strategies as rules that associate to each realization of productivity a specific action—in this case, an announcement—we need only to consider the incentive compatibility constraints for this strategy at the expected utility maximizing level of savings. The logic is straightforward. If at the optimal level of savings for a given strategy the agent finds it better to announce truthfully than to pursue this strategy, then she will never find it optimal to adopt this strategy, and truthful announcement is guaranteed.

Because there are only two levels of productivities, there are only four possible strategies, the first strategy, to always announce truthfully one's productivity, being the one we must induce the agents to adopt. To understand the main argument, define a strategy as a mapping from types to announcements $\sigma^k : \{H, L\} \to \{H, L\}$, then, $\sigma^k(j)$ is the announcement prescribed by strategy k if one realizes type j. It is trivial to verify that with only two types four different strategies are possible: *i)* $\sigma^*(i) = i$, for $i = H, L$; *ii)* $\sigma^o(i) = L$, for $i = H, L$; *iii)* $\sigma^1(i) = H$, for $i = H, L$, and; *iv)* $\sigma^2(i) = j \neq i$, for $i = H, L$. Strategy σ^*, prescribing truthful announcement for all realizations, is, as we said, the strategy we want to induce.

Each strategy defines a strictly concave savings problem. Therefore, we associate to each σ^k, $k = *, o, 1, 2.$ a unique optimal level of savings s_t^k,

$$s_t^k \equiv \arg\max_s \left\{ v(\boldsymbol{q}_t, y_t - s) + \sum\nolimits_{i=H,L} \pi^i v\left(\boldsymbol{q}_{t+1}, y_{\sigma^k(i),t} + R_t s\right)\right\}, \ k = *, o, 1, 2. \quad (4)$$

and a corresponding indirect utility function,

$$U^k\left(\boldsymbol{q}_t, \boldsymbol{q}_{t+1}, y_t, Y_t, (y_{i,t}, Y_{i,t})_{i=H,L}\right), \quad (5)$$

where Y_t, Y_t^H and Y_t^L were omitted from (4) since they play no role in the optimization with respect to s.

Before writing the governments' program a final intermediary result will be needed.

[7] For a discussion of this issue in a moral hazard setting see Chiappori et al. (1995). In a self-selection framework, Albanesi and Sleet (2005), Golosov and Tsyvinsky (2005) are important references.

In principle, we need to write five incentive constraints. The three incentive constraints that guarantee that the optimal strategy is to choose the truthful announcement—which we refer to as first period IC constraints—and the two IC constraints that guarantee that once uncertainty is realized agents will not change their minds and choose to announce falsely—the second period IC constraints. What we are able to show is that we need only be concerned with one first period IC constraints. All other will be slack at the optimum.

We shall first try and understand why the second period IC constraints are slack at any implementable allocation, i.e., why if it is not optimal to choose a strategy of announcing falsely and saving accordingly, it will never be optimal to announce falsely if one has chosen the equilibrium level of savings s^*. If all first period constraints are satisfied, then no alternative strategy yields a higher expected utility than the strategy of always announcing truthfully, even after one re-optimizes with respect to savings. This is true, in particular for all strategies where an agent deviates only in one state. Now, if, the expected utility under the alternative strategy, at its corresponding level of savings, is no greater than the expected utility of the truthful announcement, then the expected utility under the alternative is strictly lower at the equilibrium—i.e., the optimum for the truthful strategy—level of savings. But, at the equilibrium savings, the only difference between the expected utility of the truthful strategy and the alternative strategy is the utility in the state of the world where the deviation takes place.

Formally, rewrite the first period IC constraints as

$$v\left(\boldsymbol{q}_t, y_t - s_t^*\right) + \sum_{i=H,L} \pi^i \left[v\left(\boldsymbol{q}_{t+1}, y_{i,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{i,t}}{w_i}\right)\right]$$
$$\geq v\left(\boldsymbol{q}_t, y_t - s_t^o\right) + v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s_t^o\right) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y_{L,t}}{w_i}\right), \quad (6)$$

$$v\left(\boldsymbol{q}_t, y_t - s_t^*\right) + \sum_{i=H,L} \pi^i \left[v\left(\boldsymbol{q}_{t+1}, y_{i,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{i,t}}{w_i}\right)\right]$$
$$\geq v\left(\boldsymbol{q}_t, y_t - s_t^1\right) + v\left(\boldsymbol{q}_{t+1}, y_{H,t} + R_t s_t^1\right) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y_{H,t}}{w_i}\right), \quad (7)$$

and

$$v\left(\boldsymbol{q}_t, y_t - s_t^*\right) + \sum_{i=H,L} \pi^i \left[v\left(\boldsymbol{q}_{t+1}, y_{i,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{i,t}}{w_i}\right)\right]$$
$$\geq v\left(\boldsymbol{q}_t, y_t - s_t^2\right) + \sum_{j\neq i, i=H,L} \pi^i \left[v\left(\boldsymbol{q}_{t+1}, y_j + R_t s_t^2\right) - \zeta\left(\frac{Y_{j,t}}{w_i}\right)\right], \quad (8)$$

and assume that they are all satisfied.

Next, because s_t^o solves $\max_s \left\{ v\left(\boldsymbol{q}_t, y_t - s\right) + v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s\right)\right\}$ we have,

$$v\left(\boldsymbol{q}_t, y_t - s_t^o\right) + v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s_t^o\right) \geq v\left(\boldsymbol{q}_t, y_t - s_t^*\right) + v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s_t^*\right),$$

which implies from (6) that

$$v\left(\boldsymbol{q}_{t+1}, y_{H,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{H,t}}{w_H}\right) \geq v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{L,t}}{w_H}\right).$$

An analogous argument can be applied to (7) to show that

$$v\left(\boldsymbol{q}_{t+1}, y_{L,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{L,t}}{w_L}\right) \geq v\left(\boldsymbol{q}_{t+1}, y_{H,t} + R_t s_t^*\right) - \zeta\left(\frac{Y_{H,t}}{w_L}\right).$$

In passing, we should mention that, if the first period IC constraints are satisfied, the second period ones are usually satisfied with strict inequalities. If this is the case, i.e., if utility strictly lower at the equilibrium savings for those who deviate, the tax schedule will be interim inefficient: once the saving decision is made, agents would want the government to redesign the tax schedule. These results are akin to the ones found in the repeated moral hazard literature: the 'tax system' and, for that matter, any deterministic implementable contract, is not renegotiation-proof in the sense of Dewatripont (1988).[8]

As for the first period IC constraints, it turns out that the only constraint that binds at the optimum is the one which prevents the agents from adopting the strategy of always announcing to be of low productivity.

In a static model, the fact that only downward constraints are binding is an immediate consequence of single-crossing and the utilitarian motives of the government. Here, however, the proofs are a little more evolved—shown in appendix A.1—since one must keep track of off-equilibrium savings behavior. Moreover single-crossing alone cannot rule out the possibility of both constraints binding at the optimum without bunching.

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

The logic, however, is similar. Given the utilitarianism implicit in the maximization of expected utility, redistribution takes place from the high type to the low type. It is, then, the possibility of the high type announcing to be a low type that should be of concern here.

We are now ready to write the government's program. Using the definitions (4) and (5) we consider the problem of choosing sequences $\left\{R_t, \boldsymbol{q}_t, y_t, Y_t, (y_{i,t}, Y_{i,t})_{i=H,L}\right\}_{t=0}^{\infty}$, to solve

$$\max \sum \beta^t U^* \left(R_t, \boldsymbol{q}_t, \boldsymbol{q}_{t+1}, y_t, Y_t, (y_{i,t}, Y_{i,t})_{i=H,L}\right)$$

subject to the incentive compatibility constraints,

$$U^* \left(R_t, \boldsymbol{q}_t, \boldsymbol{q}_{t+1}, y_t, Y_t, (y_{i,t}, Y_{i,t})_i\right) \geq U^o \left(R_t, \boldsymbol{q}_t, \boldsymbol{q}_{t+1}, y_t, Y_t, (y_{i,t}, Y_{i,t})_i\right) \qquad [\beta^t \mu_t]$$

and the period by period resource constraint,

$$\begin{aligned} &\boldsymbol{p}\left[\boldsymbol{x}\left(\boldsymbol{q}_t, y_t - s^*\right) + \sum\nolimits_i \pi^i \boldsymbol{x}\left(\boldsymbol{q}_t, y_{i,t-1} + R_{t-1} s^*\right)\right] \\ &\qquad \leq F\left(K_t, Y_t + \sum\nolimits_i \pi^i Y_{i,t-1}\right) - (K_{t+1} - K_t) - \delta K_t. \qquad [\ \beta^t \lambda_t\ ] \end{aligned}$$

To find the optimal taxes, let $\partial_{k_t} f$ denote the partial derivative of a function f with respect to the price of good k at period t and let $\partial_I f$ denote the derivative with respect to disposable income (Note that $I = y - s$, in the first period and $I = y_i + Rs$, in the second period, state i.), whereas ∂_y (as well as ∂_{y_H} and ∂_{y_L}) denotes the derivative with respect to the after tax income in the savings function, $s_t\left(R_t, y_t, y_{H,t}, y_{L,t}, \boldsymbol{q}_t, \boldsymbol{q}_{t+1}\right)$.

Bearing this notation in mind, the first order conditions for this problem are, with respect to q_t^k,

$$\begin{aligned} &\beta \left\{(1+\mu_t)\,\partial_{k_t} U_t^* - \mu_t \partial_{k_t} U_t^o\right\} + \left(1+\mu_{t-1}\right)\partial_{k_t} U_{t-1}^* - \mu_{t-1}\partial_{k_t} U_{t-1}^o = \\ &\beta\lambda_t \boldsymbol{p}\left\{\partial_{k_t}\boldsymbol{x}_t^* - \partial_I \boldsymbol{x}_t^* \partial_{k_t} s_t^* + R_{t-1}\sum\nolimits_{i=H,L}\pi^i\left[\partial_{k_t}\boldsymbol{x}_{i,t-1}^* + \partial_I \boldsymbol{x}_{i,t-1}^*\partial_{k_t} s_{t-1}^*\right]\right\} - \\ &\lambda_{t-1}\boldsymbol{p}\partial_I \boldsymbol{x}_{t-1}^*\partial_{k_t} s_{t-1}^* + \beta^2\lambda_{t+1} R_t \sum\nolimits_{i=H,L}\pi^i \boldsymbol{p}\partial_I \boldsymbol{x}_{i,t}^*\partial_{k_t} s_t^* \end{aligned} \qquad (9)$$

with respect to y_t,

$$(1+\mu_t)\,\partial_{y_t} U_t^* - \mu_t \partial_{y_t} U_t^o = \lambda_t \boldsymbol{p}\left[\partial_I \boldsymbol{x}_t^* - \partial_I \boldsymbol{x}_t^* \partial_{y_t} s_t^*\right] + \lambda_{t+1}\beta R_t \sum\nolimits_i \pi^i \boldsymbol{p}\partial_I \boldsymbol{x}_{i,t}^* \partial_{y_t} s_t^*, \quad (10)$$

with respect to $y_{H,t}$,

$$\begin{aligned} &(1+\mu_t)\,\partial_{y_H} U_t^* - \mu_t \partial_{y_H} U_t^o = \\ &\qquad \beta\lambda_{t+1}\boldsymbol{p}\left[\pi^H \partial_I \boldsymbol{x}_{H,t}^* + R_t \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}_{i,t}^* \partial_{y_H} s_t^*\right] - \lambda_t \boldsymbol{p}\partial_I \boldsymbol{x}_t^* \partial_{y_H} s_t^*, \end{aligned} \qquad (11)$$

with respect to $y_{L,t}$,

$$(1+\mu_t)\,\partial_{y_L}U_t^* - \mu_t\partial_{y_L}U_t^o = \beta\lambda_{t+1}\boldsymbol{p}\left[\pi^L\partial_I\boldsymbol{x}_{L,t}^* + R_t\sum\nolimits_i \pi^i\partial_I\boldsymbol{x}_{i,t}^*\partial_{y_L}s_t^*\right] - \lambda_t\boldsymbol{p}\partial_I\boldsymbol{x}_t^*\partial_{y_L}s_t^*, \quad (12)$$

with respect to Y_t,

$$(1+\mu_t)\,\partial_Y U_t^* - \mu_t\partial_Y U_t^o = \lambda_t, \quad (13)$$

with respect to $Y_{H,t}$,

$$(1+\mu_t)\,\partial_{Y_H}U_t^* - \mu_t\partial_{Y_H}U_t^o = \pi_H\beta\lambda_{t+1}, \quad (14)$$

with respect to $Y_{L,t}$,

$$(1+\mu_t)\,\partial_{Y_L}U_t^* - \mu_t\partial_{Y_L}U_t^o = \pi_L\beta\lambda_{t+1}, \quad (15)$$

and with respect to R_t,

$$(1+\mu_t)\,\partial_R U_t^* - \mu_t\partial_R U_t^o = \beta\lambda_{t+1}\sum\nolimits_i \pi^i\boldsymbol{p}\partial_I\boldsymbol{x}_{i,t}^*\left(s_t^* + R_t\partial_R s_t^*\right) - \lambda_t\boldsymbol{p}\partial_I\boldsymbol{x}_t^*\partial_R s_t^*. \quad (16)$$

It is apparent from the expression above that a change in price q_t^k has direct effect on two generations and the resource constraint of three consecutive periods through its effects on savings of both generations. It is important also to realize that $\partial_{k_t}U_t^* \neq \partial_{k_t}U_{t-1}^*$, since for generation t, the change refers to first period prices while for generation $t-1$ it refers to second period prices. A similar reasoning applies to $\partial_{k_t}s_t^*$ and $\partial_{k_t}s_{t-1}^*$.

It is easy to verify that none of the first order conditions depend directly on time. I.e., it is possible to find stationary allocations and prices that satisfy all first order conditions above. This was the purpose of considering a stationary environment. Therefore, aside for possible non-convexities, the optimal policy is stationary. In this case, we have $\lambda_t = \lambda$, $\mu_t^k = \mu^k$, $\boldsymbol{q}_t = \boldsymbol{q}$, $y_t = y$, $Y_t = Y$, $(y_{i,t}, Y_{i,t})_{i=H,L} = (y_i, Y_i)_{i=H,L}$, and $R_t = R$.

4 Stationary Policies

In what follows we shall investigate the optimal stationary policies. First, we write

$$U^*\left(R, \boldsymbol{q}, \boldsymbol{q}', y, Y, (y_i, Y_i)_{i=H,L}\right) \text{ and } U^o\left(R, \boldsymbol{q}, \boldsymbol{q}', y, Y, y_L, Y_L\right)$$

to substitute for (5), at the two relevant strategies.

Here, we write $\boldsymbol{q}$ and $\boldsymbol{q}'$ to denote the price vectors faced by agents in the first and the second period of their lives. Under a stationary policy, $\boldsymbol{q} = \boldsymbol{q}'$, however we keep the distinction in the notation in order to keep track of the effect of changes in each moment of one's life. In any case, we shall be interested in changes in the price of each good k in every period. Hence, when we speak, for instance, on the effect of a change in the price of good k on a specific generation's utility we will be referring to $(\partial_k + \partial_{k'}) U^* (\boldsymbol{q}, \boldsymbol{q}', y, Y, (y_i, Y_i)_i)$.

Therefore, the only relevant incentive compatibility constraint is

$$v(\boldsymbol{q}, y - s^*) + \sum\nolimits_{i=H,L} \pi^i \left[v\left(\boldsymbol{q}', y_i + Rs^*\right) - \zeta\left(\frac{Y_i}{w_i}\right)\right]$$
$$\geq v(\boldsymbol{q}, y - s^o) + v\left(\boldsymbol{q}', y_L + Rs^o\right) - \sum\nolimits_i \pi^i \zeta\left(\frac{Y_L}{w_i}\right), \qquad [\ \beta^t \mu_t^o\ ]$$

where $s^* = s^*(R, \boldsymbol{q}, \boldsymbol{q}', y, y_H, y_L)$ and $s^o = s^o(R, \boldsymbol{q}, \boldsymbol{q}', y, y_L)$. We have, opened up the expected utility expression at each strategy to facilitate the exposition of the main arguments in the discussion that follows.

Next, we note that, in the steady-state, $\partial_{k_t} U_t^* = \partial_k v(\boldsymbol{q}, y - s^*)$, while $\partial_{k_t} U_{t-1}^* = \partial_{k_{t+1}} U_t^* = \sum_{i=H,L} \pi^i \partial_k v(\boldsymbol{q}', y_i + Rs^*)$, with similar expression for $\partial_{k_t} U_t^o$, and $\partial_{k_t} U_{t-1}^o = \partial_k v(\boldsymbol{q}', y_L + Rs^o)$. Analogous derivations are valid for $s^*(R, \boldsymbol{q}, \boldsymbol{q}', y, y_H, y_L)$ and for $s^o(R, \boldsymbol{q}, \boldsymbol{q}', y, y_L)$, which allows us to group the terms related to a single generation in (9), as follows,

$$\begin{aligned}
&(1+\mu)\partial_k \left[v(\boldsymbol{q}, y - s^*) + \sum\nolimits_i \pi^i v\left(\boldsymbol{q}', y_i + Rs^*\right)\right] \\
&-\mu \partial_k \left[v(\boldsymbol{q}, y - s^o) + v\left(\boldsymbol{q}', y_L + Rs^o\right)\right] \\
&\qquad = \lambda \boldsymbol{p} \left\{ \partial_k \left[\boldsymbol{x}(\boldsymbol{q}, y - s^*) + \beta \sum\nolimits_i \pi^i \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right] + \right. \\
&\qquad\quad \left. \partial_I \left[\boldsymbol{x}(\boldsymbol{q}, y - s^*) + \beta R \sum\nolimits_i \pi^i \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right] (\partial_k + \partial_{k'}) s^* \right\}. \quad (17)
\end{aligned}$$

Similarly, we re-write (10) as

$$\begin{aligned}
&(1+\mu)\partial_I v(\boldsymbol{q}, y - s^*) - \mu \partial_I v(\boldsymbol{q}, y - s^o) = \lambda \boldsymbol{p} \partial_I \boldsymbol{x}(\boldsymbol{q}, y - s^*) \\
&\quad - \lambda \boldsymbol{p} \left[\partial_I \boldsymbol{x}(\boldsymbol{q}, y - s^*) - \beta R \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right] \partial_y s^*, \qquad (18)
\end{aligned}$$

(11) as

$$\begin{aligned}
&(1+\mu)\pi^H \partial_I v\left(\boldsymbol{q}', y_H + Rs^*\right) = \lambda \beta \pi^H \boldsymbol{p} \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_H + Rs^*\right) - \\
&\quad - \lambda \boldsymbol{p} \left[\partial_I \boldsymbol{x}(\boldsymbol{q}, y - s^*) - \beta R \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right] \partial_{y_H} s^*, \qquad (19)
\end{aligned}$$

and (12) as

$$(1+\mu)\,\pi^L \partial_I v\left(\boldsymbol{q}', y_L + Rs^*\right) - \mu \partial_I v\left(\boldsymbol{q}', y_L + Rs^o\right) = \beta\lambda\pi^L \boldsymbol{p}\partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_L + Rs^*\right)$$
$$- \lambda \boldsymbol{p}\left[\partial_I \boldsymbol{x}\left(\boldsymbol{q}, y - s^*\right) - \beta R \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right] \partial_{y_L} s^*. \qquad (20)$$

The other first order conditions will not be needed for our purposes, so we do not write them down.

4.1 Optimal Commodity Taxes

To advance in the characterization of optimal commodity taxes we multiply expression (18) by $x_k\left(\boldsymbol{q}, y - s^*\right)$, (19) by $x_k\left(\boldsymbol{q}', y_H + Rs^*\right)$ and (20) by $x_k\left(\boldsymbol{q}', y_L + Rs^*\right)$. Next, we add the three and combine with (17) using Roy's identity, $x_k = -\,\partial_k v / \partial_I v$, and the Slutsky equation, $\partial_j x_k = \partial_j h_k - \partial_I x_k x_j$, to get

$$x_k\left(\boldsymbol{q}, y - s^*\right) - x_k\left(\boldsymbol{q}, y - s^o\right) + R^{-1}\left[x_k\left(\boldsymbol{q}', y_L + Rs^*\right) - x_k\left(\boldsymbol{q}', y_L + Rs^o\right)\right]$$
$$= \frac{\lambda}{\mu \partial_I v\left(\boldsymbol{q}, y - s^o\right)} \boldsymbol{p}\left\{\left[\partial_k \boldsymbol{h}\left(\boldsymbol{q}, u\right) + \beta \sum\nolimits_i \pi^i \partial_k \boldsymbol{h}\left(\boldsymbol{q}', u^i\right)\right]\right.$$
$$\left. + \left[\partial_I \boldsymbol{x}\left(\boldsymbol{q}, y - s^*\right) - R\beta \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right]\left(\partial_k + \partial_{k'}\right)\hat{s}^*\right\}, \qquad (21)$$

where

$$\left(\partial_k + \partial_{k'}\right)\hat{s}^* = \left(\partial_k + \partial_{k'}\right) s^* + \partial_y s^* x_k\left(\boldsymbol{q}, y - s^*\right) + \sum\nolimits_i \pi^i \partial_{y_i} s^* x_k\left(\boldsymbol{q}', y_i + Rs^*\right),$$

which is a form of 'compensated' savings.

By definition, $\boldsymbol{p} = \boldsymbol{q} - \boldsymbol{t}$. We may, then, use the the symmetry and homogeneity properties of Hicksian demands, $\boldsymbol{q}\partial_k \boldsymbol{h} = \boldsymbol{q}\partial_{\boldsymbol{q}} h^k = 0$, and the adding up property of marshallian demands, $\boldsymbol{q}\partial_I \boldsymbol{x} = 1$ to write the term inside the curly brackets in the right hand side of (21) in the form of the discouragement of consumption of good k,

$$-\boldsymbol{t}\left[\partial_k \boldsymbol{h}\left(\boldsymbol{q}, u\right) + \beta \sum\nolimits_i \pi^i \partial_k \boldsymbol{h}\left(\boldsymbol{q}', u^i\right)\right] + (1 - R\beta)\left(\partial_k + \partial_{k'}\right)\hat{s}^*$$
$$- \boldsymbol{t}\left[\partial_I \boldsymbol{x}\left(\boldsymbol{q}, y - s^*\right) - R\beta \sum\nolimits_i \pi^i \partial_I \boldsymbol{x}\left(\boldsymbol{q}', y_i + Rs^*\right)\right]\left(\partial_k + \partial_{k'}\right)\hat{s}^*. \quad (22)$$

Discouragement of consumption of good i is how we refer to the linear approximation of the reduction in compensated demand induced by the tax system, a terminology employed by Mirrlees. The fact that cross price effects are presents makes it the relevant measure of how the tax system treats each good. The discouragement expression has two components: the first one, also found in Mirrlees (1976), is captured by the Hicksian demand term $\partial_k \boldsymbol{h}$, which, due to symmetry, is equal to the gradient of h^k, $\partial_{\boldsymbol{q}} h^k$, while the second captures the indirect discouragement due to changes in savings.

How much the consumption of each good ought to be discouraged depends on the expression in the left hand side of (21). It compares the consumption of good k for an agent who chose to adopt the truthful announcement strategy with the consumption of the same good for an agent who chose to always announce to be of low productivity, both in the first period and in the second period, state L. Goods which consumption increase relatively more with deviant behavior are exactly those which must be more strongly discouraged, since this helps punishing off-equilibrium choices thus relaxing IC constraints.

4.1.1 Investigating AS

Expression (21) provides a general rule for optimal taxes. To evaluate the conditions under which uniform taxes are optimal, we rewrite (22) under uniform taxes, and check the restrictions on preferences that allow for (21) to be satisfied.

Homogeneity of demands guarantee that we may take uniform commodity taxes to mean zero taxes without loss in generality. In this case, (22) simplifies to

$$[1 - \beta R] \left(\partial_k + \partial_{k'}\right) \hat{s}^*.$$

When prices change, the marginal utility of consumption is potentially affected, even when this change is compensated by keeping real income constant. This change will alter savings choices and increase the marginal excess burden of taxes, when compared to the case where savings are directly controlled by the government. In fact, it is a simple, although a bit tedious, comparative statics exercise—see appendix A.2—to show that

$$\left(\partial_k + \partial_{k'}\right) \hat{s}^* = \kappa \sum_i \pi^i \partial_I v\left(\boldsymbol{q}', y_i + Rs^*\right) \left[\partial_I x_k\left(\boldsymbol{q}, y - s^*\right) - \partial_I x_k\left(\boldsymbol{q}', y_i + Rs^*\right)\right]$$

where κ is a positive scalar.

Hence, if $\partial_I x_k\left(\boldsymbol{q}, y - s^*\right) - \partial_I x_k\left(\boldsymbol{q}', y_i + Rs^*\right) \neq 0$, (21) implies that

$$\frac{x_k\left(\boldsymbol{q}, y - s^*\right) - x_k\left(\boldsymbol{q}, y - s^o\right) + \left[x_k\left(\boldsymbol{q}', y_L + Rs^*\right) - x_k\left(\boldsymbol{q}', y_L + Rs^o\right)\right] R^{-1}}{R \sum_i \pi^i \partial_I v\left(\boldsymbol{q}', y_i + Rs^*\right) \left[\partial_I x_k\left(\boldsymbol{q}, y - s^*\right) - \partial_I x_k\left(\boldsymbol{q}', y_i + Rs^*\right)\right]} \tag{23}$$

must be independent of k for the optimal solution to be characterized by uniform taxes. This is *not* the case, in general, i.e., (23) is *not* independent of k.

Were we in a traditional Mirrlees' setup and separability alone would suffice. Nonetheless, the independence condition required for uniform taxation to be optimal is *in addition* to separability. Preferences must be such that Marshallian demands satisfy constancy across goods of (23).

To understand this condition, we recall that $x_k(\boldsymbol{q}, y - s^o) - x_k(\boldsymbol{q}, y - s^*)$ is the difference in first period consumption of good k for an agent who chooses the relevant alternative strategy and the analogous choice for an agent who chooses the truthful strategy, while the term within brackets is the equivalent expression for second period consumption in the low productivity state. In the first period, productivities have not been revealed, hence, it is only through differences in savings that consumption choices are affected. Similarly, in the second period, for the consumption of good k due to choosing a different strategy. The agent always gets the after tax income of a low productivity agent, but her available income differs due to different savings.

Separability guarantees that the amount of leisure an agent gets does not affect demand for goods *conditional on a given level of expenditures*. In a static setting, or if the government has full control of agents' savings, all demands are conditioned on the expenditures determined by the allocation chosen by the agent. This means that the consumption patterns of a low productivity agent and of a high productivity agent who claims to be of low productivity are identical. In our model, however, savings are not directly controlled by the government and are added to after tax labor income to determine total available income, i.e., after tax labor income may *not* be taken as identical to expenditures. If savings differ for agents who intend to announce falsely and those who intend to abide by the rules—lemma 3, in the appendix, shows that this is indeed the case—, then available income will differ and the consumption pattern will signal one's deviating behavior, if goods have different income elasticity of demand.

4.1.2 The role of homotheticity

Since income effects are key in generating differences in consumption choices, the natural candidate for restoring the results are preferences for which income effects are in some sense neutral, e.g., homothetic preferences.

When preferences are homothetic the shares of expenditure, ω_k, are constant, and the left hand side of (21) becomes

$$\mu\partial_I v\left(\boldsymbol{q}, y-s^o\right)\left[\omega_k\left(s^*-s^o\right)+R^{-1}\omega_k R\left(s^o-s^*\right)\right]/p_k=0.$$

Finally note that under homotheticity, we have that the expression for $\left(\partial_k+\partial_{k'}\right)\hat{s}^*$ in (32) is proportional to

$$\partial_I v\left(\boldsymbol{q}, y-s^*\right)-R\sum\nolimits_i \pi^i\partial_I v\left(\boldsymbol{q}', y_i+Rs^*\right)=0.$$

Hence, the right hand of (22) also vanishes. Homotheticity restores the uniform tax prescription.

Increased savings means more second period (and less first period) available income. Therefore, consumption of at least one good must be increased, thus signaling the agent's lying. One may then wonder why isn't it then the case that the simple observation of more consumption in the second period does not justify the violation of AS. That is, why is homotheticity capable of restoring the result?

There are two things to consider. First, AS is about differential taxation, and under homothetic preferences any departure of uniform taxes will hurt agents along the equilibrium path just as badly as those off the equilibrium. Second, inter-temporal taxation—broadly interpreted as taxation of savings—handles this part of the effects of deviant behavior, as we shall show next.

4.2 The Taxation of Savings

When goods are uniformly taxed, the first order conditions with respect to y, y_H and y_L are, respectively,

$$\left(1+\mu\right)\partial_I v^*\left(\boldsymbol{q}, y-s^*\right)-\mu\partial_I v\left(\boldsymbol{q}, y-s^o\right)=\lambda\left(1-\left(1-R\beta\right)\partial_y s^*\right), \tag{24}$$

$$\left(1+\mu\right)\pi^H\partial_I v\left(\boldsymbol{q}, y_H+Rs^*\right)=\lambda\left(\beta\pi^H-\left(1-R\beta\right)\partial_{y_H}s^*\right), \tag{25}$$

and,

$$\left(1+\mu\right)\pi^L\partial_I v\left(\boldsymbol{q}, y_L+Rs^*\right)-\mu\partial_I v\left(\boldsymbol{q}, y_L+Rs^o\right)=\lambda\left(\beta\pi^L-\left(1-R\beta\right)\partial_{y_L}s^*\right). \tag{26}$$

Taking the difference between (24) and β^{-1} times the sum of (25) and (26) we get $\partial_I v\left(\boldsymbol{p}, y-s^*\right)-\beta^{-1}\sum_i \pi^i\partial_I v\left(\boldsymbol{q}, y_i+Rs^*\right)=$

$$\lambda\frac{1-R\beta}{1+\mu}\left[\partial_y s^*-\beta^{-1}\sum\nolimits_i \partial_{y_i}s^*\right]+\frac{\mu}{1+\mu}\left[\partial_I v\left(\boldsymbol{q}, y-s^o\right)-\beta^{-1}\partial_I v\left(\boldsymbol{q}', y_L+Rs^o\right)\right]$$

This must be true for any feasible R. Thus, letting $R \neq \beta^{-1}$, the first order conditions of the agents problem can be used to verify that the equality above implies that

$$\sum_i \pi^i \partial_I v\left(\boldsymbol{q}', y_i + Rs^*\right) - \frac{\mu}{1+\mu}\partial_I v\left(\boldsymbol{q}', y_L + Rs^o\right) = -\frac{\lambda}{(1+\mu)}\left[\beta\partial_y s^* - \sum_i \partial_{y_i} s^*\right]$$

From (31), in appendix A.2, we have, $\beta\partial_y s^* - \sum_i \partial_{y_i} s^* > 0$, which using the first order conditions for the savings choice both along and off the equilibrium path yields

$$\partial_I v^*\left(\boldsymbol{q}, y - s^*\right) - \frac{\mu}{1+\mu}\partial_I v\left(\boldsymbol{q}, y - s^o\right) < 0.$$

The fact that $s^o > s^*$ suffices to guarantee that $\partial_I v^*\left(\boldsymbol{q}, y - s^*\right) < \partial_I v\left(\boldsymbol{q}, y - s^o\right)$. What the expression above shows is that this difference is large enough to overcome the fact that the latter is multiplied by a number that is less than one.

We can finally use (24) to verify that $1 - (1 - R\beta)\,\partial_I s^* < 0$. Therefore, $1 - \beta R > (\partial_y s^*)^{-1} > 0$, i.e., $\beta^{-1} < R$. Savings are taxed at the optimum.

A remark is due here. While the violation of AS does not hinge upon which IC constraint binds at the optimum, the sign of the marginal tax on savings does. The logic is straightforward. The role of commodity and savings taxation in the presence of an optimally designed non-linear labor income tax schedule, is to relax the IC constraints. If agents who intend to announce falsely save more than those who intend to abide by the rules, reducing the gains from savings hurts this off-equilibrium behavior, thus playing a role in relaxing IC constraints. What guarantees that off-equilibrium savings is higher than equilibrium savings is the fact that the relevant deviating strategy is the one that prescribes always announcing to be of low type, for it generates a higher expected marginal utility of income.

Finally it is worth pointing out to the fact that the main difficulty in extending the savings taxation result for a more complex setting is exactly the fact that a full-characterization of deviating strategies and equilibrium allocations is needed. However, as long as the relevant deviation is downwards, i.e., announcing to be less productive than one really is, Golosov and Tsyvinsky (2005), show that the result remains valid.

5 Conclusion

We derive the properties of an optimal tax system comprised of a non-linear labor income tax and linear taxes on goods and on savings for an overlapping generations economy, where, following the new dynamic public finance literature, we assume that each agents' productivity changes through time and is private information, but, contrary to most of the literature, we assume that the government does not directly control agents's savings.

The problem is solved using a mechanism design approach, where understanding off-equilibrium savings is shown to be crucial for the identification of incentive compatible allocations. At the same time, it is the very same behavior regarding savings of agents who opt not to tell the truth which determines how taxation of goods and capital income may supplement the optimal income tax schedule. We show that the simple fact that savings differ off and along the equilibrium path generates a violation of Atkinson and Stiglitz' (1976) uniform taxation prescription: which is in contrast with the rest of the literature, where the result is shown to extend to a dynamic environment.

To understand our results it is useful to cite a recent paper by Cremer et al. (2001). They show that uniform taxation is usually not optimal in Mirrlees' setup when multi-dimensional heterogeneity due to differences in the (non-observable) endowments of some goods is present. Income effects become important and, as in our case, homotheticity must be added to separability for AS to obtain.

The similarity of results is not accidental: there is a subtle way in which unobserved heterogeneity shows up in our framework. From a second period perspective, agents off the equilibrium path have different 'unobserved endowments' when compared to agents who abide by the rules, since they save a different amount. The point here is that agents optimally anticipate their false announcements and save accordingly. Optimal savings for any off-equilibrium strategy will, in general, differ from the optimum savings under truthful announcement, making the consumption pattern differ for agents who tell the truth and agents who lie if goods have different income elasticities of demand.[9]

The fact that off-equilibrium savings are greater than savings along the equilibrium path drives the optimal prescription for savings: the tax rate on savings is positive! Even though the literature has already shown that an inverse Euler equation which implies a wedge between current and future marginal utility of income characterizes the optimum allocation, this need not translate into an expected marginal tax rate on capital income. In our case, however, because taxes are not state-dependent, the wedge

[9] The fact that the argument only depends on savings being different and not on how they differ indicates that the result should extend to more complex environments than the one considered here.

must be created by a positive tax on savings.

A Appendix

A.1 The Relevant IC constraints

Claim 1 *Constraint (8) is not binding at the optimum.*

Proof of Claim 1. Fix any level of savings s, then note that

$$u(y_{H,t}+s)-\zeta\left(\frac{Y_{H,t}}{w_H}\right) \geq u(y_{L,t}+s)-\zeta\left(\frac{Y_{L,t}}{w_H}\right), \text{ and} \tag{27}$$

$$u(y_{L,t}+s)-\zeta\left(\frac{Y_{L,t}}{w_L}\right) \geq u(y_{H,t}+s)-\zeta\left(\frac{Y_{H,t}}{w_L}\right) \tag{28}$$

imply that

$$\zeta\left(\frac{Y_{L,t}}{w_L}\right)-\zeta\left(\frac{Y_{H,t}}{w_L}\right) \leq \zeta\left(\frac{Y_{L,t}}{w_H}\right)-\zeta\left(\frac{Y_{H,t}}{w_H}\right) \tag{29}$$

Next, note that $Y_{H,t} < Y_{L,t}$, if and only if

$$\frac{Y_{L,t}-Y_{H,t}}{w_H} < \frac{Y_{L,t}-Y_{H,t}}{w_L}$$

which implies

$$\zeta\left(\frac{Y_{L,t}}{w_L}\right)-\zeta\left(\frac{Y_{H,t}}{w_L}\right) > \zeta\left(\frac{Y_{L,t}}{w_H}\right)-\zeta\left(\frac{Y_{H,t}}{w_H}\right)$$

from the convexity of $\zeta(\cdot)$, and the fact that $Y_{H,t}/w_L > Y_{H,t}/w_H$. This contradicts (29) and shows that any implementable allocation must be increasing. Because s was arbitrary in (27) and (28) the argument can be reversed to show that (8) cannot be binding, at any implementable allocation. ∎

Claim 2 *Absent commodity taxes, only IC constraint (6) binds at the optimum.*

Proof of Claim 2. Assume that the optimal allocation is such that $Y_{H,t} < y_{H,t}$ and $Y_{L,t} > y_{L,t}$. Then, for any s the allocation is a mean preserving spread over the allocation that the agent would obtain in autarchy by choosing s and consuming $Y_{H,t}$ and $Y_{L,t}$ conditional on having innate ability w_H and w_L, respectively. Because agents are risk averse, utility is lower in the first case. Hence, this cannot be optimal. Notice also that adding a constant at the left hand side of the inequality does not alter the argument, which means that transfers from the previous generations will not alter the argument.

Next consider the case where $Y_{H,t} \geq y_{H,t}$ and $Y_{L,t} \leq y_{L,t}$. If the IC constraint is binding, the expected utility delivered by the optimal tax scheme is (dropping price vectors as arguments of $u(\cdot)$ for notational simplicity):

$$u(y_t - s_t^*) + \sum_{i=H,L} \pi^i \left[u(y_{i,t} + R_t s_t^*) - \zeta\left(\frac{Y_{i,t}}{w_i}\right)\right] =$$

$$u(y_t - s_t^o) + u(y_{H,t} + R_t s_t^o) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y_{H,t}}{w_i}\right)$$

Consider now pooling agents at an allocation $(\hat{y}_t, Y_{H,t})$ where $\hat{y}_t$ is defined by adding $Y_{H,t} - y_{H,t} - \max\left\{Y_{H,t} - y_{H,t} - \left(\tilde{Y}_{L,t} - y_{L,t}\right); 0\right\}$ to $y_{H,t}$ to define a new disposable income $\hat{y}_t$. This is clearly feasible. It is trivial to see that the total utility the agent obtains at this new allocation is greater than what is obtained with the previous allocation. This allocation is also feasible for it generates all additional output it generates. Finally note that the allocation is trivially incentive compatible since the agents are pooled. ■

There are two important things to retain from this claim. First is the fact that, starting from a position where goods are not taxed then it is only constraint (6) that binds at the optimum. This guarantees that commodity taxes are introduced to relax IC constraint (6). Second, once commodity taxes are introduced some indirect effects defined over conditions on demands which are hard to interpret show up in the arguments and must be accounted for. These indirect effects, which are pervasive in the supplementary commodity taxation literature, are very unlike to change the basic result, even though no formal proofs of this being the case will be offered here.[10]

We end this discussion with the following claim.

Claim 3 *At any implementable allocation $s^o > s^*$.*

Proof of Claim 3. Just note that $y_H > y_L$ which means that,

$$\sum\nolimits_{i=H,L} \pi^i u'(y_i + s) < u'(y_L + s),$$

at any s. ■

[10] Nor do we know of this having been done in any other place. In fact, Lagrangian multipliers of incentive compatibility constrains are only signed in a Mirrleesian framework in the absence of supplementary commodity taxes—e.g., Ebert (1992) and Brunner (1993)—while classic papers in supplementary commodity taxation—e.g., Mirrlees (1976), Cooter (1978)—overlook the issue.

A.2 Price variations and Savings

Consider the following problem

$$\max_s \left\{ v(\boldsymbol{q}, y-s) + \sum\nolimits_{i=H,L} \pi^i v\left(\boldsymbol{q}', y_i + \hat{R}s\right)\right\} \tag{30}$$

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

$$-\partial_I v + R\left[\sum\nolimits_{i=H,L} \pi^i \partial_I v\,(i)\right] = 0,$$

and

$$\Theta \equiv -\partial^2_{II} v - \sum\nolimits_{i=H,L} \pi^i \partial^2_{II} v\,(i) > 0.$$

We may then use these to find

$$\begin{aligned}\Theta\left(\partial_k + \partial_{k'}\right)s &= -\partial^2_{Ik} v + R\sum\nolimits_{i=H,L} \pi^i \partial^2_{Ik} v\,(i) \\ &= \left(\partial^2_{II} v x_k + \partial_I v \partial_I x_k\right) - R\sum\nolimits_{i=H,L} \pi^i \left[\partial^2_{II} v\,(i)\, x_k\,(i) + \partial_y v\,(i)\,\partial_I x_k\,(i)\right]\end{aligned}$$

where Roy's identity was used for the latter part of the expression.

The same procedure for $\partial_y s$, $\partial_{y^H} s$ and $\partial_{y^L} s$, yields

$$\Theta\partial_y s = -\partial^2_{II} v,\ \Theta\partial_{y^H} s = \hat{R}\pi^H \partial^2_{II} v\,(H)\,,\ \text{and}\ \Theta\partial_{y^L} s = \hat{R}\pi^L \partial^2_{II} v\,(L)\,, \tag{31}$$

respectively.

Hence,

$$\begin{aligned}\Theta\left(\partial_k + \partial_{k'}\right)\hat{s} &= \left(\partial_k + \partial_{k'}\right)s + x_k \partial_y s + \sum\nolimits_{i=H,L} x_k\,(i)\,\partial_{y_i} s \\ &= \partial_I v \partial_I x_k - R\sum\nolimits_{i=H,L} \pi^i \partial_I v\,(i)\,\partial_I x_k\,(i) \\ &= \sum\nolimits_{i=H,L} \pi^i \partial_I v\,(i)\left[\partial_I x_k\,(i) - \partial_I x_k\right]\end{aligned} \tag{32}$$

Moreover,

$$\partial_y s^* - \beta^{-1}\sum\nolimits_i \partial_{y_i} s^* = \frac{\partial^2_{II} v\left(\boldsymbol{q}, y-s\right) + R\beta^{-1}\sum_i \pi^i \partial^2_{II} v\left(\boldsymbol{q}', y_i + \hat{R}s\right)}{\partial^2_{II} v\left(\boldsymbol{q}, y-s\right) + R^2\sum_{i=H,L} \pi^i \partial^2_{II} v\left(\boldsymbol{q}', y_i + \hat{R}s\right)} > 0.$$

References

[1] Albanesi, S. and C. Sleet (2005) "Dynamic Optimal Taxation with Private Information," *The Review of Economic Studies,* 72: 1-29.

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

[3] Brunner, J. (1993) "A Note on the Optimum Income Tax" *Journal of Public Economics* 50, 445-451.

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

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

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

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

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

[9] da Costa, C. and I. Werning, (2000) "A Note on Pre-Committed Goods." mimeo., University of Chicago.

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

[11] Ebert, U. (1992) "A Reexamination of the Optimal Non-linear Income Tax" *Journal of Public Economics*, 49: 47-73

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

[13] Golosov, M., and A. Tsyvinski (2005) "Optimal Taxation with Endogenous Insurance Markets" NBER, working paper no. 11185.

[14] Kocherlakota, N. (2004) "Wedges and Taxes" *American Economic Review Papers and Proceedings*, 94 (2): 109-113.

[15] Kocherlakota, N. (2004) "Figuring out the impact of hidden savings on optimal unemployment insurance" *Review of Economic Dynamics* 7 541–554

[16] Kocherlakota, N. (2005) "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation" *Econometrica*, 73: 1587-1621.

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

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

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

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