



F U N D A Ç Ã O
GETULIO VARGAS

EPGE

Escola de Pós-Graduação
em Economia

Ensaio Econômico

Escola de

Pós Graduação

em Economia

da Fundação

Getúlio Vargas

Nº 596

ISSN 0104-8910

Yet Another Reason to Tax Goods

Carlos E. da Costa

Julho de 2005

Os artigos publicados são de inteira responsabilidade de seus autores. As opiniões neles emitidas não exprimem, necessariamente, o ponto de vista da Fundação Getulio Vargas.

Yet Another Reason to Tax Goods*

Carlos E. da Costa[†]
Getulio Vargas Foundation

July 6, 2005

Abstract

The optimal taxation of goods, labor and capital income is considered in a two period model where: *i*) private information changes through time; *ii*) savings are not observed, and; *iii*) savings affect preferences conditional on the realization of types. The simultaneous appearance of these three elements cause optimal commodity taxes to depend on off-equilibrium savings. As a consequence, separability no longer suffices for the uniform taxation prescription of Atkinson and Stiglitz (AS) to obtain. If preferences are homothetic AS is partially restored: taxes are uniform within periods, however, future consumption is taxed at a higher rate than current consumption.

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

JEL Classification: H21, D82.

1 Introduction

The main result concerning the supplementary 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 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.

To understand the rationale of the result we recall the derivation of an optimal non-linear income tax schedule. Without loss—as assured by 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

*I thank Luis Braido, Bev Dahlby, Lucas Maestri and seminar participants at PUC-RJ, IBMEC, EPGE, the 2004 Latin American and European Meetings of the Econometric Society and the 2004 Public Economics and Economic Theory meeting held by the APET in Rio de Janeiro 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

they must supply and net income to which they are entitled.¹ Truthful announcement of productivities is guaranteed if incentive compatibility constraints are satisfied.

Commodity taxation is useful in this world if it relaxes 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.

This result has been under some attack due to its dependence on the specific structure of Mirrlees' (1971, 1976) setup. Two assumptions have been shown to be crucial for the result to be valid: perfect substitutability of different skills and unidimensional heterogeneity. The first assumption is relaxed by Naito (1999) in a model where labor supply of skilled workers is not a perfect substitute of labor supply of unskilled workers. Taxation of goods affect incentive compatibility constraints through changes in relative prices of skills: an important implication of the Stopler-Samuelson theorem for the design of optimal taxes.

As for multi-dimensional heterogeneity, Saez (2001), Cremer et al. (2001), etc., have shown how consumption patterns of mimickers and the agents they mimic may differ, even under separability, if other dimensions of heterogeneity are present. Taxation of goods may still play a role, in this case.

In the opposite direction to these challenges to AS, recent advances on dynamic optimal taxation have shown that the uniform tax prescription is still valid in a world where skills evolve stochastically over time and tax instruments may depend on complex informations—e.g. Cremer and Gahvari (1995,1999), da Costa and Werning (2000), Golosov et al. (2003).

The present paper takes on the same topic from the dynamic optimal taxation perspective. We investigate AS in a two periods version of Stiglitz's (1982) representation of the optimal income taxation problem. We follow the new dynamic public finance literature in including an evolving information set for the agents as the crucial dynamic element for the model. We diverge, however, from most of the same literature in assuming that private savings are not directly controlled by the government.

This simple and compelling restriction on policy instruments is sufficient to overturn AS, which is in contrast with the result found in previous dynamic taxation models. To make our results comparable with the same literature we take AS to mean the uniform taxation of goods within each period and show that the result is restored if homotheticity is added to separability.

Interestingly enough AS breaks down despite the fact that we maintain unidimensional heterogeneity and perfect substitutability between skills: the only two hypothesis that are known to break AS when assumed away.

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-

¹The fact that these allocations are associated to a tax function is a direct consequence of the taxation principle.

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’ than agents along the equilibrium path, since savings differ.

As we have argued, supplementary commodity taxation is useful if it allows for relaxing incentive compatibility constraints. Separability rules out differences in choices *conditional on available income*, which is the relevant difference between mimickers and the agents they mimic in Mirrlees’ setup. Here, however, agents who anticipate deviating behavior differ from agents who intend to abide by the rules in terms of their second period available income, for they change their savings pattern in the first period. Deviating behavior thus implies different consumption choices generated by off-equilibrium savings.

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. It turns out that knowing how they differ allows us to tell whether optimal taxes on current consumption goods are higher or lower than those in future consumption goods. Taxing goods in different periods at different rates is equivalent to imposing an anonymous tax on capital income, therefore, being able to sign these taxes is of paramount relevance.

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 more heavily second period consumption—i.e., taxing capital income.²

In a world where savings are directly controlled by the government, an inverse Euler equation describes the optimal inter-temporal allocation.³ Because the marginal utility of today’s income is smaller than expected marginal utility of tomorrow’s income, government must induce agents to consume early more than they would privately choose, either by obliging agents to do so or by imposing state dependent taxes on capital in-

²In a parallel and independent work, Golosov and Tsyvinsky (2003) also consider a multi-period version of this problem and investigate optimal capital income taxation in a decentralized economy. This extension to a multi-period environment comes at a cost. To handle the problem they assume that shocks to productivity are i.i.d., which makes the nature of uncertainty quite distinct from what we have here. Upon making explicit assumptions about which constraints bind at the optimum, they show that it is optimal to tax capital income. They do not discuss the differential taxation of goods.

³The inverse Euler equation result was first derived in a Mirrleesian framework, to the best of our knowledge, in da Costa and Werning (2000) though, it could be derived quite straightforwardly from the findings of Cremer and Gahvari (1999) where it was associated to the subsidization of pre-committed goods. Golosov et al (2003) generalized in many dimensions this result which is reminiscent of Rogerson’s (1985) findings in a dynamic moral hazard setting.

come.⁴ Here, the only way in which government may affect the timing of consumption is by taxing goods at different rates according to the period in which they are consumed. This is, of course, equivalent to having a positive anonymous withholding tax on capital income.

The remainder of this paper is organized as follows. The economy is presented in Section 2. Then, in Section A.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 shows the policy implications of the model for retirement savings, while section 5 concludes. All results are proved in the appendices.

2 The Environment

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

$$u(\mathbf{x}^0) - \zeta(l^0) + \sum \pi^i [u(\mathbf{x}^i) - \zeta(l^i)],$$

where π^i is the probability associated with state i . Consumption vectors for first period and state i of second period, are represented by $\mathbf{x}^0, \mathbf{x}^i \in \mathbb{R}^n$, respectively,⁵ while $l^0, l^i \in \mathbb{R}$ 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 to investigate Atkinson and Stiglitz's uniform taxation result. We also assume that ζ is strictly increasing and convex while u is strictly increasing and concave and that both functions are smooth.

Uncertainty arises in this problem because in the first period agents do not know their 'adult' productivities, w , which we call their types. In the first period, w is identical for all agents and normalized to 1, while in the second we follow Stiglitz (1982) in considering only two possible types, or states: H (for high productivity) and L (for low) with $w^H > w^L$. We also assume that shocks are independent to use the law of large numbers convention that equates the cross-sectional distribution of types with the probability distribution faced by each agent in the first period.

Asymmetric information in this model is due to the fact that, once uncertainty is realized, each agent's productivity is only observed by the agent herself.

Despite temporary utility being the same differences in productivities affect the way agents of different productivities rank consumption and supply of efficiency units, Y . To understand the issue, note that an agent of productivity w needs $l = Y/w$ hours to supply Y labor efficiency units: the higher her productivity, the more leisure she gets for the same Y she supplies. As a consequence, preferences over bundles of \mathbf{x} and Y differ,

⁴E.g., Kocherlakota (2004).

⁵Bold is used to represent vectors. We adopt the convention that prices are row and quantities, column vectors.

according to an agent's productivity. This is not a minor issue since trade takes place for these objects, and not directly l .

Technology is very simple. The only input used in production is efficiency units of labor. These are sold by the agents to a representative firm and transformed in consumption goods \mathbf{x} through a linear technology. Units are normalized so that marginal cost of all goods equal to 1. Finally, goods are sold back to agents with competition driving (producer) prices to marginal costs.

The key difference between trade in goods and trade in efficiency units is that we assume that agents can conduct side trade of goods at no transaction cost, while for efficiency units the transaction costs are assumed to be prohibitively high. Finally, savings are assumed not to be observed.

A benevolent government who inhabits this economy maximizes the agents' expected utilities. 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 rule out any form of non-linearity in their taxation. Similarly, savings cannot be directly observed, thus cannot be taxed. However, a uniform higher tax for second period goods can mimic a withholding tax based on anonymous transactions on capital income. This type of instrument is also considered in Cremer et al. (2001), where, however, savings do not affect the ranking of bundles in the second period.

To find the optimal tax schedule we define a truthful direct mechanism and derive the allocation that maximizes the government objective function. 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.

2.0.1 The Nature of the Game

The game played by the government and the agents is a Stackelberg game, where the Government, the Stackelberg leader, moves first by choosing: a first period allocation, (y, Y) , a second period budget set, $\mathbb{B} \equiv \{(y, Y); (y, Y) = (y^i, Y^i) \text{ for } i = H \text{ or } i = L\}$, and tax rates $\boldsymbol{\tau} \equiv (\mathbf{p} - \boldsymbol{\iota})$ and $\boldsymbol{\theta} \equiv (\mathbf{q} - \boldsymbol{\iota})$, where \mathbf{p} and \mathbf{q} are first and second period consumer prices and $\boldsymbol{\iota}$ is a vector of ones. It should also be clear that $T(Y) \equiv Y - y$.

In the first period, an agent must supply Y efficiency units and is left with after-tax income y , which implies non-linear taxation of labor income in the first period, as well. She then chooses what to consume and how much to save. All this before nature defines the agent's type. Once her individual productivity is realized, the agent chooses her bundle of net and gross income— y and Y , respectively—among those contained in the budget set, \mathbb{B} , made available by the government. Finally, she adds her savings to the after tax income, y , that corresponds to the bundle she chose, and uses her available income to buy her preferred basket of goods.

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

to her type, from the budget set, \mathbb{B} .

As usual, one may guarantee that an allocation is implementable if it satisfies the associated incentive compatibility, henceforth IC, constraints. Only considering IC constraints at the equilibrium level of savings will not suffice, however. Off-equilibrium savings must be taken into account 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. This is the so called *double-deviation* problem.⁶

The problem here is that savings affects the way an agent ranks bundles of (y, Y) , which means that whether a bundle is preferred to another depends on these first period choices. Conversely, the expected marginal utility of income, which is what drives optimal savings, depends on which bundles are to be picked in each state of the world.

This circularity creates a potential problem for solving the model. 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 the 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 is to always announce truthfully one's productivity. This is the strategy we want to induce the agents to adopt. What we show in the next section is that all we need to guarantee is that this strategy is no worse than always announcing to have low productivity. When this is done, the other two strategies are dominated by the strategy of truthful announcement at the optimum.

To understand the main argument, define a strategy as a mapping from types to announcements $\sigma^k : \{H, L\} \rightarrow \{H, L\}$, 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: $\sigma^*(H) = H$, $\sigma^*(L) = L$; $\sigma^o(H) = L$, $\sigma^o(L) = L$; $\sigma^{oo}(H) = H$, $\sigma^{oo}(L) = H$ and $\sigma^{ooo}(H) = L$, $\sigma^{ooo}(L) = H$. Strategy σ^* , prescribing truthful announcement for all realizations, is, as we said, the strategy we want to induce.

To further advance towards the solving the problem, define

$$v(\mathbf{p}, I) \equiv \max_{\mathbf{x}} u(\mathbf{x}) \text{ s.t. } \mathbf{p}\mathbf{x} \leq I, \quad (1)$$

with $\mathbf{x}(\mathbf{p}, I)$ as the corresponding (conditional) Marshallian demand. An analogous definition applies for second period choices.

⁶For a discussion of this issue in a moral hazard setting see Chiappori et al. (1995). In a self-selection framework very nice discussions are found in Golosov and Tsyvinsky (2003) and in Kocherlakota (2004).

Because each strategy defines a strictly concave savings problem, we associate to each σ^k a unique optimal level of savings s^k . That is, for $k = *,^o,^{oo}$ and ooo , we define s^k ,

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

where $y^{\sigma^k(i)}$ is net income received by an agent of type i using strategy k . Notice that, under our separability assumption, labor supply plays no role whatsoever in determining the optimal level of savings.

When looking for implementable allocations, we need only to compare the expected utility for pairs of announcement and savings— $\{(\sigma^k, s^k)\}_{k=*,^o,^{oo},^{ooo}}$ —with the expected utility for the pair (σ^*, s^*) we want to induce: if σ^* (along with s^*) is the expected utility maximizing strategy, we are assured that there is no savings choice that, coupled with an alternative strategy, yields higher expected utility than truthful announcement.⁷

In principle, to set up the program to be solved by the government we should include five IC constraints, namely: *i*) the two second period IC constraints that guarantee that, after uncertainty is resolved, 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, what we prove in appendix A.1 is that we need only to considered one incentive compatibility constraint: the one that ensures that always announcing truthfully—pair (σ^*, s^*) —yields an expected utility at least as high as always announcing to be of low productivity—pair (σ^o, s^o) . If this constraint is satisfied so are the other four at the optimum.

With regards to the second period IC constraints, if the first period ones are satisfied than the second period ones are usually satisfied with strict inequalities. As a consequence tax schedules will be interim inefficient, in the sense that 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:⁸ the ‘tax system’—for that matter, any deterministic implementable contract—is not renegotiation-proof in the sense of Dewatripont (1988). These ideas are apparent from the proofs found in A.1.

As for the first period IC constraints, intuitively, given the utilitarianism implicit in the maximization of expected utility, redistribution takes place from the high type to the low type. Therefore, it is the possibility of the high type announcing to be a low type that should be of concern here. When savings are controlled by the government, the fact that only downward constraint are binding is an immediate consequence of single-crossing. Here, however, the proof is a little more evolved and since the understanding of the main results of the paper do not require a step by step understanding of the proofs regarding which IC constraints bind we send them to the appendix.

⁷In a very general inter-temporal setting, Fernandez and Phelan (2000) show how to handle this type of inter-temporal links created by savings. By using value functions as state variables they make the problem recursive by generalizing the methodology of Abreu et al. (1990). Unfortunately, as pointed out by Kocherlakota (2004) it is still not known how to computationally implement this type of procedure in practice for very general settings. Hence the convenience of working with a finite horizon and finite states.

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

3 Optimal Taxation

The government provides life-time insurance for agents subject to the resource constraints of the economy and to an incentive compatibility constraint. This latter constraint guarantees that it is optimal for the agent to announce her productivity truthfully regardless of how productive she turns out to be.

As argued before, and proven in appendix A.1, the only relevant incentive compatibility constraint is the one which imposes that announcing truthfully is a strategy that yields as much expected utility as always announcing to be of low productivity (and choosing savings, s , accordingly).

Therefore we can write the government's program as

$$\max_{\mathbf{p}, \mathbf{q}, (y, Y), \{(y^i, Y^i)\}} v(\mathbf{p}, y - s^*) - \zeta(Y) + \sum_{i=H,L} \pi^i \left[v(\mathbf{q}, y^i + s^*) - \zeta\left(\frac{Y^i}{w^i}\right) \right],$$

subject to the relevant IC constraint

$$v(\mathbf{p}, y - s^*) + \sum_{i=H,L} \pi^i \left[v(\mathbf{q}, y^i + s^*) - \zeta\left(\frac{Y^i}{w^i}\right) \right] \geq v(\mathbf{p}, y - s^o) + v(\mathbf{q}, y^L + s^o) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y^L}{w^i}\right), \quad [\mu]$$

and the resource constraint

$$\iota \mathbf{x}(\mathbf{p}, y - s^*) + \iota \sum_{i=H,L} \pi^i \mathbf{x}(\mathbf{q}, y^i + s^*) \leq Y + \sum_{i=H,L} \pi^i Y^i, \quad [\lambda]$$

where we write the Lagrange multipliers inside brackets to the right of each constraint. Here, s^o is the optimum level of savings if the agent anticipates the strategy of always announcing to be a low productivity type, while s^* is the optimum level associated to the strategy of always telling the truth.

When compared to a standard optimal taxation problem some interesting new features are apparent. First, there is an extra term in the objective function which is the first period temporary utility. In practice, the social preferences are not subject to much controversy since agents are ex-ante identical and all that the government does is to maximize everyone's expected utilities.

Second, the s term that appears not only in the objective function but also in the IC constraints is an endogenous variable that must be accounted for when solving the optimization problem. 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 the first period labor supply does not appear in the IC constraint (it would appear in both sides of $[\mu]$). It does not mean that labor supply is not distorted in the first period. In fact, it is easy to verify that, at the

optimum, $v_y(\mathbf{p}, y) \neq \zeta'(Y)$.⁹ This result is new in the literature, and is in contrast with the prescription for the case where government controls savings.

To derive optimal commodity taxes we follow the standard procedure for the investigation of optimal supplementary commodity taxation first presented in Mirrlees (1976). We differentiate the Lagrangian with respect to prices and derive the necessary conditions for an interior optimum.

First Period Goods The first step is to, with the aid of the envelope theorem, differentiate the Lagrangian with respect to first period price p_j , to derive the first order condition with respect to price p_j ,

$$-v_y^* x^{j*} (1 + \mu) + \mu v_y^o x^{j*} - \lambda \mathbf{l}(\mathbf{x}_j - \mathbf{x}_y s_j) - \lambda \sum_i \pi^i \mathbf{l}_{\mathbf{x}_y}(i) s_j = 0, \quad (3)$$

where we used Roy's identity, to substitute $-v_y^* x^{j*}$ and $-v_y^o x^{j*}$ for v_j^* and v_j^o , respectively.

Next, we take the first order condition with respect to y

$$v_y^* (1 + \mu) - \mu v_y^o - \lambda \mathbf{l}(\mathbf{x}_y - \mathbf{x}_y s_y) - \lambda \sum_i \pi^i \mathbf{l}_{\mathbf{x}_y}(i) s_y = 0$$

and multiply by x^{j*} to get

$$v_y^* x^{j*} (1 + \mu) - \mu v_y^o x^{j*} - \lambda \mathbf{l}(\mathbf{x}_y - \mathbf{x}_y s_y) x^{j*} - \lambda \sum_i \pi^i \mathbf{l}_{\mathbf{x}_y}(i) s_y x^{j*} = 0. \quad (4)$$

Finally, add (3) with (4) to get

$$\mu v_y^o (x^{j*} - x^{j*}) - \lambda [\mathbf{l}(\mathbf{x}_j - \mathbf{x}_y x^{j*} + \mathbf{x}_y (s_j + s_y x^{j*}))] = \lambda \sum_i \pi^i \mathbf{l}_{\mathbf{x}_y}(i) (s_j + s_y x^{j*})$$

Define $\hat{s}_j \equiv s_j + s_y x^{j*}$ and $\mathbf{h}_j^* \equiv \mathbf{x}_j + \mathbf{x}_y x^{j*}$, to obtain the optimal tax prescription for first period goods,

$$\mu v_y^o (x^{j*} - x^{j*}) = -\lambda \mathbf{l} \left\{ \mathbf{h}_j^* + \left(\sum_i \pi^i \mathbf{x}_y(i) - \mathbf{x}_y \right) \hat{s}_j \right\}, \quad (5)$$

It is easy to verify that \mathbf{h}_j^* is the vector of derivatives of hicksian demands with respect to price p_j . As for \hat{s}_j , it is a form of 'compensated savings', whose sign only depend on whether good j is normal or inferior.

Second Period Goods Differentiating now the Lagrangian with respect to second period price q^j and using Roy's identity state by state one gets

$$\sum_i \pi^i v_y^*(i) x^{*j}(i) - \frac{\mu v_y^o(L)}{1 + \mu} x^{oj}(L) = \frac{-\lambda}{1 + \mu} \mathbf{l} \left[\mathbf{x}_y s_j - \sum_i \pi^i (\mathbf{x}_j(i) + \mathbf{x}_y(i) s_j) \right] \quad (6)$$

⁹The result is best seen in the case where there are no commodity taxes. It is then easy to show that

$$v_y^* + \mu [v_y^* - v_y^o] = -\zeta'(Y)$$

where v_y^* and v_y^o are, respectively, the marginal utility of income in first period for those who adopt the truthfull strategy and for those who always announce low.

Analogously to what was done in the previous section, multiply the first order condition with respect to y^H , by $x^{*j}(H)$ to obtain,

$$\pi^H v_y^*(H) x^{*j}(H) = \frac{\lambda x^{*j}(H)}{1 + \mu} \boldsymbol{\iota} \left[\pi^H \mathbf{x}_y(H) + \left(\sum_i \pi^i \mathbf{x}_y(i) - \mathbf{x}_y \right) s_{y^H} \right] \quad (7)$$

A similar procedure for y^L yields

$$\left[\pi^L v_y^*(L) - \frac{\mu v_y^o(L)}{1 + \mu} \right] x^{*j}(L) = \frac{\lambda x^{*j}(L)}{1 + \mu} \boldsymbol{\iota} \left[\pi^L \mathbf{x}_y(L) + \left(\sum_i \pi^i \mathbf{x}_y(i) - \mathbf{x}_y \right) s_{y^L} \right] \quad (8)$$

Now we add (6) and (7) subtract from (8) to get

$$\mu v_y^o(L) [x^{*j}(L) - x^{oj}(L)] = -\lambda \boldsymbol{\iota} \left\{ -\mathbf{x}_y \tilde{s}_j + \sum_i (\mathbf{h}_j(i) + \mathbf{x}_y(i) \tilde{s}_j) \pi^i \right\} \quad (9)$$

where $\tilde{s}_j = s_j + \sum_{i=H,L} \pi^i s_{y^i} x^{*j}(i)$ and $\mathbf{h}_j(i) = \mathbf{x}_j(i) - \mathbf{x}_y(i) x^j(i)$. These are the tax prescriptions for second period goods. Let us consider expressions (5) and (??).

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. The discouragement of consumption of good i has two components. The first is captured by the Hicksian demand term \mathbf{h}_j —which, due to symmetry, is equal to the gradient of h^j —both for first and second period consumption. The second term captures the indirect discouragement due to changes in savings.

The left hand side of both equations represent the change in consumption of good j , due the adoption of a strategy different from truth-telling. In principle, one could be intrigued by the fact that the formula does not seem to include the choice of a high type pretending to be a low type, as we usually see in this literature. But it does! The point here is that separability makes the consumption choices of high and low types, conditional on a given level of disposable income, identical. As a consequence, the relevant comparison is not between high types who lie and low types, but between agents who have anticipated to always announce a low type (be they true low types or mimickers) and those who have chosen to abide by the rules.

Notice also that the prescription is to discourage mostly goods that are more affected by differences in income, i.e. those with higher income elasticity of demand.

3.0.2 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.

Along these lines, assume that $\mathbf{p} = \boldsymbol{\iota}$ and $\mathbf{q} = \rho \boldsymbol{\iota}$, where ρ is an arbitrary scalar an interpretation of which we shall postpone for a while. Symmetry and homogeneity of Hicksian demands guarantee that, at $\mathbf{p} = \boldsymbol{\iota}$, $\mathbf{q} = \rho \boldsymbol{\iota}$, $\sum_k h_k^{*j} = \rho \sum_k \pi^i h_k^{*j}(i) = 0$, $i = H, L$,

while $\sum_k x_y^{*k} = \rho \sum_k x_y^{*k} (H) = \rho \sum_k x_y^{*k} (L) = 1$ as a consequence of Engel's aggregation - which holds for each period and each state of nature, conditional on chosen level of income.

Expression (5) thus collapses to

$$\mu v_y^o (x^{jo} - x^{j*}) = -\lambda (1 - 1/\rho) \hat{s}_j, \quad (10)$$

while (9) becomes

$$\mu v_y^o (L) [x^{*j} (L) - x^{oj} (L)] = -\lambda (1 - 1/\rho) \tilde{s}_j. \quad (11)$$

We shall say that AS holds if there is a number $\rho > 0$ for which equations (10) and (11) are satisfied for all j , or equivalently, we need the terms in the left hand side of

$$\frac{x^{jo} - x^{j*}}{\hat{s}_j} = \frac{-\lambda (1 - 1/\rho)}{\mu v_y^o} \text{ and } \frac{x^{*j} (L) - x^{oj} (L)}{\tilde{s}_j} = \frac{-\lambda (1 - 1/\rho)}{\mu v_y^o (L)}.$$

to be independent of j , for AS to hold.

We state this more formally in proposition 1, where an expression for \hat{s}_j^* in terms of a more familiar object is also provided—the expression is formally derived in appendix A.2. 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 1 *There is a ρ that satisfies equations (10) and (11) only if*

$$\frac{(x^{oj} - x^{*j})}{v_y^* x_y^{*j}} \text{ and } -\frac{(x^{oj} (L) - x^{*j} (L))}{\sum_{i=H,L} \pi^i v_y^* (i) x_y^{*j} (i)} \quad (12)$$

are constant across goods.

Were we in a traditional Mirrlees' setup and separability alone would suffice. Nonetheless, the condition required for uniform taxation to be optimal in Proposition 1 is *in addition* to separability. Preferences must be such that Marshallian demands satisfy constancy across goods of (12).

To understand what this conditions implies, we recall that $x^{oj} - x^{*j}$ is the difference in first period consumption of good j 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, $x^{oj} (L) - x^{*j} (L)$ is the change in the second period consumption of good j due to choosing a different strategy. In this case the agent always gets the after tax income of a low productivity agent, but has different available income due to differences in savings.

Separability guarantees that the amount of leisure an agent gets does not affect demand for goods *conditional on a given level of expenditures*. On the one hand, this means that consumption pattern of a low productivity agent and a high productivity agent who claims to be of low productivity are identical. On the other, contrary to a situation where after tax labor income may be taken as identical to available income, in our model, savings are not directly controlled by the government and are added to after tax labor income to determine total expenditures.

Because differences in savings associated to each different strategies determine whether expression (12) is expected to be constant across goods, and ultimately determine the validity of AS, it becomes crucial to determine how savings differ for the two relevant strategies. Fortunately, claim 4, in the appendix, states that savings are greater for the strategy of always announcing low than for the truth-telling strategy, i.e., $s^o > s^*$.

As previously mentioned, separability and convexity of the utility function guarantee that leisure is normal. Agents who anticipate that they will always announce to be of a low type will, then, choose to save more than agents who opt for a truthful strategy.

Intuitively, when an agent decides that she will announce a low type no matter what, she increases her savings according to claim 4. 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.

Note that 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 inter-temporal taxation—broadly interpreted as capital income taxation—handles this part of the effects of deviant behavior. We shall come back to this issue, later.

To summarize, if goods have different income elasticity of demand, independence condition (12) 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 1 *If preferences are separable and homothetic, then uniform taxation of goods is optimal.*

Proof of Corollary 1. If preferences are homothetic, $p^j x^{oj} = \omega^j (y - s^o)$, $p^j x^{*j} = \omega^j (y - s^*)$, $q^j x^{oj}(L) = \rho p^j x^{oj}(L) = \omega^j (y^L + s^o)$, $q^j x^{*j}(L) = \rho p^j x^{*j}(L) = \omega^j (y^L + s^*)$, $p^j x_y^{*j} = \omega^j$ and $q^j x_y^j(i) = \rho p^j x_y^j(i) = \omega^j$, $i = H, L$, where ω^j is the (constant) proportion of income spent on good j .

In this case,

$$\frac{(x^{oj} - x^{*j})}{v_y^* x_y^{*j}} = \frac{\omega^j (y - s^o) - \omega^j (y - s^*)}{v_y^* \omega^j} = \frac{s^* - s^o}{v_y^*}, \text{ and} \quad (13)$$

$$-\frac{(x^{oj}(L) - x^{*j}(L))}{\sum_i \pi^i v_y^*(i) x_y^{*j}(i)} = -\frac{\omega^j (y^L + s^o) - \omega^j (y^L + s^*)}{\omega^j \sum_i \pi^i v_y^*(i)} = \frac{s^* - s^o}{\sum_i \pi^i v_y^*(i)}. \quad (14)$$

Both are independent of j , which, according to proposition 1 is sufficient for uniform taxation to be optimal. ■

At this point it is interesting to compare our results to the ones in Cremer et al. (2001).¹⁰ 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: 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.

3.0.3 Inter-temporal Taxation

When defining AS in our setup we did not require prices in both periods to be identical. If this were the case, ρ would be 1 and the terms in the left hand side would vanish. Such a solution would only be possible if savings were not changed by the adoption of the alternative strategy. As it should be clear, at this point, ρ is the price of future consumption in terms of today’s goods. This statement being precise when goods are not taxed at different rates within a period.

Hence, assuming homotheticity, which we have shown to guarantee uniform taxation in both periods, we shall now prove that second period prices are uniformly higher than first period prices. When savings are positive, in equilibrium, this is equivalent to an

¹⁰See their example in page 790, second paragraph.

anonymous tax on capital income taxation. When they are not, optimal policy still implies the subsidization of credit for early consumption.

The first step of the proof, found in the appendix, is to show that the sign of $1 - 1/\rho$ is the same as the sign of $s^* - s^o$. That is, the optimal inter-temporal prices are driven by the way savings change according to the strategy adopted. This naturally depends on σ^o being the relevant alternative strategy. Therefore, from a policy perspective, it is important to identify the binding IC constraint since the sign of the (implicit) optimal tax rate on capital income will be different if it is constraint (16) rather than (15) that binds at the optimum. Because we are able to show that—see claim 3, if goods are not differentially taxed within periods, only constraint (15) binds at the optimum, under the same conditions that deliver AS, namely homotheticity, we can prove our results for (implicit) taxation on capital income.

The following proposition is an immediate consequence of corollary 1 and claim 3.

Proposition 2 *If preferences are separable and homothetic: i) goods are uniformly taxed within each period and; ii) tax rate on second period goods are higher than on first period goods, i.e., $\rho > 1$.*

Proof of Proposition 2. If preferences are homothetic, $\mathbf{p} = \boldsymbol{\iota}$ and $\mathbf{q} = \rho\boldsymbol{\iota}$,

$$\frac{(x^{oj} - x^{*j})}{v_y^*} = \frac{-\lambda(1 - 1/\rho)}{\mu v_y^o} x_y^{*j}$$

Adding over j yields, in this case,

$$\frac{s^* - s^o}{v_y^*} = \frac{-\lambda(1 - 1/\rho)}{\mu v_y^o} \sum_j x_y^{*j} = \frac{\lambda(1/\rho - 1)}{\mu v_y^o}$$

Because $s^* < s^o$, according to claim 4, then it must be the case that $\rho > 1$.

The same result obtains if we instead add the following expression over j

$$x^{oj}(L) - x^{*j}(L) = \frac{-\lambda(1 - 1/\rho)}{\mu v_y^o(L)} \sum_{i=H,L} \pi^i v_y^*(i) x_y^{*j}(i).$$

This 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, goods are taxed at a higher rate in the second period. This, which is a form of linear tax on capital income, is related to other findings in multi-period Mirrlees' settings. ■

The sign of the marginal tax on capital income, i.e., the difference in taxes for consumption goods in the two periods, depends on which specific IC constraint binds at the optimum.¹¹ The logic is straightforward. The role of commodity and capital income

¹¹ Golosov and Tsyvinsky (2003), upon assuming the direction of binding IC constraints in a multi-period setup akin to the one found herein, also find that it is optimal to tax capital income.

taxation in the presence of an optimally designed non-linear labor income tax schedule, is to relax the IC constraints.

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.

4 Retirement

In this section we extend the model presented in section 3 to include a brief discussion about the taxation of retirement funds.

It is needless to say that the topic is of paramount importance from a policy perspective. Notwithstanding the fact that our model is very stylized, the discussion becomes particularly relevance once we recall that our main result regarding inter-temporal taxation goes in a different direction from the literature on capital income taxation. One may, then wonder whether we are also prescribing a policy for taxation of retirement funds, which differs from what is advocated by Atkinson and Stiglitz (1976) as a corollary of their uniform tax result.

The answer is yes and no. Yes, all investments, including investments made early in life for retirement purposes, are taxed in our model, which is a novel prescription. And no, if we take the prescription of Atkinson and Stiglitz in a more strict sense, we will show that we also advocate the exemption of taxes for retirement funds. Basically, once an agent has fully realized her earnings potential, taxation of capital income has no role in the optimal tax policy, exactly like in Atkinson and Stiglitz.

The addition of a retirement period is a straightforward exercise. First, assume that preferences are defined as

$$u(\mathbf{x}) - \zeta(l) + \sum_{i=H,L} \pi^i [u(\mathbf{x}_1^i) + u(\mathbf{x}_2^i) - \zeta(l^i)].$$

where, \mathbf{x}_2^i is the vector of consumption goods at retirement for an agent of type i while \mathbf{x}_1^i is her consumption at 'adult' age.

The model studied in the previous section can be easily adapted to handle this modification. For us to free ride on the results derived therein, just notice that we may define the following maximization problem and related indirect utility function,

$$\nu(\mathbf{q}, I) \equiv \max_{\mathbf{x}} u(\mathbf{x}_1) + u(\mathbf{x}_2) \text{ s.t. } \mathbf{q}\mathbf{x} \leq I,$$

where $\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2)$ and $\mathbf{x} = (\mathbf{x}_1', \mathbf{x}_2')'$.

Nothing in our proofs required the price vectors or the indirect utility functions to coincide. That is, we could redo all the same steps with

$$v(\mathbf{p}, y - s^*) - \zeta(Y) + \sum_{i=H,L} \pi^i [\nu(\mathbf{q}, y^i + s^*) - \zeta(Y^i/w^i)].$$

replacing (1) as the equilibrium expected utility and

$$v(\mathbf{p}, y - s^o) - \zeta(Y) + \nu(\mathbf{q}, y^L + s^o) - \sum_{i=H,L} \pi^i \zeta(Y^L/w^i).$$

as the relevant alternative.

Proposition 2, then, implies the following. First, if preferences are homothetic, consumption goods within periods are taxed uniformly. Second, consumption in the first period is taxed at a lower rate than consumption in the second period, but second period consumption is taxed at the same rate as retirement consumption, as in Atkinson and Stiglitz (1976).

The first part of the result is a straightforward consequence of conditions (12). The second part, however, is due to the fact that, once uncertainty is realized, and conditional on their preferred bundle, agents will not alter their consumption pattern as a consequence of deviating behavior.

5 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 goods and (implicitly) on capital income.

The simultaneous appearance of the following features 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, and;¹² finally, 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, imply that the set of implementable allocations, from a second period perspective, is 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 be for another. Rational agents anticipate their pattern of announcements conditional on realized types, and manipulate their preferences—in the $y \times Y$ space—by means of convenient choices of savings, to take advantage of the possibilities made available by the second period budget set.

Because off-equilibrium savings is crucial in defining the set of incentive compatible allocations, it will play a role in defining conditions under which IC constraints may be affected by commodity taxes. We show that non-observability of savings generates a violation of Atkinson and Stiglitz’ (1976) uniform taxation prescription. Homotheticity must be added to separability for the result to hold.

¹²This sensible restriction on tax instruments is also imposed in Cremer et al. (2003), where, however savings do not affect how agents rank bundles ex post.

Differential taxation of goods across time is used to emulate the non-existing taxation of capital income. We show that first period goods are to be taxed at a lower rate than second period goods, at the optimum. Implicitly, we are prescribing the taxation of savings (or subsidization of early consumption). From a purely theoretical perspective, this result resemble the inverse Euler equation result found in Golosov et al. (2003), Cremer and Gahvari (1995) and da Costa and Werning (2000).

Finally, we apply the model to the discussion of tax policy regarding retirement funds. We show that retirement funds are not to be taxed if investment is made after an agent has fully realized her earnings potential. This partially, and only partially, retrieves the famous result of no taxation of retirement funds due to Atkinson and Stiglitz (1976).

The model is admittedly simplistic, which means that policy prescriptions derived herein are to be considered with some caution. The point is not to portrair an accurate picture of the workings of a real economy but rather to emphasize what forces are at play and to what direction they drive optimal policies. Incorporating more periods (possibly infinite periods) and/or states (possibly a continuum of types) is not an easy task and, as of this moment, it seems that a computationally feasible method for doing it is still out of reach.¹³

A Appendix

A.1 The Relevant IC constraints

A.1.1 The Redundancy of Second Period IC constraints

Let $k = *, ^o, ^{oo}, ^{ooo}$, as in section 2.0.1, and define

$$\begin{aligned} U^k(i) &\equiv v(q, y^i + s^k) - \zeta(Y^i/w^i), \\ U^k(i|j) &\equiv v(q, y^i + s^k) - \zeta(Y^i/w^j), \text{ and} \\ U^k &\equiv v(p, y - s^k) - \zeta(Y), \end{aligned}$$

for $j, i = H, L$.

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

$$U^* + \sum_{i=H,L} \pi^i U^*(i) \geq U^o + \pi^H U^o(L|H) + \pi^L U^o(L), \quad (15)$$

$$U^* + \sum_{i=H,L} \pi^i U^*(i) \geq U^{oo} + \pi^H U^{oo}(H) + \pi^L U^{oo}(H|L), \text{ and} \quad (16)$$

$$U^* + \sum_{i=H,L} \pi^i U^*(i) \geq U^{ooo} + \pi^H U^{ooo}(L|H) + \pi^L U^{ooo}(H|L), \quad (17)$$

are satisfied, so are second period ones,

$$U^*(H) \geq U^*(L|H), \text{ and} \quad (18)$$

$$U^*(L) \geq U^*(H|L). \quad (19)$$

¹³See comment in footnote 7.

To see this, assume that (15) holds. Because savings are optimally chosen in (15), the right hand side of this equation is no less than the same expression evaluated at any level of savings. In particular,

$$\begin{aligned} U^* + \sum_{i=H,L} \pi^i U^*(i) &\geq U^* + \pi^H U^*(L|H) + \pi^L U^*(L) \\ \therefore U^*(H) &\geq U^*(L|H) \end{aligned}$$

That is, (18) also holds.

An analogous argument can be used to show that (16) implies (19). Hence, if the first period constraints are satisfied, so are the second period ones.

A.1.2 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 shall see next is that, if there is a level of savings that, at the same time, makes strategy σ^{ooo} optimal, and yields an expected utility at least as large as the one obtained with the adoption of the truthful announcement strategy, then, one of the other two constraints is violated.

In fact, assume that constraint (17) is binding at the optimal allocation. Then, there is a level of savings s^{ooo} such that

$$U^* + \sum_{i=H,L} \pi^i U^*(i) = U^{ooo} + \pi^H U^{ooo}(H|L) + \pi^L U^{ooo}(L|H).$$

Notice that, because this is the optimal strategy, agents find it in their best interest to make these announcements after their productivities are realized.

Therefore, it must be the case that

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

What the claim 2 shows however, is that if such a situation occurs, one of the other two first period IC constraints is violated. Therefore, whenever constraints (15) and (16) are satisfied, (17) is satisfied as a strict inequality. Hence, we may always leave it in the background.

Before presenting claim 2 we need the following result.

Claim 1 *Monotonicity, i.e., $(y^H, Y^H) \geq (y^L, Y^L)$, is necessary for an allocation to be implementable.*

Proof of Claim 1. Take an arbitrary choice of savings, and define the following set, for each allocation (y, Y) ,

$$\mathbb{Z}_{\sim}^H(y, Y) \equiv \{(y', Y') \in \mathbb{R}_+^2; (y', Y') \succsim_H (y, Y)\}.$$

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

$$\mathbb{Z}_{\sim}^L(y, Y) \equiv \{(y', Y') \in \mathbb{R}_+^2; (y', Y') \succsim_L (y, Y)\}$$

Define also

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

the set of bundles for which both quantities are at least as great as (y, Y) , with at least one entry strictly greater. Similarly,

$$\mathbb{Z}_-(y, Y) \equiv \{(y', Y') \in \mathbb{R}_+^2; (y', Y') < (y, Y)\},$$

is the set of bundles for which both quantities are no greater than (y, Y) , with at least one entry strictly smaller. We can see that single-crossing implies

$$\mathbb{Z}_{\sim}^L(y, Y) \cap \mathbb{Z}_+(y, Y) \subset \text{int}(\mathbb{Z}_{\sim}^H(y, Y)) \quad (20)$$

and

$$\mathbb{Z}_{\sim}^H(y, Y) \cap \mathbb{Z}_-(y, Y) \subset \text{int}(\mathbb{Z}_{\sim}^L(y, Y)) \quad (21)$$

where $\text{int}(A)$ is the interior of set A . Because preferences are strongly monotonic, $\text{int}(\mathbb{Z}_{\sim}^i(y, Y))$ is the set of bundles strictly preferred to (y, Y) . For an allocation to be implementable, it must be the case that, at the equilibrium choice of savings, $(y^i, Y^i) \in \mathbb{Z}_{\sim}^i(y^j, Y^j)$ $i, j = H, L$. Transitivity, naturally means that $(y, Y) \in \text{int}(\mathbb{Z}_{\sim}^i(y', Y'))$ then $(y', Y') \notin \mathbb{Z}_{\sim}^i(y, Y)$ while monotonicity of preferences implies that $(y, Y) + (\alpha, \beta) \in \text{int}(\mathbb{Z}_{\sim}^i(y^j, Y^j))$ $i, j = H, L$, for all $\alpha \geq 0, \beta \leq 0$ (with at least one strict) and $(y, Y) + (\alpha, \beta) \notin \mathbb{Z}_{\sim}^i(y^j, Y^j)$ $i, j = H, L$, for all $\alpha \leq 0, \beta \geq 0$ (with at least one strict).

Now assume that $(y^L, Y^L) \in \mathbb{Z}_+(y^H, Y^H)$ and $(y^L, Y^L) \in \mathbb{Z}_{\sim}^L(y^H, Y^H)$. Naturally, $(y^H, Y^H) \notin \mathbb{Z}_+(y^L, Y^L)$ which implies that either $(y^H, Y^H) \in \mathbb{Z}_-(y^L, Y^L)$, $(y^H, Y^H) = (y^L, Y^L) + (\alpha, \beta)$ where, either $\alpha \geq 0$ and $\beta \leq 0$ (with at least one strict) or $\alpha \leq 0, \beta \geq 0$ (with at least one strict). The last two options can be ruled out on the grounds that the latter does not belong in $\mathbb{Z}_{\sim}^H(y^L, Y^L)$ while, under the former, $(y^L, Y^L) \notin \mathbb{Z}_{\sim}^L(y^H, Y^H)$, contradicting our initial assumption. Therefore, $(y^H, Y^H) \in \mathbb{Z}_-(y^L, Y^L)$. But, in this case, $(y^H, Y^H) \in \mathbb{Z}_{\sim}^H(y^L, Y^L) \cap \mathbb{Z}_-(y, Y)$, which implies from (21) that $(y^H, Y^H) \subset \text{int}(\mathbb{Z}_{\sim}^L(y, Y))$ contradicting $(y^L, Y^L) \in \mathbb{Z}_{\sim}^L(y^H, Y^H)$. ■

Claim 2 *If there is a savings choice s^{ooo} such that*

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

then constraints (19) and (18) are not satisfied.

Proof of Claim 2. Let $\ddot{\mathbb{Z}}_{\sim}^i(y, Y)$ be as in the proof of claim 1 but defined for the level of savings s^{ooo} , while $\hat{\mathbb{Z}}_{\sim}^i(y, Y)$ denotes the same sets at $s = s^*$. Let $(y^H, Y^H) \in \ddot{\mathbb{Z}}_{\sim}^L(y^L, Y^L)$ and $(y^L, Y^L) \in \ddot{\mathbb{Z}}_{\sim}^H(y^H, Y^H)$. Now, assume that $(y^H, Y^H) \in \mathbb{Z}_+(y^L, Y^L)$. Because (20) was derived for an arbitrary choice of savings, it must be the case that

$(y^H, Y^H) \in \text{int} \left(\ddot{\mathbb{Z}}_{\sim}^H (y^L, Y^L) \right)$ according to (20). This, however, contradicts $(y^L, Y^L) \in \ddot{\mathbb{Z}}_{\sim}^H (y^H, Y^H)$ and implies $(y^H, Y^H) \in \mathbb{Z}_- (y^L, Y^L)$. Claim 1 then guarantees that the allocation is not implementable. ■

We are down to two first period IC constraints, (15) and (16), associated, respectively, to the strategies of always announcing to be of a low type and always announcing to be of a high type. The next claim shows that absent commodity taxes, only constraint (15) binds at the optimum.

Claim 3 *Absent commodity taxes, only IC constraint (15) binds at the optimum.*

Proof of Claim 3. First note that the government cannot gain from raising revenue from the first period and sending to the second period. All that matters for an agent when deciding what strategy to follow is the present value of income and the timing of labor supply. Hence, we can consider the case where agents are not taxed in the first period, without loss in generality.

Then, assume that the optimal allocation is such that $Y^H < y^H$ and $Y^L > y^L$. 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 producing Y^H and Y^L 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.

Next consider the case where $Y^H \geq y^H$ and $Y^L \leq y^L$. 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):

$$\begin{aligned} [u(Y - s^*) - \zeta(Y)] + \sum_{i=H,L} \pi^i \left[u(y^i + s^*) - \zeta\left(\frac{Y^i}{w^i}\right) \right] = \\ [u(Y - s^o) - \zeta(Y)] + u(y^H + s^o) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y^H}{w^i}\right) \leq \\ [u(Y - s^o) - \zeta(Y)] + u(Y^H + s^o) - \sum_{i=H,L} \pi^i \zeta\left(\frac{Y^H}{w^i}\right), \end{aligned}$$

with strict inequality if $Y^H > y^H$.

But, if $Y^H = y^H$ the allocation is feasible in autarchy and, in general, non-optimal. Once again, the government policy lowers utility when compared to what can be attained in autarchy. ■

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 (15) that binds at the optimum. This guarantees that commodity taxes are introduced to relax IC constraint (15). Second, once commodity taxes are introduced some indirect effects defined over conditions on demands that 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.¹⁴

We end this discussion with the following claim.

Claim 4 *If leisure is normal, then $s^o > s^*$.*

Proof of Claim 4. Let

$$\mathbb{Z}_+^{H*}(y^H, Y^H) \equiv \{(y', Y') \in \mathbb{R}_+^2; (y', Y') \succ_{H*} (y^H, Y^H)\},$$

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

$$\mathbb{Z}_+^{Ho}(y^H, Y^H) \equiv \{(y', Y') \in \mathbb{R}_+^2; (y', Y') \succ_{Ho} (y^H, Y^H)\},$$

at the optimal level of savings for strategy σ^o , s^o .

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

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

A contradiction ■

A.2 Price variations and Savings

We cannot understand what is necessary for conditions (10) and (11) to be satisfied without, first, understanding \hat{s}_j . Let us, then, explore the its meaning.

Consider the following problem:

$$\max_s \left\{ v(\mathbf{p}, y - s) + \sum_{i=H,L} \pi^i v(\mathbf{q}, y^i + s) \right\} \quad (22)$$

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

$$-v_y + \left[\sum_{i=H,L} \pi^i v_y^*(i) \right] = 0,$$

and

$$\Theta \equiv -v_{yy} - \sum_{i=H,L} \pi^i v_{yy}^*(i) > 0.$$

¹⁴Nor do we know of this having been done in any other place. In fact, Lagrangian multipliers of incentive compatibility constraints 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.

We may then use this to find $\partial s / \partial p_j$:

$$\Theta \frac{ds}{dp_j} = -v_{yj} = (v_{yy}x^j + v_yx_y^j)$$

and $\partial s / \partial q_j$:

$$\Theta \frac{ds}{dq_j} = \sum_{i=H,L} \pi^i v_{yj}^*(i) = - \sum_{i=H,L} \pi^i [v_{yy}^*(i) x^j(i) + v_y(i) x_y^j(i)]$$

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

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

$$\frac{ds}{dy} = \frac{v_{yy}}{\Theta}, \quad \frac{ds}{dy^H} = \frac{\pi^H v_{yy}(H)}{\Theta}, \quad \text{and} \quad \frac{ds}{dy^L} = \frac{\pi^L v_{yy}(L)}{\Theta}, \quad \text{respectively.}$$

Hence,

$$\hat{s}_j = s_j + s_y x^j = \frac{v_{yy}x^j + v_yx_y^j - v_{yy}x^j}{\Theta} = \frac{v_yx_y^j}{\Theta}$$

and

$$\begin{aligned} \tilde{s}_j &= s_j + \sum_{i=H,L} \pi^i s_{y^i} x^{*j}(i) \\ &= - \frac{\sum_{i=H,L} \pi^i [v_{yy}^*(i) x^j(i) + v_y(i) x_y^j(i)]}{\Theta} + \frac{\sum_{i=H,L} \pi^i v_{yy}(i) x^j(i)}{\Theta} \\ &= - \frac{\sum_{i=H,L} \pi^i v_y(i) x_y^j(i)}{\Theta} \end{aligned}$$

which gives (12).

References

- [1] Atkinson A. and J. Stiglitz "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6, (1976) 55-75.
- [2] Brunner, J. "A Note on the Optimum Income Tax" *Journal of Public Economics* 50, (1993) 445-451.
- [3] Chiappori, P.-A., I. Macho, P. Rey and B. Salani , "Repeated Moral Hazard: the Role of Memory, Commitment and the Access to Credit Markets", *European Economic Review*, 38 (1994) 1527-53.
- [4] Cooter, R. "Optimal Tax Schedules and Rates: Mirrlees and Ramsey" *The American Economic Review*, 68 (1978) 756-768.
- [5] Cremer, H and F. Gahvari, "Uncertainty, Optimal Taxation and the Direct Versus Indirect Tax Controversy." *The Economic Journal* 105, (1995) 1165-1179

- [6] Cremer, H. and F. Gahvari, "Uncertainty, Commitment and Optimal Taxation." *Journal of Public Economic Theory*, 1 (1999) 51-70
- [7] Cremer, H., Pestieau, P. and J.-C. Rochet "Capital Income Taxation when Inherited Wealth is not Observable" *Journal of Public Economics*, 87: (2003) 2475-2490.
- [8] 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.
- [9] da Costa, C. and I. Werning, "A Note on Pre-Committed Goods." mimeo., University of Chicago, 2000.
- [10] Dewatripont, M. "Commitment through Renegotiation-proof Contracts with Third Parties." *Review of Economic Studies* 55, (1988) 377-90.
- [11] Ebert, U. "A Reexamination of the Optimal Non-linear Income Tax" *Journal of Public Economics*, 49 (1992): 47-73.
- [12] Fernandes, A. and C. Phelan, "A recursive formulation for repeated agency with history dependence." *Journal of Economic Theory* 99, (2000) 223-247.
- [13] Guesnerie, R. "A Contribution to the Pure Theory of Taxation" (Cambridge: Cambridge University Press, 1995)
- [14] Golosov, M., N. Kocherlakota, and A. Tsyvinski "Optimal Indirect and Capital Taxation" *Review of Economic Studies* 70, (2003): 569-587.
- [15] Golosov, M., and A. Tsyvinski (2004) "Optimal Taxation with Endogenous Insurance Markets" Mimeo, UCLA.
- [16] Kocherlakota, N. "Wedges and Taxes" *American Economic Review Papers and Proceedings*, (2004) forthcoming.
- [17] Kocherlakota, N. "Figuring out the impact of hidden savings on optimal unemployment insurance" *Review of Economic Dynamics* 7 (2004) 541-554
- [18] Mirrlees, J. "An Exploration on the Theory of Optimum Income Taxation" *Review of Economic Studies* 38 (1971) 175-208.
- [19] Mirrlees, J. "Optimal Tax Theory: A Synthesis," *Journal of Public Economics*, 6 (1976) 327-58.
- [20] 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.
- [21] Rogerson, W. "Repeated Moral Hazard" *Econometrica*, 58 (1985) 69-76.

- [22] Saez, E. “The Desirability of Commodity Taxes with Non-linear Income Taxation and Heterogeneous Tastes” *Journal of Public Economics*, 83 (2002) 217-230.
- [23] Stiglitz, J. E. “Self-selection and Pareto Efficient Taxation,” *Journal of Public Economics*, 17 (1982) 213-240.

Últimos Ensaio Econômicos da EPGE

- [570] Rubens Penha Cysne. *On the Statistical Estimation of Diffusion Processes – A Partial Survey (Revised Version, Forthcoming Brazilian Review of Econometrics)*. Ensaio Econômicos da EPGE 570, EPGE-FGV, Out 2004.
- [571] Aloisio Pessoa de Araújo, Luciano I. de Castro Filho, e Humberto Luiz Ataíde Moreira. *Pure strategy equilibria of multidimensional and Non-monotonic auctions*. Ensaio Econômicos da EPGE 571, EPGE-FGV, Nov 2004.
- [572] Rubens Penha Cysne e Paulo César Coimbra Lisboa. *Imposto Inflacionário e Transferências Inflacionárias no Mercosul e nos Estados Unidos*. Ensaio Econômicos da EPGE 572, EPGE-FGV, Nov 2004.
- [573] Renato Galvão Flôres Junior. *Os desafios da integração legal*. Ensaio Econômicos da EPGE 573, EPGE-FGV, Dez 2004.
- [574] Gustavo M. de Athayde e Renato Galvão Flôres Junior. *Do Higher Moments Really Matter in Portfolio Choice?*. Ensaio Econômicos da EPGE 574, EPGE-FGV, Dez 2004.
- [575] Germán Calfat e Renato Galvão Flôres Junior. *The EU–Mercosul free trade agreement: Quantifying mutual gains*. Ensaio Econômicos da EPGE 575, EPGE-FGV, Dez 2004.
- [576] Andrew W. Horowitz e Renato Galvão Flôres Junior. *Beyond indifferent players: On the existence of Prisoners Dilemmas in games with amicable and adversarial preferences*. Ensaio Econômicos da EPGE 576, EPGE-FGV, Dez 2004.
- [577] Rubens Penha Cysne. *Is There a Price Puzzle in Brazil? An Application of Bias–Corrected Bootstrap*. Ensaio Econômicos da EPGE 577, EPGE-FGV, Dez 2004.
- [578] Fernando de Holanda Barbosa, Alexandre Barros da Cunha, e Elvia Mureb Salum. *Competitive Equilibrium Hyperinflation under Rational Expectations*. Ensaio Econômicos da EPGE 578, EPGE-FGV, Jan 2005.
- [579] Rubens Penha Cysne. *Public Debt Indexation and Denomination, The Case of Brazil: A Comment*. Ensaio Econômicos da EPGE 579, EPGE-FGV, Mar 2005.
- [580] Gina E. Acosta Rojas, Germán Calfat, e Renato Galvão Flôres Junior. *Trade and Infrastructure: evidences from the Andean Community*. Ensaio Econômicos da EPGE 580, EPGE-FGV, Mar 2005.

- [581] Edmundo Maia de Oliveira Ribeiro e Fernando de Holanda Barbosa. *A Demanda de Reservas Bancárias no Brasil*. Ensaios Econômicos da EPGE 581, EPGE–FGV, Mar 2005.
- [582] Fernando de Holanda Barbosa. *A Paridade do Poder de Compra: Existe um Quebra-Cabeça?*. Ensaios Econômicos da EPGE 582, EPGE–FGV, Mar 2005.
- [583] Fabio Araujo, João Victor Issler, e Marcelo Fernandes. *Estimating the Stochastic Discount Factor without a Utility Function*. Ensaios Econômicos da EPGE 583, EPGE–FGV, Mar 2005.
- [584] Rubens Penha Cysne. *What Happens After the Central Bank of Brazil Increases the Target Interbank Rate by 1%?*. Ensaios Econômicos da EPGE 584, EPGE–FGV, Mar 2005.
- [585] GUSTAVO GONZAGA, Naércio Menezes Filho, e Maria Cristina Trindade Terra. *Trade Liberalization and the Evolution of Skill Earnings Differentials in Brazil*. Ensaios Econômicos da EPGE 585, EPGE–FGV, Abr 2005.
- [586] Rubens Penha Cysne. *Equity–Premium Puzzle: Evidence From Brazilian Data*. Ensaios Econômicos da EPGE 586, EPGE–FGV, Abr 2005.
- [587] Luiz Renato Regis de Oliveira Lima e Andrei Simonassi. *Dinâmica Não–Linear e Sustentabilidade da Dívida Pública Brasileira*. Ensaios Econômicos da EPGE 587, EPGE–FGV, Abr 2005.
- [588] Maria Cristina Trindade Terra e Ana Lucia Vahia de Abreu. *Purchasing Power Parity: The Choice of Price Index*. Ensaios Econômicos da EPGE 588, EPGE–FGV, Abr 2005.
- [589] Osmani Teixeira de Carvalho Guillén, João Victor Issler, e George Athanasopoulos. *Forecasting Accuracy and Estimation Uncertainty using VAR Models with Short– and Long–Term Economic Restrictions: A Monte–Carlo Study*. Ensaios Econômicos da EPGE 589, EPGE–FGV, Abr 2005.
- [590] Pedro Cavalcanti Gomes Ferreira e Samuel de Abreu Pessôa. *The Effects of Longevity and Distortions on Education and Retirement*. Ensaios Econômicos da EPGE 590, EPGE–FGV, Jun 2005.
- [591] Fernando de Holanda Barbosa. *The Contagion Effect of Public Debt on Monetary Policy: The Brazilian Experience*. Ensaios Econômicos da EPGE 591, EPGE–FGV, Jun 2005.
- [592] Rubens Penha Cysne. *An Overview of Some Historical Brazilian Macroeconomic Series and Some Open Questions*. Ensaios Econômicos da EPGE 592, EPGE–FGV, Jun 2005.
- [593] Luiz Renato Regis de Oliveira Lima e Raquel Menezes Bezerra Sampaio. *The Asymmetric Behavior of the U.S. Public Debt..* Ensaios Econômicos da EPGE 593, EPGE–FGV, Jul 2005.

- [594] Pedro Cavalcanti Gomes Ferreira, Roberto de Góes Ellery Junior, e Victor Gomes. *Produtividade Agregada Brasileira (1970–2000): declínio robusto e fraca recuperação*. Ensaios Econômicos da EPGE 594, EPGE–FGV, Jul 2005.